



Universidad
de Alcalá

COMISIÓN DE ESTUDIOS OFICIALES
DE POSGRADO Y DOCTORADO

ACTA DE EVALUACIÓN DE LA TESIS DOCTORAL

Año académico 2018/19

DOCTORANDO: **DELGADO AZUARA, FRANCISCO DE ASÍS**
D.N.I./PASAPORTE: ****3848E

PROGRAMA DE DOCTORADO: **D442-INGENIERÍA DE LA INFORMACIÓN Y DEL CONOCIMIENTO**
DPTO. COORDINADOR DEL PROGRAMA: **CIENCIAS DE LA COMPUTACIÓN**
TITULACIÓN DE DOCTOR EN: **DOCTOR/A POR LA UNIVERSIDAD DE ALCALÁ**

En el día de hoy 08/04/19, reunido el tribunal de evaluación nombrado por la Comisión de Estudios Oficiales de Posgrado y Doctorado de la Universidad y constituido por los miembros que suscriben la presente Acta, el aspirante defendió su Tesis Doctoral, elaborada bajo la dirección de **JOSÉ RAMÓN HILERA GONZÁLEZ // RAUL JULIAN RUGGIA FRICK**.

Sobre el siguiente tema: *PROPUESTA DE UN ESTANDAR PARA SOLUCIONAR LOS PROBLEMAS DE INTEROPERABILIDAD SEMÁNTICA ENTRE SISTEMAS DE TRATAMIENTO DE DATOS DE SEGURIDAD SOCIAL*

Finalizada la defensa y discusión de la tesis, el tribunal acordó otorgar la CALIFICACIÓN GLOBAL¹ de (no apto, aprobado, notable y sobresaliente): SOBRESALIENTE

Alcalá de Henares, 8 de Abril de 2019

EL PRESIDENTE

Fdo.: JOSE JAVIER MARTINEZ HERRAIZ
MUÑOZ-REJA

EL SECRETARIO

Fdo.: MARIA TERESA VILLALBA DE BENITO

EL VOCAL

Fdo.: ISMAEL CABALLERO

Con fecha 24 de abril de 2019 la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado, a la vista de los votos emitidos de manera anónima por el tribunal que ha juzgado la tesis, resuelve:

- Conceder la Mención de "Cum Laude"
 No conceder la Mención de "Cum Laude"

La Secretaria de la Comisión Delegada

FIRMA DEL ALUMNO,

Fdo.: DELGADO AZUARA, FRANCISCO DE ASÍS

¹ La calificación podrá ser "no apto" "aprobado" "notable" y "sobresaliente". El tribunal podrá otorgar la mención de "cum laude" si la calificación global es de sobresaliente y se emite en tal sentido el voto secreto positivo por unanimidad.

INCIDENCIAS / OBSERVACIONES:



Universidad
de Alcalá

COMISIÓN DE ESTUDIOS OFICIALES
DE POSGRADO Y DOCTORADO

En aplicación del art. 14.7 del RD. 99/2011 y el art. 14 del Reglamento de Elaboración, Autorización y Defensa de la Tesis Doctoral, la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado y Doctorado, en sesión pública de fecha 24 de abril, procedió al escrutinio de los votos emitidos por los miembros del tribunal de la tesis defendida por *DELGADO AZUARA, FRANCISCO DE ASÍS*, el día 8 de abril de 2019, titulada *PROPUESTA DE UN ESTANDAR PARA SOLUCIONAR LOS PROBLEMAS DE INTEROPERABILIDAD SEMÁNTICA ENTRE SISTEMAS DE TRATAMIENTO DE DATOS DE SEGURIDAD SOCIAL*, para determinar, si a la misma, se le concede la mención "cum laude", arrojando como resultado el voto favorable de todos los miembros del tribunal.

Por lo tanto, la Comisión de Estudios Oficiales de Posgrado **resuelve otorgar** a dicha tesis la

MENCIÓN "CUM LAUDE"

Alcalá de Henares, 24 de abril de 2019

EL VICERRECTOR DE INVESTIGACIÓN Y TRANSFERENCIA



F. Javier de la Mata

F. Javier de la Mata de la Mata

Copia por e-mail a:

Doctorando: DELGADO AZUARA, FRANCISCO DE ASÍS

Secretario del Tribunal: MARIA TERESA VILLALBA DE BENITO

Directores de Tesis: JOSÉ RAMÓN HILERA GONZÁLEZ // RAUL JULIAN RUGGIA FRICK





Universidad
de Alcalá

ESCUELA DE DOCTORADO
Servicio de Estudios Oficiales de
Posgrado

DILIGENCIA DE DEPÓSITO DE TESIS.

Comprobado que el expediente académico de D./D^a _____
reúne los requisitos exigidos para la presentación de la Tesis, de acuerdo a la normativa vigente, y habiendo
presentado la misma en formato: soporte electrónico impreso en papel, para el depósito de la
misma, en el Servicio de Estudios Oficiales de Posgrado, con el nº de páginas: _____ se procede, con
fecha de hoy a registrar el depósito de la tesis.

Alcalá de Henares a _____ de _____ de 20 _____



Fdo. El Funcionario



Universidad de Alcalá

Programa de Doctorado

Ingeniería de la Información y del Conocimiento

Departamento de Ciencias de la Computación

TESIS DOCTORAL

Propuesta de un estándar para solucionar los problemas de interoperabilidad semántica entre sistemas de tratamiento de datos de Seguridad Social

Autor:

D. Francisco de Asís Delgado Azuara

Directores:

Dr. D. José Ramón Hilera González

Dr. D. Raul K. Ruggia Frick

2019



Universidad
de Alcalá

DPTO. DE CIENCIAS DE LA COMPUTACIÓN

Escuela Politécnica Superior

Campus Universitario. Edificio Politécnico

28871 Alcalá de Henares (Madrid) España

Teléfonos: +34 91 885 6645

Fax: +34 91 885 6646

e-mail: dpto.ccomputacion@uah.es

Dr. D. José Javier Martínez Herraiz, Profesor Titular de Universidad del Área de Ciencias de la Computación e Inteligencia Artificial, en calidad de Coordinador del Programa de Doctorado en Ingeniería de la Información y del Conocimiento de la Universidad de Alcalá.

CERTIFICO: Que la Comisión Académica del Programa ha aprobado la presentación de la Tesis Doctoral titulada “PROPUESTA DE UN ESTÁNDAR PARA SOLUCIONAR LOS PROBLEMAS DE INTEROPERABILIDAD SEMÁNTICA ENTRE SISTEMAS DE TRATAMIENTO DE DATOS DE SEGURIDAD SOCIAL” realizada por D. FRANCISCO DE ASIS DELGADO AZUARA, dirigida por el Dr. D. José Ramón Hílera González y co-dirigida por el Dr. D. Raúl Julián Ruggia Frick.

La Tesis Doctoral reúne los requisitos científicos de originalidad y rigor metodológicos para ser defendida ante un tribunal. Esta Comisión ha tenido también en cuenta la evaluación positiva anual del doctorando, habiendo obtenido las correspondientes competencias establecidas en el Programa.

Y para que así conste, firmo la presente en Alcalá de Henares, a 16 de enero de 2019.



Dr. José Javier Martínez Herraiz

Dr. D. José Ramón Hilera González, Catedrático de Universidad del Área de Lenguajes y Sistemas Informáticos del Departamento de Ciencias de la Computación de la Universidad de Alcalá.

Dr. D. Raúl Julián Ruggia Frick, Profesor de la Universidad de la República (Uruguay).

HACEN CONSTAR:

Que, una vez concluido el trabajo de tesis doctoral titulado: “PROPUESTA DE UN ESTÁNDAR PARA SOLUCIONAR LOS PROBLEMAS DE INTEROPERABILIDAD SEMÁNTICA ENTRE SISTEMAS DE TRATAMIENTO DE DATOS DE SEGURIDAD SOCIAL” realizado por D. FRANCISCO DE ASIS DELGADO AZUARA, dicho trabajo tiene suficientes méritos teóricos, que se han contrastado adecuadamente mediante validaciones experimentales y que son altamente novedosos. Por todo ello consideran que procede su defensa pública.

Y para que así conste, firman la presente, a 20 de diciembre de 2018.

El Director de la Tesis

El Codirector de la Tesis



Dr. José Ramón Hilera González



Dr. Raúl Julián Ruggia Frick

A mi esposa, Cristina

A mi hija, Cris

Agradecimientos

A los directores de la tesis, por sus orientaciones y paciencia en el largo recorrido de este trabajo.

A José Ángel Martínez Luque, que con su trabajo fin de carrera proporcionó la primera prueba del contenido fundamental de la tesis.

A Marlon Santiago Viñan Ludeña, que con su trabajo fin de Master sobre “Interoperabilidad en instituciones de la seguridad social” construyó la primera arquitectura web en la que fueron intercambiados los metadatos estandarizados propuestos en esta tesis.

Al Grupo de Investigación TIFYC de la Universidad de Alcalá, por la creación e instalación de los vocabularios de apoyo al estándar propuesto, publicados en forma de *dataset* con el nombre de “socialsecurity”.

A mi mujer Cristina, por su revisión exhaustiva y sus consejos para mejorar la presentación de esta tesis.

A mi hija Cris, por su habilidad en mejorar las imágenes que se resistían.

Resumen

Los movimientos de población, especialmente los movimientos migratorios, son un fenómeno de gran impacto social y económico. Garantizar el ejercicio de los derechos en materia de Seguridad Social de las personas que se mueven por el mundo exige la puesta en marcha de una serie de acuerdos entre las instituciones competentes en diferentes países. El que esos acuerdos tengan una efectividad real exige que se realicen intercambios de datos entre sistemas de información totalmente heterogéneos. En ese contexto, las técnicas de interoperabilidad juegan un papel clave para que estos sistemas sean capaces de entenderse.

La presente tesis va a tratar los problemas de interoperabilidad en este ámbito, muy especialmente los relacionados con los aspectos semánticos, proponiéndose una solución basada en la estandarización de un conjunto de metadatos que acompañen a cada uno de los datos intercambiados con el fin de evitar cualquier ambigüedad o interpretación equívoca.

Como punto de partida para la propuesta de estándar se ha elegido *Dublin Core*, por ajustarse especialmente a las necesidades detectadas y por su amplia implantación en todos los ámbitos, incluido el de la administración pública. Tal como se ha hecho en otros dominios, se ha adaptado este conjunto de elementos, modificando varios de ellos y añadiendo tres extensiones.

Un aspecto importante es que el conjunto de metadatos se basa en una serie de vocabularios controlados que permiten definir y acotar sus posibles valores. Dichos vocabularios tienen diferente grado de desarrollo. Su revisión y extensión se propone como futura línea de investigación que dé continuidad a lo propuesto en esta tesis.

Abstract

Population movements, especially migratory movements, are a phenomenon of great social and economic impact. Guaranteeing the exercise of Social Security rights for people moving around the world requires the implementation of a series of agreements between the competent institutions in different countries. The fact that these agreements have a real effectiveness requires the exchange of data between completely heterogeneous information systems. In this context, interoperability techniques play a key role for these systems to be able to understand each other.

This thesis will address the interoperability problems that arise in this area, especially those related to semantic aspects, proposing a solution based on the standardization of a set of metadata that accompany each of the data exchanged in order to avoid any ambiguity or equivocal interpretation.

As a starting point for the proposed standard, *Dublin Core* has been chosen, as it is specially adjusted to the needs detected and its wide implementation in all areas, including the public administration. As has been done in other domains, this set of elements has been adapted, modifying several of them and adding three extensions.

An important aspect is that the metadata set is based on a series of controlled vocabularies that allow defining and delimiting their possible values. These vocabularies have a different degree of development. Its revision and extension is proposed as a future line of research that gives continuity to what is proposed in this thesis.

Índice del documento

Capítulo 1. Introducción	17
1.1 Contexto general de la tesis.....	18
1.1.1 Los movimientos migratorios	18
1.1.2 Trabajadores migrantes y derechos en materia de Seguridad Social	20
1.1.3 Recomendaciones y reglamentos internacionales.....	21
1.1.4 Acuerdos internacionales sobre Seguridad Social.....	22
1.1.5 Acuerdos internacionales existentes en materia de Seguridad Social.....	23
1.1.6 Experiencias existentes en la utilización de las tecnologías de la información como apoyo a los intercambios de datos en materia de Seguridad Social.....	28
1.2 Contexto tecnológico	36
1.2.1 Situación de los intercambios en materia de Seguridad Social.....	36
1.2.2 Interoperabilidad entre sistemas de información de Seguridad Social	37
1.2.3 Problemas semánticos en los intercambios de datos de Seguridad Social....	42
1.3 Hipótesis y objetivos de la tesis doctoral	46
1.4 Organización del trabajo.....	47
Capítulo 2. Estado del Arte	49
2.1 Interoperabilidad	51
2.1.1 Definiciones	51
2.1.2 Modelos de madurez de interoperabilidad.....	53
2.2 La interoperabilidad en las Administraciones Públicas	61
2.2.1 Modelos de interoperabilidad en las Administraciones Públicas	61
2.2.2 La interoperabilidad en las Administraciones Públicas: casos y referencias .	66
2.3 La interoperabilidad entre las organizaciones de Seguridad Social	70
2.3.1 Modelo de aplicación de interoperabilidad para la Seguridad Social.....	70
2.3.2 Casos de interoperabilidad en la Seguridad Social.....	76
2.3.2.1 Programas de protección social y servicios sociales	77
2.3.2.1 Sistemas integrados de seguro de salud	78
2.3.2.2 Recaudación de cotizaciones de Seguridad Social y control del cumplimiento de las obligaciones	79
2.3.2.3 Mejora de la eficiencia institucional, de la calidad de la información y los servicios	

2.3.2.4	Acuerdos internacionales de Seguridad Social	80
2.3.3	Dimensiones de interoperabilidad en la Seguridad Social: situación actual..	81
2.3.3.1	Interoperabilidad política y legal.....	81
2.3.3.2	Dimensión organizativa	82
2.3.3.3	Dimensión semántica	89
2.3.3.4	Dimensión técnica	101
2.4	Vocabularios	104
2.4.1	Conceptos generales	104
2.4.2	Tipos de Vocabularios Controlados	105
2.4.3	Vocabularios controlados en la Unión Europea	106
2.5	Metadatos.....	113
2.5.1	Definición.....	113
2.5.2	Historia de los metadatos.....	114
2.5.3	Tipos de metadatos	116
2.5.4	Modelo de madurez de metadatos	117
2.6	Estándares.....	118
2.6.1	AGLS.....	118
2.6.2	<i>DUBLIN CORE</i>	118
2.6.3	Esquema Nacional de Interoperabilidad de España.....	119
2.6.4	HL7 (ISO 21090)	120
2.6.5	IEEE/LOM: ISO/ISO 19788	122
2.6.6	ISO/IEC 11179	122
2.6.7	INTEROPS	122
2.6.8	OMG (<i>Object Management Group</i>).....	123
2.7	Análisis crítico de la situación	123
2.7.1	La interoperabilidad	123
2.7.2	Modelos de interoperabilidad.....	124
2.7.3	La interoperabilidad en los distintos escenarios de la Seguridad Social.....	125
2.7.4	Situación de las distintas dimensiones de interoperabilidad.....	127
2.7.4.1	Situación de la dimensión política y legal	127
2.7.4.2	Situación organizacional.....	127

2.7.4.3	Situación técnica	128
2.7.4.4	Situación semántica	129
2.7.5	Análisis de vocabularios existentes	130
2.7.5.1	Build 6 del proyecto TESS	130
2.7.5.2	<i>Core Vocabularies</i>	131
2.7.6	Análisis de los estándares de metadatos	131
2.7.7	Situación de las principales experiencias de intercambios de datos de Seguridad Social en lo que a interoperabilidad se refiere	132
Capítulo 3.	Propuesta.....	134
3.1	Razones para una nueva especificación: recapitulación del problema.....	135
3.1.1	Alcance del problema	135
3.1.2	Automatización de los intercambios	135
3.1.3	Concepto de “dato”	137
3.2	Desarrollo de la solución.....	138
3.2.1	Solución propuesta.....	138
3.2.2	Elección de un estándar	139
3.2.3	Especificación propuesta: ESSIM	143
3.2.4	Elementos de ESSIM	145
3.2.4.1	<i>Creator Institution</i>	147
3.2.4.2	<i>Data Set</i>	148
3.2.4.3	<i>Data Term</i>	150
3.2.4.4	<i>Date Created</i>	151
3.2.4.5	<i>Date Modified</i>	151
3.2.4.6	<i>Description</i>	152
3.2.4.7	<i>Format</i>	153
3.2.4.8	<i>Identifier</i>	154
3.2.4.9	<i>Language</i>	155
3.2.4.10	<i>Mandate</i>	155
3.2.4.11	<i>References</i>	156
3.2.4.12	<i>Relation with Terms</i>	157
3.2.4.13	<i>Sender Institution</i>	158
3.2.4.14	<i>Social Security Category</i>	159

3.2.5	Vocabularios controlados utilizados por ESSIM	161
3.2.5.1	Literales	161
3.2.5.2	Glosario de Términos	161
3.2.5.3	<i>Institution Repository</i>	162
3.2.5.4	<i>Core Vocabularies</i>	164
3.2.5.5	ISO 639	164
3.2.5.6	EUR-Lex.....	165
3.2.5.7	<i>Official Documents</i>	167
3.2.6	Publicación de los vocabularios creados	168
Capítulo 4.	Demostración	170
4.1	Introducción.....	171
4.2	Tecnologías utilizadas	172
4.2.1	La Web semántica	172
4.2.2	<i>Resource Description Framework (RDF)</i>	174
4.2.3	Servicios web	175
4.3	Representación de los vocabularios con RDF	177
4.4	Prueba 1: desarrollo de una aplicación de creación de metadatos	188
4.5	Prueba 2: arquitectura basada en servicios web.....	199
4.5.1	Estudio de caso	199
4.5.2	Esquema de la arquitectura	203
4.5.3	Servicios web desarrollados	204
4.5.4	Descripción de los servicios web	206
4.5.4.1	Servicio WS1: Verificación de acuerdo en vigor	206
4.5.4.2	Servicio WS2: Creación de metadatos	209
4.5.4.3	Servicio WS3: Validación de metadatos.....	215
4.5.4.4	Servicio WS4: Solicitud de trámite administrativo.....	222
4.5.4.5	Servicio WS5: Respuesta a un trámite	226
4.6	Conclusiones de las pruebas.....	229
Capítulo 5.	Conclusiones y trabajos futuros	231
5.1	Recapitulación de objetivos.....	232
5.2	Aportaciones originales	233

5.3	Líneas de trabajo futuro.....	235
5.3.1	Ámbito profesional.....	235
5.3.2	Proyectos de investigación.....	236
	Referencias.....	239
	Anexos.....	255
	Anexo 1: propuesta de estandar.....	256
	Anexo 2: textos completos referenciados.....	271
	AC 370/12REV.....	272
	AC 12-446es.....	275
	AC 522/13REV.....	281
	AC 738/15.....	286
	EC (1997).....	291
	Pekka, R. (2013).....	296
	Nota TF 063/07.....	306

Índice de figuras

Figura 1. Afiliados extranjeros en España (Activa, 2018)	19
Figura 2. Esquema de SIBERO. Fuente: OISS	33
Figura 3. Modelo conceptual de interoperabilidad. EIF (2007).....	39
Figura 4. Arquitectura de información propuesta por el DMG (Pekka, 2013)	44
Figura 5. Bases del modelo de interoperabilidad LISI (Bueno de la Fuente, 2018)	55
Figura 6. Modelo de referencia LISI (LISI, 2018).....	55
Figura 7. Niveles de interoperabilidad LISI (Morris et al., 2004).....	57
Figura 8. Relación entre los niveles de interoperabilidad LISI (LISI, 2018).....	63
Figura 9. Relación entre los modelos LCI y LISI (Clark, 1999).....	65
Figura 10. Capas de <i>Coalition Interoperability</i> (LCI, 2003)	66
Figura 11. Modelo conceptual de Interoperabilidad para la Seguridad Social (ISSA, 2012) 72	
Figura 12. <i>Code List Database</i> (PAI, 2015).....	85
Figura 13. Ejemplo de consulta a la <i>Code List Database</i>	86
Figura 14. Arquitectura del directorio maestro EESSI (Delgado et al., 2012)	87
Figura 15. Diagrama de flujo de documentos del proyecto (TF, 2007).....	94
Figura 16. Formulario E-106 en su versión de 2007	96
Figura 17. Modelo de datos relacionados con la persona y su institución (TF, 2007)	100
Figura 18. Modelo en capas de NZ e-GIF para la interoperabilidad técnica	103
Figura 19. Fragmento del formulario P1000 definido al amparo de lo previsto en el Reglamento 883/2004 (OJEU, 2004).	138
Figura 20. Formulario de intercambio según lo previsto en la Decisión 202 (EC, 2005) ...	146
Figura 21. Ejemplo de institución registrada en el <i>Institution Repository</i> (EC, 2018).....	148
Figura 22. Sitio web del <i>Core Person Vocabulary</i> (SEMIC, 2012)	149
Figura 23. Algunas categorías asociadas a una institución en el IR (EC, 2018)	160
Figura 24. Extracto del vocabulario <i>DCMI Metadata Terms</i> (DCMI, 2018).....	165
Figura 25. Web con datos abiertos sobre ISO639-1 (LC, 2018).....	165
Figura 26. Extracto de información sobre un documento en EUR-Lex (EU, 2018)	166
Figura 27. Publicación de los vocabularios creados en GitHub.....	168
Figura 28. Publicación de los vocabularios creados en GitHub.....	169

Figura 29. Arquitectura de la Web Semántica (Shadbolt et al., 2006).....	173
Figura 30. Esquema general de descripción de datos utilizando metadatos ESSIM.....	177
Figura 31. Diagrama conceptual de los vocabularios.....	179
Figura 32. Creación de un <i>dataset</i> en el SPARQL Endpoint.....	187
Figura 33. Ejemplo de consulta SPARQL.....	188
Figura 34. Aplicación web de creación de metadatos.....	189
Figura 35. Selección de una institución para el metadato <i>creatorInstitution</i>	190
Figura 36. Selección de un valor para el metadato <i>dataSet</i>	190
Figura 37. Selección de un valor para el metadato <i>dataTerm</i>	191
Figura 38. Selección de un valor para el metadato <i>dateCreated</i>	191
Figura 39. Introducción de valores para los metadatos.....	191
Figura 40. Selección de un valor para el metadato <i>language</i>	192
Figura 41. Valores para los metadatos <i>mandate</i> y <i>references</i>	192
Figura 42. Valores para el metadato <i>relationWithTerms</i>	192
Figura 43. Selección de un valor para el metadato <i>senderInstitution</i>	193
Figura 44. Valor para el metadato <i>socialSecurityCategory</i>	193
Figura 45. Proceso de concesión de una pensión.....	200
Figura 46. Formulario oficial P5000.....	201
Figura 47. Arquitectura de intercambio propuesta.....	203
Figura 48. Ejemplo de <i>workflow</i> del proceso de negocio.....	205
Figura 49. Especificación del servicio WS1 con Swagger Editor.....	207
Figura 50. Especificación visual del servicio WS1 en Swagger Editor.....	208
Figura 51. Parámetros de entrada del método GET metadata.....	214
Figura 52. Formato del resultado del método GET metadata.....	215
Figura 53. Formato del resultado del método GET validation.....	219
Figura 54. Virtuoso Jena RDF Data Provider (OpenLink, 2018).....	220

Índice de tablas

Tabla 1. Tendencias en la migración internacional. UN (2015).....	18
Tabla 2. Formularios portables durante el año 2016 (De Wispelaere, 2017)	24
Tabla 3. Pensiones en el año 2016 (De Wispelaere, 2017)	28
Tabla 4. Campos (atributos) del directorio maestro en su versión inicial.....	88
Tabla 5. Campos del formulario E-106	96
Tabla 6. Conceptos de información	98
Tabla 7. Elementos de <i>Core Person</i> (EU, 2017)	108
Tabla 8. Definición del término nombre (TF, 2007)	112
Tabla 9. Vocabularios recomendados por el ENI (BOE, 2013)	120
Tabla 10. Elementos propuestos para ESSIM.....	144
Tabla 11. Elemento <i>creatorInstitution</i>	147
Tabla 12. Elemento <i>dataSet</i>	148
Tabla 13. Elemento <i>dataTerm</i>	150
Tabla 14. Término <i>Surname</i> incluido en el GOT (TF, 2007).....	150
Tabla 15. Elemento <i>dateCreated</i>	151
Tabla 16. Elemento <i>dateModified</i>	152
Tabla 17. Elemento <i>description</i>	152
Tabla 18. Elemento <i>format</i>	153
Tabla 19. Ejemplos de términos incluidos en el GOT (TF, 2007)	153
Tabla 20. Elemento <i>identifier</i>	154
Tabla 21. Elemento <i>language</i>	155
Tabla 22. Elemento <i>mandate</i>	156
Tabla 23. Elemento <i>references</i>	157
Tabla 24. Elemento <i>relationWithTerms</i>	158
Tabla 25. Elemento <i>senderInstitution</i>	158
Tabla 26. Elemento <i>socialSecurityCategory</i>	159
Tabla 27. Propiedades de los términos del GOT	162
Tabla 28. Relaciones entre términos en el GOT	162
Tabla 29. Propiedades de las instituciones	163
Tabla 30. Propiedades del concepto <i>Category</i>	164

Tabla 31. Propiedades propuestas para EUR-Lex.....	166
Tabla 32. Vocabulario sobre tipos de documentos oficiales.....	167
Tabla 33. Vocabulario ESSIM en formato RDF-Turtle.....	181
Tabla 34. Vocabulario GOT en formato RDF-Turtle y ejemplos de instancias	182
Tabla 35. Vocabulario IR en formato RDF-Turtle y ejemplos de instancias	184
Tabla 36. Vocabulario EUR-LEX en formato RDF-Turtle y ejemplos de instancias.....	185
Tabla 37. Vocabulario OD en formato RDF-Turtle y ejemplos de instancias	186
Tabla 38. Descripción del campo <i>Surname</i> utilizando los elementos de ESSIM (formato RDF/XML).....	194
Tabla 39. Descripción del campo <i>Surname</i> con formato Turtle.....	194
Tabla 40. Descripción del campo <i>Surname</i> con formato JSON-LD.....	195
Tabla 41. Metadatos con la descripción del campo <i>Surname</i> incluidos en el propio documento utilizando JSON-LD.....	197
Tabla 42. Metadatos con la descripción del campo <i>Surname</i> utilizando la tecnología RDFa	198
Tabla 43. Metadatos con la descripción del campo <i>Surname</i> utilizando la tecnología Microdata	199
Tabla 44. Ejemplo de contenido para el campo <i>Place of birth</i>	202
Tabla 45. Descripción del campo <i>Place of birth</i> con metadatos ESSIM	203
Tabla 46. Servicios web/métodos incluidos en la arquitectura	204
Tabla 47. Descripción del servicio WS1 con notación OpenAPI.....	207
Tabla 48. Código JavaScript que invoca un método GET del servicio WS2.....	210
Tabla 49. Descripción del servicio WS2 con notación OpenAPI	214
Tabla 50. Descripción del servicio WS3 con notación OpenAPI.....	218
Tabla 51. Descripción del servicio WS2 con notación OpenAPI.....	221
Tabla 52. Descripción del servicio WS4 con notación OpenAPI.....	223
Tabla 53. Ejemplo de datos de una solicitud.....	226
Tabla 54. Descripción del servicio WS5 con notación OpenAPI.....	227
Tabla 55. Ejemplo simplificado de datos de una respuesta a una solicitud.....	228

Lista de publicaciones

A continuación, se detallan las publicaciones asociadas a la tesis y el capítulo en el que están recogidas.

Artículos en revistas de impacto	Capítulos
Delgado, F., Hilera, JR., Ruggia, R. (2012). Soluciones para el intercambio electrónico de información de Seguridad social a nivel internacional. El profesional de la Información. 2012, Julio-agostoDDC (2013). Incluida en el JCR (Social Science Citation Index) de 2012, factor de impacto 0.44, posición 53/85:Q3, categoría INFORMATION SCIENCE & LIBRARY SCIENCE	2.3, 3.2 y 4.15
Delgado, F., Hilera, JR. y Ruggia, R. (2013). Proposal of a controlled vocabulary to solve semantic interoperability problems in social security information exchanges. Library Hi Tech, Vol. 31 Iss: 4, pp.602 - 619. Incluida en el JCR (Social Sciences Citation Index) de 2013, factor de impacto 0,62 posición 46/84:Q3, categoría INFORMATION SCIENCE & LIBRARY SCIENCE	1.1
Delgado, F., Hilera, J.R., Ruggia, R., Otón, S., Amado-Salvatierra, H.R. (2018). Using microdata for international e-Government data exchange: The case of Social Security domain. Enviado a Journal of Information Science. Enviado el 19/12/2018, pendiente de revisión.	4.3
Comunicaciones en congresos	
Delgado F, Salvador Otón, Raul Ruggia, José Ramón Hilera, R. Barchino (2013). Improving information system interoperability in social sector through advanced Metadata. 22nd International Conference on Information Systems Development (ISD2013), Sevilla, Spain, September 2-4, 2013. Core A	1.2 , 3.2, 3.3 y 5.1
Delgado, F., Otón, S., Ruggia, R., Hilera, JR. Y Gutiérrez, JM. (2013). Proposal of an Interoperability model for Social Security Information Systems. Proceedings of the 15th international conference on enterprise information systems vol. 2 (ICEIS 2013, Angers, Francia, 4-7 Julio 2013). Core B	1.1 y 1.2
Delgado, F., Hilera, J., Ruggia, R. y Otón, S. (2014). A Semantic Interoperability Standard for social security information exchange. 12th European Conference oneGovernment (EGEC 2014, Brasov, Rumanía, 12-13 Junio 2014). Core C	5.1
	5.1

Delgado, F. (2013). Modelos y estándares para la Interoperabilidad en los intercambios de información entre sistemas de Seguridad Social. Actas del V Congreso Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones Avanzadas (ATICA 2013). Huancayo (Perú), 2-4 de Octubre de 2013 5.1

Viñán, M., Otón, S., Delgado, F., Ruggia, Raul. (2014). Interoperabilidd en Instituciones de Seguridad Social. Actas del VI Congreso Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones Avanzadas (ATICA 2014). Universidad de Alcalá. Alcalá de Henares (España), 29 - 31 de octubre de 2014. ISBN: 978-84-16133-42-0. 4.1

Delgado, F. (2018). Propuesta de estándar para mejorar los aspectos semánticos en los intercambios de datos de Seguridad Social. IX Congreso Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones Avanzadas y Accesibilidad (ATICA 2018). Ciudad de Guatemala (Guatemala), 7 al 9 de Noviembre de 2018. ISBN: 978-84-17729-63-9. 3.2 y 5.1

Capítulos de libros

Delgado, F (2017). Improving Management of Social Security International Agreements. Intercultural Relations and Migration Processes. Nova Science Publisher, New York. ISBN: 978-1-53610-778-4. 1.1

Capítulo 1. Introducción

1.1 Contexto general de la tesis

El contexto en el que se sitúa esta tesis es el de los intercambios informatizados de datos en materia de Seguridad Social. Para comprender el origen y la naturaleza del problema es necesario conocer en toda su extensión los fenómenos que dan lugar a que estos intercambios se produzcan, sean necesarios y estén en franca expansión. De ahí que se dediquen las primeras páginas de esta tesis a los movimientos migratorios y sus principales consecuencias, como son el establecimiento de acuerdos en materia de Seguridad Social, el intercambio electrónico de datos para ponerlos en práctica y los problemas de interpretación que aparecen como consecuencia de la complejidad de la intercomunicación entre sistemas de información heterogéneos.

1.1.1 Los movimientos migratorios

Los movimientos migratorios han sido un fenómeno importante durante los últimos siglos. A medida que los medios de transporte y comunicación, el trabajo y los intercambios diversos han ido aumentando, las personas se han desplazado de un lugar a otro en busca de mejores condiciones de vida. La intensidad de la migración ha ido variando según las regiones del mundo y a lo largo del tiempo, pero en cualquier caso, su impacto ha sido significativo. Según las Naciones Unidas (UN, 2015) (Tabla 1), alrededor de 243 millones de emigrantes vivían en el mundo en 2015.

Major area, region, country or area of destination	International migrant stock at mid-year					
	1990	1995	2000	2005	2010	2015
WORLD	152,5	160,8	172,7	191,3	221,7	243,7
Developed regions	82,3	92,3	103,3	117,2	132,6	140,4
Developing regions	70,1	68,5	69,3	74,0	89,2	103,2
Africa	15,7	16,3	14,8	15,1	16,8	20,6
Asia	48,1	46,5	49,3	53,3	65,9	75,0
Europe	49,2	52,8	56,2	64,0	72,3	76,1
Latin America and the Caribbean	7,2	6,7	6,7	7,2	8,2	9,2
Central America	1,8	1,3	1,1	1,4	1,7	2,0
South America	4,2	4,2	4,2	4,5	5,1	5,8
Northern America	27,6	33,3	40,3	45,3	51,2	54,5
Oceania	4,7	5,0	5,3	6,0	7,1	8,1
Australia and New Zealand	4,5	4,7	5,0	5,7	6,8	7,8
Melanesia	0,1	0,9	0,1	0,1	0,1	0,1
Micronesia	0,1	0,1	0,1	0,1	0,1	0,1
Polynesia	0,1	0,1	0,1	0,1	0,1	0,1

Tabla 1. Tendencias en la migración internacional. UN (2015)

Alrededor de un 58% de los trabajadores migrantes residen en países desarrollados y la mitad de ellos en 10 países: Estados Unidos (46,6 millones), Federación Rusa (11,6 millones), Alemania (12 millones), Arabia Saudí (10,1 millones), Emiratos Árabes Unidos (8 millones), Reino Unido (8,5 millones), Francia (7,7 millones), Canadá (7,8 millones), Australia (6,7 millones) y España (5,8 millones).

Mientras que los principales vectores de migración tienden hacia los países de mayores rentas, cada vez hay una mayor diversidad. En los últimos años ha habido un considerable incremento entre los países de rentas bajas. Según Avato (2010), alrededor de 130 millones de emigrantes (70%) proceden de países con bajas o medias-bajas rentas, mientras que en ellos reside un 40% de los emigrantes. Alrededor del 51% (al menos 67 millones) de los emigrantes de países de rentas bajas o medias-bajas viven en otro país de rentas de estos mismos niveles. Es más, según Ratha (2007), en el año 2005, casi la mitad de los emigrantes de los países desarrollados residía en otro país desarrollado; casi el 80% de ellos en países vecinos. Y de ellos, una proporción importante de emigrantes se desplaza como trabajador temporal, atendiendo a necesidades del mercado laboral.

En el caso concreto de España, según se puede ver en la figura 1, en enero de 2018 la cifra de trabajadores de origen extranjero afiliados a la Seguridad Social se situó en enero en 1.815.092, (Activa, 2018).



Figura 1. Afiliados extranjeros en España (Activa, 2018)

Respecto al año 2017, el número de afiliados extranjeros aumenta un 7,56%. Del total de trabajadores extranjeros que cotizan en el Sistema de la Seguridad Social, los grupos más numerosos proceden de Rumania (321.022), Marruecos (232.151), China (100.158), Italia (95.971) y Ecuador (67.988). Les siguen los que provienen de Reino Unido (61.585), Bulgaria (55.786), Colombia (54.175) y Bolivia (48.065).

1.1.2 Trabajadores migrantes y derechos en materia de Seguridad Social

La protección social intenta reducir la vulnerabilidad y los riesgos económicos de las personas, los hogares y las comunidades. Esto se consigue a través de las distintas prestaciones de Seguridad Social como las de jubilación, prestaciones familiares, prestaciones económicas por enfermedad y servicios médicos, accidentes de trabajo, etc. Tales prestaciones, establecidas por la Convención 102 de la Organización Internacional del Trabajo (OIT, 2014) son ampliamente admitidas en el ámbito internacional.

Para tener derecho a estas prestaciones, las personas tienen que cotizar a los sistemas de Seguridad Social o bien contribuir con sus impuestos, dependiendo de la organización de cada país. Proteger los derechos de los trabajadores migrantes exige, entre otras medidas, asegurar la igualdad de trato y mantener los derechos adquiridos o en curso de adquisición con respecto a los trabajadores locales del país en que se encuentre.

Independientemente del escenario de la migración, los trabajadores migrantes a menudo tienen que hacer frente a limitaciones en la cobertura de Seguridad Social y el derecho a las prestaciones comparado con los trabajadores que viven y desarrollan su vida laboral en un país determinado. La falta de acceso a los servicios sociales y los derechos sociales portables para emigrantes no solo da lugar a problemas y vulnerabilidades de los emigrantes, sino que también crea distorsiones en el mercado laboral y en el comportamiento de las migraciones (Avato, 2010). Si los emigrantes no obtienen las prestaciones completas a partir de sus cotizaciones o impuestos porque los beneficios asociados a largo plazo no son accesibles o no son exportables, pueden decidir no cotizar y trabajar en situación ilegal o por un importe menor. Si los emigrantes han cotizado pero sus derechos no son exportables, su decisión de regresar al país de origen o decidir donde desarrollan su actividad laboral puede verse mediatizada por sus menores expectativas de recibir prestaciones o de recibirlas en menor cuantía.

El marco legal que proporciona portabilidad a los derechos sociales de los trabajadores migrantes lo constituyen los acuerdos internacionales en esta materia, especialmente aquellos que aseguran que los periodos de empleo en otros países signatarios son tenidos en cuenta en el reconocimiento de derecho de las prestaciones, mediante el reconocimiento de las cotizaciones para reconocer los periodos exigidos.

1.1.3 Recomendaciones y reglamentos internacionales

Cada vez se exige más a los gobiernos su colaboración a través de las fronteras organizacionales, sectoriales y administrativas (Janowski, 2012). Para proteger el derecho a la protección social de los trabajadores migrantes, la Oficina Internacional del Trabajo (OIT, en inglés ILO) ha desarrollado desde 1925 una serie de recomendaciones y estándares. Los más recientes son:

- El Convenio sobre la conservación de los derechos en materia de Seguridad Social (Nº 157), publicado en 1982 (OIT, 1982), orientado al mantenimiento de los derechos adquiridos o en curso de adquisición en materia de Seguridad Social de trabajadores migrantes. Esta convención se amplía con la Recomendación nº 167 (publicada en 1983), que contiene modelos para la conclusión de instrumentos bilaterales o multilaterales de Seguridad Social.
- El Marco Multilateral para Migraciones laborales de la OIT, desarrollado en 2006 (OIT, 2006). Consiste en principios no vinculantes y directrices para migraciones con orígenes laborales, en una amplia recopilación de los instrumentos políticos relevantes y las prácticas más destacadas en todo el mundo.

Estas recomendaciones definen cinco principios básicos (OIT, 1983) que constituyen la columna vertebral de la mayoría de los acuerdos bilaterales o multilaterales en materia de Seguridad Social:

- Igualdad de trato, que significa que el trabajador migrante debería tener, tanto como sea posible, los mismos derechos y obligaciones que los nacionales del país en el que desarrolla su actividad laboral.
- Determinación de la legislación aplicable, que establece que las reglas para determinar qué reglamentos de Seguridad Social tienen que ser aplicados en cada caso y a qué régimen deben contribuir (desde el país de nacionalidad hasta el país de residencia y trabajo).
- Mantenimiento de los derechos adquiridos y provisión de los derechos en el extranjero, que significa que todo derecho adquirido debería ser garantizado al trabajador migrante en cualquier territorio, incluso si se ha adquirido en otro. Por ejemplo, no debería haber restricciones en el pago de prestaciones.
- Mantenimiento de los derechos en curso de adquisición, que significa que donde un derecho está condicionado a que se complete un periodo de carencia, la acumulación debería tener en cuenta los periodos acreditados por el trabajador en cada país.

- Reciprocidad, que es un principio subrayado en todas las convenciones y significa que cada país que ha suscrito un acuerdo se compromete a aplicar los mismos mecanismos que cada una de las partes para hacer más accesibles las prestaciones de Seguridad Social a los trabajadores migrantes.

1.1.4 Acuerdos internacionales sobre Seguridad Social

Los acuerdos en materia de Seguridad Social establecen un marco legal para coordinar los sistemas de distintos países. Muchos acuerdos son bilaterales, por establecerse entre dos países para coordinar sus reglamentaciones concretas. Sin embargo, hay algunos acuerdos multilaterales que permiten a varios países coordinar partes de sus regulaciones de Seguridad Social. Cuando no hay un acuerdo bilateral ni multilateral que pueda aplicarse a nacionales de ciertos países, la mayoría de los países anfitriones establecen medidas unilaterales para proporcionar Seguridad Social a los trabajadores migrantes.

Los acuerdos son documentos normalmente estructurados en secciones y artículos seguidos de recomendaciones que proceden de la Recomendación 167 de la OIT (OIT, 1983). Sin embargo, difieren en las condiciones concretas acordadas por los países, así como en las características de sus sistemas o regímenes de Seguridad Social. Por ejemplo, pueden variar en los modelos de cotización, las condiciones de derecho a prestaciones basadas en los periodos de trabajo o de residencias, la forma de registrar las cotizaciones y pagos, entre otras. Las instituciones que proporcionan cobertura de Seguridad Social también cambian. Por otra parte, los problemas relacionados con el pago, como la fluctuación de las monedas, son factores de complejidad que deben ser abordados por las agencias o departamentos gubernamentales.

La aplicación de las funciones de Seguridad Social a través de los acuerdos internacionales implica dos tipos principales de solicitudes:

- Información de otras instituciones que participan en el acuerdo, como pueden ser las instituciones de enlace.
- Solicitudes relacionadas con prestaciones para las personas, tanto trabajadores como pensionistas. Estas últimas probablemente den lugar posteriormente a solicitudes de información dirigidas a las instituciones de enlace.

En general, la implementación global de un acuerdo de Seguridad Social implica varios pasos y diferentes clases de actividades. Por una parte, abordar los aspectos políticos y legales significa entablar discusiones previas y negociaciones, preparar el texto del acuerdo, firmar y ratificar el acuerdo, concluir los acuerdos administrativos y decidir cuándo el acuerdo será aplicable (Hirose et. al, 2011). Por otra parte, los aspectos organizacionales exigen establecer procedimientos administrativos para tratar las solicitudes y las operaciones relacionadas, así como definir los roles y las responsabilidades de estas tareas.

Pero hacer realidad la aplicación de los acuerdos requiere definir y preparar los mecanismos para realizar los intercambios de datos con otras instituciones. En particular, cuando esto se desea realizar utilizando tecnologías de la información, exige definir los formatos y la semántica de los datos a intercambiar, los mecanismos de autenticación (como firmas electrónicas), el protocolo de petición-respuesta de los intercambios y sus plazos de respuesta máximos, etc.

En consecuencia, la efectividad, fiabilidad y eficiencia de la operatividad del acuerdo dependerá de la forma en que estas tareas sean ejecutadas. Los datos intercambiados serán principalmente los siguientes:

- Datos personales del trabajador.
- Periodos de trabajo agrupados por empleador y categorías laborales agrupados de forma relevante para el acuerdo.
- Futuros periodos de trabajo para trabajadores destacados así como los empleadores en el país de origen.
- Certificados de fallecimiento y situación familiar.
- Informes médicos en el caso de que el acuerdo se refiera a prestaciones de salud.
- Periodos de residencia para los países con sistemas de Seguridad Social universal, es decir, con reconocimiento de prestaciones basados en periodos de residencia más que en periodos de trabajo.

1.1.5 Acuerdos internacionales existentes en materia de Seguridad Social

La situación de los acuerdos en vigor en materia de Seguridad Social varía mucho en función de las regiones (ILO Natlex, 2014). La red de acuerdos es particularmente densa entre países industrializados (Europa occidental, USA, Canadá, Australia, Nueva Zelanda, etc.). En particular, el acuerdo multilateral de la Unión Europea (UE) es el más amplio, abarcando a todos los países de la UE así como Noruega, Islandia, Liechtenstein y Suiza. En rigor, en este caso no debería hablarse de acuerdo sino de aplicación de una normativa común, como son los Reglamentos (EC) Nos. 883/2004 (OJEU, 2004) y 987/2009 (OJEU, 2009) sobre coordinación de los regímenes de Seguridad Social. Efectivamente, la libre circulación de los trabajadores es una de las cuatro libertades fundamentales de la Unión Europea (AISS, 2014). Se ha previsto una serie de medidas legislativas y de disposiciones complementarias con objeto de apoyar a los trabajadores migrantes. El uso de estas herramientas permite proteger a los trabajadores que se desplazan entre Estados miembros, con arreglo a tres principios fundamentales:

- La no discriminación entre nacionales de los Estados miembros de la UE y otros países.
- Garantía de transferibilidad de las prestaciones de Seguridad Social entre Estados miembros.
- La totalización de los periodos de cobertura para determinación del derecho a una prestación.

	Issued				Received			
	Flow: In 2016		Stock: Total and still valid		Flow: In 2016		Stock: Total and still valid	
	Number	% of column total	Number	% of column total	Number	% of column total	Number	% of column total
BE	21,753	3.2%	159,872	11.3%	52,254	10.8%	249,392	19.3%
BG	3,534	0.5%	7,174	0.5%	1,414	0.3%	3,464	0.3%
CZ	20,119	3.0%	52,550	3.7%	33,591	6.9%	82,495	6.4%
DK	969	0.1%	1,895**	0.1%	0	0.0%	0	0.0%
DE	99,223	14.6%	288,907	20.4%	91,988	19.0%	206,131	16.0%
EE	774	0.1%	1,374	0.1%	803	0.2%	1,955	0.2%
IE	1,578	0.2%	2,792	0.2%	259	0.1%	791	0.1%
EL	734	0.1%	3,337	0.2%	4,645	1.0%	54,041	4.2%
ES	4,812	0.7%	8,297	0.6%	18,970	3.9%	167,387	13.0%
FR	2,742	0.4%	6,281	0.4%	11,630	2.4%	72,971	5.7%
HR	528	0.1%	2,251	0.2%	5,206	1.1%	27,311	2.1%
IT	14,953	2.2%	23,888	1.7%	4,714	1.0%	19,548	1.5%
CY	372	0.1%	814	0.1%	1,752	0.4%	15,111	1.2%
LV	765	0.1%	1,387	0.1%	2,101	0.4%	607	0.0%
LT	439	0.1%	951	0.1%	4,115	0.9%	5,050	0.4%
LU	181,903	26.7%	203,998	14.4%	1,946	0.4%	5,463	0.4%
HU	3,816	0.6%	10,010	0.7%	27,463	5.7%	59,963	4.6%
MT	510	0.1%	550	0.0%	441	0.1%	3,936	0.3%
NL	141,956	20.9%	205,163	14.5%	12,106	2.5%	37,812	2.9%
AT	51,732	7.6%	140,027	9.9%	9,975	2.1%	40,048	3.1%
PL	4,833	0.7%	14,006	1.0%	81,133	16.8%	139,108	10.8%
PT	1,881	0.3%	4,015	0.3%	2,909	0.6%	11,759	0.9%
RO	5,359	0.8%	20,667	1.5%	8,238	1.7%	12,924	1.0%
SI	1,100	0.2%	9,238	0.7%	6,939	1.4%	15,138	1.2%
SK	12,950	1.9%	12,627	0.9%	80,486	16.6%	40,117	3.1%
FI	959	0.1%	5,515	0.4%	247	0.1%	758	0.1%
SE	6,221	0.9%	0	0.0%	0	0.0%	0	0.0%
UK	15,356	2.3%	157,937	11.1%	2,403	0.5%	5,111	0.4%
IS	265	0.0%	401	0.0%	37	0.0%	64	0.0%
LI	1,223	0.2%	496	0.0%	0	0.0%	0	0.0%
NO	17,147	2.5%	0	0.0%	129	0.0%	138	0.0%
CH	60,310	8.9%	70,563	5.0%	15,727	3.3%	12,167	0.9%
Total	680,816	100.0%	1,416,983	100.0%	483,621	100.0%	1,290,760	100.0%

Tabla 2. Formularios portables durante el año 2016 (De Wispelaere, 2017)

Las estadísticas publicadas por la Unión Europea (De Wispelaere, 2017) nos dan una idea del orden de magnitud de la información intercambiada como consecuencia de la aplicación de esos acuerdos. Por ejemplo, la tabla 2 recoge el número de formularios portables (entregados al interesado) durante el año 2016. Eso nos puede dar una idea de las magnitudes de los movimientos y sus posibles repercusiones.

Además, la UE tiene también acuerdos con Marruecos, Túnez y Argelia en el contexto de la iniciativa Euro-Mediterránea (Euro-Mediterranean Partnership, EMP). Asimismo, varios miembros del UE tienen acuerdos sobre Seguridad Social con países de América, África y Asia.

En Latinoamérica y el Caribe (LAC), hay tres acuerdos multilaterales principales que cubren la mayor parte de las regiones:

- El Acuerdo sobre Seguridad Social de la Comunidad del Caribe (CARICOM).

Fue firmado en 1996 por 13 de los 25 Estados miembros (CARICOM, 2001). Comprende prestaciones por jubilación, discapacidad y fallecimiento, pero excluye las prestaciones por enfermedad. El acuerdo incluye acuerdos de totalización de los periodos de cotización, aunque de menor cuantía que en otras regiones, ya que los periodos de cotización son relativamente más breves. Algunos regímenes de Seguridad Social de los países de la región tienen estructuras de prestaciones con tasas de acumulación concertadas en los primeros años, lo que puede dar lugar a que se obtengan beneficios en algunos casos.

- El acuerdo MERCOSUR.

Abarca Argentina, Brasil, Paraguay y Uruguay (MERCOSUR, 1991). La iniciativa va encaminada a racionalizar los procedimientos administrativos, por ejemplo, autorizando a que las solicitudes de pago se presenten en el país de residencia en vez de obligar a que se haga en el país de origen. Sin embargo, el acuerdo no abarca las normas sobre prestaciones.

- El Convenio Multilateral Iberoamericano de Seguridad Social (CIM, 2011).

Constituye un instrumento de coordinación de las legislaciones de Seguridad Social en materia de pensiones de los diferentes Estados Iberoamericanos que lo han firmado (España, Portugal y 12 países hispanoamericanos). Promueve la protección en caso de vejez, discapacidad y fallecimiento de cualquier trabajador migrante y miembro de su familia. Este Convenio permite totalizar los periodos de cotización o de empleo acreditados en los diferentes Estados para la determinación de las prestaciones a las que tenga derecho, brindando opciones para percibir prestaciones en otros países distintos de donde haya prestado servicios. El principio que subyace al acuerdo es la igualdad de trato entre todos los trabajadores con independencia de su nacionalidad. Actualmente, las prestaciones médicas están excluidas pero son válidos los acuerdos voluntarios bilaterales o multilaterales al respecto entre países.

En Norteamérica, los Estados Unidos y Canadá otorgan cobertura sobre la base de la residencia y, por consiguiente, a determinados trabajadores migrantes, incluyendo aquellos que provienen de México.

Otro acuerdo multilateral de extensión de la protección (*Insurance Protection Extension, IPE*), proporciona cobertura de Seguridad Social a los nacionales de los países pertenecientes al Consejo de Cooperación del Golfo (*Gulf Cooperation Council, GCC*). Este acuerdo tiene como fin la creación de un mercado común y de seguridad regional (GCC, 2006). Los miembros actuales son: Arabia Saudí, Bahrein, Emiratos Árabes Unidos, Kuwait, Qatar y Omán. El Alto Consejo del GCC dirige la implementación de los reglamentos que aseguran la igualdad entre los ciudadanos los países del GCC en el campo de los seguros sociales y de retiro. De acuerdo con esto, todo Estado del GCC debería exigir a sus empresarios que paguen las cotizaciones de los trabajadores de otros estados encuadrados en los regímenes de seguros sociales o retiro civil. En 2006, la institución de Arabia Saudí (GOSI) adoptó reglamentos para poner en práctica estos acuerdos mediante mecanismos que faciliten el movimiento de la fuerza laboral entre los Estados del GCC.

En la región de Asia y el Pacífico, la mayor parte de los acuerdos han sido establecidos por Australia y la República de Corea, seguidos por la India, Japón, Nueva Zelanda y Filipinas. Es importante subrayar que la mayor parte de estos acuerdos se establecen entre países desarrollados y no entre países de la misma región.

En el sudeste de Asia, entre los países miembros de la Asociación de Naciones de Asia Sudoriental (ASEAN), no hay acuerdos bilaterales o multilaterales de Seguridad Social. La Declaración de la ASEAN sobre la protección y promoción de los derechos de los trabajadores migrantes, adoptada en 2007, comprometía a los países miembros de ASEAN a intensificar sus esfuerzos para solucionar los problemas. ASEAN está integrada por Indonesia, Filipinas, Malasia, Singapur, Brunei Darussalam, Myanmar, Camboya, República Democrática Popular de Lao y Vietnam

Los Estados miembros de la Comunidad de Estados Independientes (CEI) firmaron en 2008 un Convenio sobre la condición jurídica de los trabajadores migratorios y de sus familiares y lo ratificaron en 2011. Este Convenio, que agrupa a diez de las quince repúblicas exsoviéticas, tiene por objeto fortalecer la cooperación en el ámbito de la protección de los trabajadores migrantes en los países de la CEI. El Convenio establece que los trabajadores migrantes, incluidos los trabajadores temporales y los fronterizos, tienen los mismos derechos en cuanto a las condiciones laborales que los ciudadanos del país de acogida. El convenio se aplica a todos los trabajadores migrantes y a los miembros de sus familias que se establezcan de forma legal en el país de destino (OIT 2012b). Asimismo, algunos países de la CEI han firmado acuerdos bilaterales con otros países de la CEI, con lo que atribuyen a los países receptores la responsabilidad de la cobertura. Sin embargo, las reformas emprendidas en muchos de estos países desde finales de los años 90 dieron lugar a situaciones en las que los sistemas de jubilación eran radicalmente distintos entre unos países y otros, lo que ha originado que se produzcan grandes distorsiones, por ejemplo, disparidad de requerimientos en cuanto a las aportaciones mínimas, la edad de jubilación, etc. (ISSA, 2012).

En el ámbito de las iniciativas regionales, hay que citar el proyecto MIGSEC de la OIT. Se trata de un proyecto de extensión de la cobertura de la Seguridad Social a los

trabajadores migrantes y sus familias en África (MIGSEC, por sus siglas en inglés), que tiene la finalidad de fortalecer la capacidad institucional mediante la promoción de acuerdos bilaterales y multilaterales. En el proyecto participan trece países, y el informe publicado en 2010 por la OIT señaló la existencia de una serie de trabas, pero también de iniciativas, en relación a las cuestiones que atañen específicamente a África (McGillivray, 2010). En el informe se señalan como problemas destacados los requisitos para causar derecho, la falta de exportabilidad de las prestaciones y la situación del sector no regulado. Se observa asimismo que, en 2010, se suscribieron veinte acuerdos bilaterales en materia de seguridad social entre países africanos y más de 50 acuerdos con países no africanos.

En el África subsahariana, la situación más habitual consiste en que existan acuerdos con los principales países destino de la emigración (Francia, Países Bajos y Portugal). Por otra parte, un acuerdo multilateral (CIPRES) se ha firmado entre 15 países francófonos en el África central y occidental, pero ha sido ratificado sólo por cinco países. La CIPRES reúne a 15 países y 19 administraciones de la Seguridad Social en África occidental y central. Tiene por objeto coordinar las iniciativas, así como mejorar la provisión y administración de las prestaciones de la Seguridad Social. La CIPRES facilita la instauración y la administración de un sistema que reconozca los derechos de los trabajadores migrantes.

En lo que se refiere al África occidental, en 2012 la Comunidad de los Estados de África Occidental (CEDEAO) adoptó el Convenio General sobre Seguridad Social que propiciará que los trabajadores migrantes jubilados que hayan trabajado en cualquiera de los países de la CEDEAO ejerzan su derecho a gozar de un seguro social en su país de origen. El citado Convenio suprime las restricciones a la concesión de prestaciones y establece los requisitos de coordinación inherentes al acuerdo (Dimechkie, 2013).

En cuanto a la aplicación de los acuerdos, el acuerdo multilateral de la Unión Europea es el que da cobertura a la mayoría de los casos actuales. Una idea general nos la da la tabla 3 que refleja el número de personas y las cuantías (en euros) pagadas por el Estado miembro de residencia a los pensionistas que recibían en 2016 pensión de dos o más Estados Miembros.

Todas las experiencias referenciadas muestran un crecimiento esperado en las operaciones relacionadas con los acuerdos en materia de Seguridad Social que posteriormente se transformarán en un crecimiento de los intercambios de datos (Delgado, 2017). Una eficiente implementación de estas operaciones, especialmente en lo que se refiere al intercambio de datos entre instituciones, requerirá a su vez un adecuada utilización y apoyo de las Tecnologías de la Información y de las Comunicaciones (TIC).

	Old-age pension		Survivors' pension		Invalidity pension		Total		
	Number of persons	Annual amount paid	Number of persons	Annual amount paid	Number of persons	Annual amount paid	Number of persons	Annual amount paid	Average annual amount per pensioner
BE	141,848	1,404,391,236	52,398	516,175,344	2,074	18,703,221	158,991	1,939,269,801	12,197
BG	7,630	12,645,753	970	637,077	872	1,084,492	9,472	14,367,323	1,517
CZ									
DK	2,153	22,271,407			204	2,782,300	2,357	25,053,707	10,629
DE									
EE	7,315	19,747,631	377	503,979	1,615	3,376,537	9,307	23,628,147	2,539
IE	6,493		945		544		7,982		
EL									
ES	2,147	51,079,344	317	5,653,566	262	3,192,000	2,726	59,924,910	21,983
FR							464,489	2,215,055,148	4,769
HR	7,583	18,408,970	290	520,379	266	395,883	8,139	19,325,232	2,374
IT	225,957	2,260,739,615	82,334	361,649,193	9,863	37,171,064	310,128	2,659,559,873	8,576
CY									
LV	538	1,357,725	3	3,516	75	134,838	609	1,496,079	2,457
LT	9,550	24,427,869	1,938	1,328,118	3,223	4,937,103	14,711	30,693,090	2,086
LU	11,636	330,540,960	3,764	80,152,980	3,818	87,277,284	19,218	497,971,236	25,912
HU	521	1,718,976	169	229,812	3,791	9,056,652	4,481	11,005,440	2,456
MT	1,875	12,360,000	868	1,890,000			2,743	14,250,000	5,195
NL	89,011	732,055,073	1,438	6,742,032			90,449	738,797,105	8,168
AT	125,516	1,849,456,926	28,623	251,103,465	7,649	91,159,126	161,788	2,191,719,517	13,547
PL									
PT	185,797	591,243,459	74,217	151,696,308	12,069	38,233,630	272,083	781,173,397	2,871
RO	1,105	3,313,977	64	86,058	171	334,106	1,340	3,734,140	2,787
SI	21,778	109,815,852	1,807	6,434,184	1,155	2,018,364	24,740	118,268,400	4,780
SK									
FI	57,059	1,211,913,686	9,557	57,918,698	5,649	52,930,489	63,469	1,322,762,873	20,841
SE									
UK									
IS	1,101	29,944,314			208	4,013,331	1,309	33,957,645	25,942
LI									
NO									
CH									
Total	906,613	8,687,432,773	260,079	1,442,724,710	53,508	356,800,420	1,166,042	10,486,957,916	8,994

Tabla 3. Pensiones en el año 2016 (De Wispelaere, 2017)

1.1.6 Experiencias existentes en la utilización de las tecnologías de la información como apoyo a los intercambios de datos en materia de Seguridad Social

El avance las tecnologías de la información en las últimas décadas ha hecho emerger una nueva ola de métodos que pueden aportar nuevos enfoques cualitativos y cuantitativos (Kim et. al, 2013). Como se ha dicho anteriormente, la puesta en práctica de los acuerdos internacionales se basa en gran parte en los intercambios de datos, que deberían ser seguros y fiables, así como permitir el seguimiento de las solicitudes y respuestas. Además, los mecanismos basados en las TIC tienen que estar coordinados entre todas las partes que participen para asegurar la compatibilidad de los intercambios de datos y la correcta interacción de los procesos afectados.

La Asociación Internacional de la Seguridad Social señala (AISS, 2013) que el éxito en la aplicación de los acuerdos internacionales depende en gran medida de que exista un intercambio eficaz y fiable de datos entre las organizaciones participantes. Los datos que se necesitan no suelen ser sólo los que constan en la información que se maneja habitualmente sobre el empleado, sino que incluye otros datos como los periodos trabajados en otros países o notificaciones de los cambios en su situación familiar.

Las experiencias existentes de intercambios de datos para la implementación de los acuerdos internacionales basados en las TIC varían desde los complejos sistemas para implementar acuerdos multilaterales que afectan a muchos países, como es el caso del EESSI (*Electronic Exchange of Social Security Information*) de la UE (EESSI, 2012) hasta los más simples que dan apoyo a los acuerdos bilaterales.

Para situar en contexto el grado de utilización de las tecnologías de la información en este ámbito, a continuación, se va a dar una idea general de las principales experiencias, empezando por la experiencia de la Unión europea, sin duda la más avanzada, y a la que se ha visto especialmente vinculada la experiencia profesional del autor de esta tesis. Estas experiencias están recogidas en el artículo del autor de impacto JCR (Delgado et al., 2013c), en el que se encuentran citadas las fuentes con información de los proyectos referenciados.

- Unión europea

El primer precedente que se debería citar es el proyecto EDIS (*Electronic Data Interchange in the Social Domain*). El estudio de viabilidad de este proyecto se realizó entre abril de 1991 y enero de 1992, con el fin principal de mejorar los intercambios de información en materia de pensiones. Estaba directamente relacionado con la iniciativa ENS (*European Nervous System*) de la Comisión europea. Contaba con el apoyo de TEDIS (*Trade Electronica Data Interchange System*) que pretendía el desarrollo de mensajes EDIFACT para el ámbito de la Seguridad Social. De esta forma, se trataba de establecer un estándar mundial en este dominio.

El proyecto EDIS no fue más allá del estudio de viabilidad y dio paso al proyecto SOSENET (*Social Security Network*) lanzado en 1992, también encuadrado en el ENS y con la pretensión de extenderse a todos los sectores de la Seguridad Social y no sólo al de pensiones. Su objetivo principal era simplificar y acelerar los procedimientos administrativos con el fin de mejorar la adquisición de los derechos, así como la concesión y el pago de las prestaciones de Seguridad Social.

Ese mismo año arrancaba el proyecto TESS (*Telematics for Social Security*), continuación natural de SOSENET, del que se consideraba segunda fase. En cuanto al contenido del programa TESS, se definieron las necesidades de los entonces doce Estados miembros para los cuatro sectores de Seguridad Social (pensiones, enfermedad, desempleo y accidentes de trabajo). Su objetivo general era crear y construir un modelo de referencia de servicios telemáticos que permitiera

coordinar los sistemas de Seguridad Social de todos los Estados miembros en los que era aplicable el Reglamento (CEE) 1408/71 (DOUE, 1971), entonces vigente, y en el cual las instituciones pasaran de un intercambio de papel a un intercambio electrónico. Sobre esa base, se fijó un primer conjunto de normas de arquitectura común. Esas normas se evaluaron mediante el intercambio de formularios de la serie E-500 (formularios de vida laboral) entre siete instituciones del sector de pensiones de jubilación de distintos Estados miembros. Los intercambios se realizaban al amparo de lo previsto en las Decisiones 117 y 118 (INSS, 1993). Dichas decisiones marcaban pautas para almacenar las referencias sobre la vida laboral del trabajador migrante en el país de nacionalidad, con vistas a su reconstrucción un año antes de que alcanzase la edad teórica de jubilación. Los estándares técnicos elegidos en esos momentos fueron X.25, X.400 y EDIFACT. Los resultados obtenidos por el proyecto mostraron la viabilidad de este tipo de intercambios.

Como paso siguiente, en 1994 se aprueba el Plan General del programa TESS. Este programa pretendió consolidar el trabajo realizado complementando el conjunto de normas de arquitectura común y la transferencia de la red piloto creada a un entorno de producción.

El Plan General establecía una serie de proyectos denominados "Build". Los tres primeros, Build 1 al 3, recogían las acciones realizadas en la primera fase, es decir, el establecimiento de la red, la conectividad y el desarrollo del piloto de intercambio de los formularios E-500. Los restantes eran:

- Build 3+, se proponía consolidar los resultados del Build 3, extendiendo su utilización a todas las instituciones y todos los Estados miembros.
- Build 4, ampliaba el proyecto al intercambio relacionados con los formularios E200 (solicitud de prestación, vida laboral, etc.) dentro del área de pensiones.
- Build 5, extendía el Build 3+ al sector de la asistencia sanitaria, con el intercambio de los formularios de la serie E100 (facturación de gastos y reconocimiento de derecho, entre otros).
- Build 6, se dedicaba al desarrollo de los mensajes para Build 4 y Build 5 mediante un enfoque intersectorial.
- Build 7, dedicado al problema específico de los formularios portátiles.
- Build 8, trataba de los servicios comunes de gestión necesarios para los servicios telemáticos operativos.

El éxito de los distintos bloques fue desigual: Build 3 dio lugar a intercambios entre diversas instituciones y Estados pero sin una periodicidad fija y con volúmenes escasos. Build 4 sólo se llegó a poner en funcionamiento entre instituciones muy concretas de Alemania, Italia y Francia, conviviendo en cualquier caso con los intercambios en papel. Build 5 ha sido el que mayor cantidad de intercambios ha propiciado. Se trata de intercambios de ficheros planos a través del servidor FTP que proporciona la red europea sTesta. Los intercambios se realizan con carácter

semestral. A título de ejemplo, puede señalarse que España envía al año a los restantes Estados miembros más de 600.000 formularios en formato electrónico en lo que a este sector se refiere. Build 6 elaboró los mensajes EDIFACT planificados, pero nunca llegaron a ponerse en funcionamiento salvo en los casos citados del Build 3 y el Build 4. Sin embargo, elaboró un producto especialmente interesante para el tema tratado en esta tesis: el **Glosario de Términos**. En el capítulo 2 se hablará más extensamente de este glosario. Finalmente, los Build 7 y 8 no llegaron a desarrollarse más allá de unos entregables teóricos.

El proyecto TESS se dio por cerrado con los cambios legislativos de los Reglamentos (EC) números 883/2004 (OJEU, 2004) y 987/2009 (OJEU, 2009) sobre coordinación de los regímenes de Seguridad Social. Dichos Reglamentos dieron lugar al nacimiento del proyecto EESSI (EESSI, 2012). El motivo hay que buscarlo dos puntos concretos de estos reglamentos.

Por una parte, el artículo 4 (2) del 987/2009 establece que *“La transmisión de datos entre instituciones u organismos de enlace deberá realizarse por medios electrónicos bien sea directa o indirectamente a través de los puntos de acceso bajo un marco común que garantice la confidencialidad y la protección de los datos intercambiados”*.

Por otra parte, el artículo 1 (2) del Reglamento 987/2009 define el punto de acceso como una entidad que debe proveer:

- Un punto de contacto electrónico;
- Un encaminamiento automático basado en la dirección, y
- Un encaminamiento inteligente basado en software que permita un encaminamiento y prueba automáticos.

En definitiva, se establece la vía electrónica como la única legalmente reconocida para los intercambios de datos en materia de Seguridad Social.

Para que esa vía única de intercambios pudiera ser una realidad, se creó el proyecto EESSI. Está financiado por la UE y su objetivo es mejorar la protección de los derechos del ciudadano, informatizando la aplicación de la normativa europea de coordinación de Seguridad Social. Gracias al intercambio informático, se facilitan y agilizan las decisiones de cálculo y pago de las prestaciones, la comprobación de datos es más eficaz, hay una interfaz más cómoda y flexible entre los distintos sistemas, y se ofrecerá una recopilación detallada de datos estadísticos sobre intercambios europeos. El objetivo inicial era que en 2012 toda la información que hasta el momento se venía intercambiando mediante impresos, se procesase electrónicamente, cosa que no se consiguió en esa fecha, como veremos más adelante. En términos de modelos de evolución del gobierno electrónico estudiados por Joan Batlle-Montserrat, Ernest Abadal y Josep Blan (Batlle-Montserrat et al., 2011), se podría considerar que esta etapa corresponde a la transformación del modelo de Baum y Di Maio, dado que no sólo se va a cambiar

la vía de los intercambios, sino que los procedimientos van a sufrir una modificación profunda que va a afectar al propio funcionamiento de las instituciones.

EESSI establece una arquitectura común que proporciona acceso público a una base de datos de instituciones (*Institution Repository*) y permite el intercambio de información entre más de 10.000 instituciones competentes mediante un protocolo de comunicación seguro. Una vez entre plenamente en producción, la información será intercambiada mediante “documentos electrónicos estructurados” (*Structured Electronic Documents, SED*).

Durante 2006 y 2007 se lleva a cabo un estudio de viabilidad, punto de partida para descripción de los intercambios existentes en esos momentos tanto en papel como electrónicos, así como para elaborar los requerimientos de una plataforma central de intercambio. En esos momentos había más de 200 formularios utilizados para la comunicación entre más de 10.000 instituciones de Seguridad Social que fueron rediseñados para su sustitución por documentos electrónicos. El modelo de EESSI abarcaba aproximadamente 150 flujos de negocio y más de 300 SED (ISA, 2014).

En 2008, se inicia la construcción del sistema, planificando su puesta en producción en mayo de 2010 momento a partir del cual debían haber empezado a utilizarlo los Estados miembros, en el transcurso de un periodo transitorio de 24 meses.

Las pruebas realizadas durante 2010 mostraron una complejidad mayor que la prevista. Una campaña intensiva de pruebas de aceptación durante 2011 puso de manifiesto defectos importantes que llevaron a una aceptación parcial a principios de 2012 y al inicio de un periodo de reflexión en marzo del mismo año. Dicho periodo se dio por cerrado a mediados de 2013 y dio paso a un nuevo desarrollo, basado en una arquitectura de la que se hablará en el capítulo 2. Oficialmente, el proyecto entró en producción en julio de 2017.

- MERCOSUR (Argentina, Brasil, Paraguay y Uruguay)

En septiembre de 2008 empezaron las operaciones del acuerdo multilateral entre los países de MERCOSUR (Argentina, Brasil, Paraguay y Uruguay), implementadas utilizando el sistema SIACI (Sistema de Acuerdos Internacionales) (Kounowsky, 2012). Este sistema pone en funcionamiento parte de los procedimientos de solicitud de pensiones, permitiendo a los ciudadanos iniciar estas formalidades en su país de residencia, mientras las instituciones de enlace interaccionan electrónicamente con las instituciones de Seguridad Social en los países donde los solicitantes han trabajado, con el fin de obtener la información necesaria para completar el procedimiento.

La arquitectura de software de SIACI responde a un modelo de “federación” de las instituciones que participan, basadas en la utilización de software similar que articula el intercambio de los datos esenciales. El intercambio de datos está

codificado en el lenguaje XML (W3C, 2008), mientras que las comunicaciones se realizan utilizando servicios web (WSDL, 2007). La identificación de los beneficiarios se gestiona por las instituciones de cada país y SIACI establece las correspondencias entre ellos. La información transmitida es firmada digitalmente por los usuarios, utilizando medidas externas con capacidad de proceso (*tokens* o *smart-cards*). Todas las operaciones son registradas por una tercera parte confiable, la Organización Iberoamericana de Seguridad Social, que administra el conjunto del sistema.

- Convenio Iberoamericano de Seguridad Social

Está en desarrollo una aplicación llamada SIBERO, de características similares a SIACI de MERCOSUR. En noviembre de 2018 se encuentra en fase de revisión por parte de los países firmantes del convenio.

El sistema está compuesto por dos aplicaciones, una que es la administradora, función que cumple la OISS, y otra aplicación que se utiliza en las entidades.

La aplicación Administradora se encarga de realizar la definición de los paquetes de datos para cada entidad, así como de asegurar la integridad de la información transmitida entre las entidades. Estos paquetes de datos, con la información a transmitir, pueden ser definidos como formularios, archivos, imágenes, audios, etc. Cualquier dato o grupo de datos que se desee transferir entre entidades, puede ser parametrizado como un tipo de paquete de datos y enviarlo como tal.

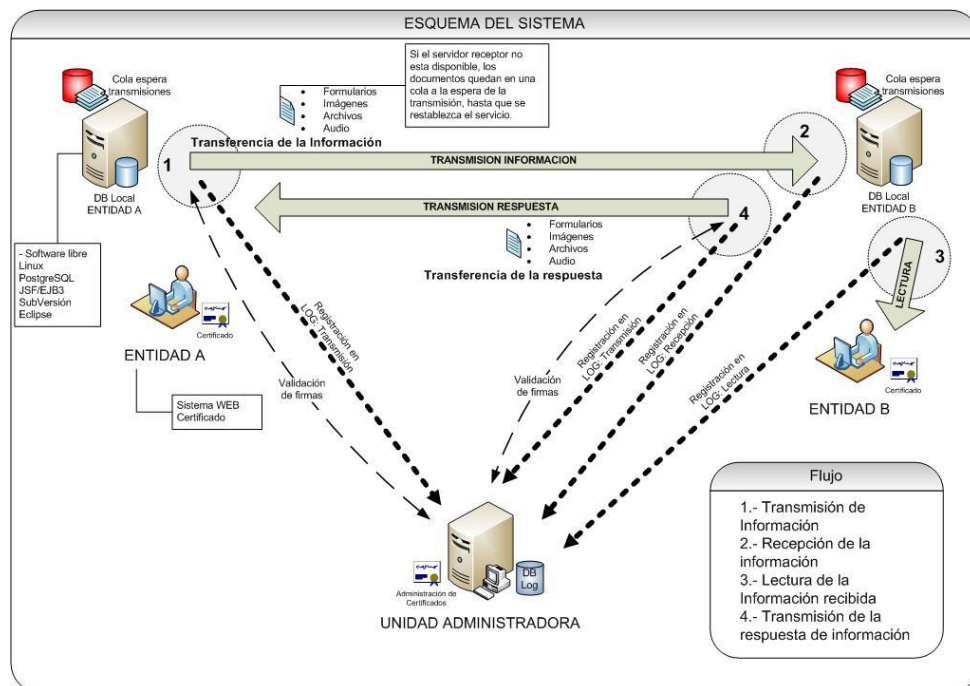


Figura 2. Esquema de SIBERO. Fuente: OISS

En la figura 2 se muestra el esquema del sistema SIBERO, donde están representadas dos Entidades y la Unidad Administradora. Se puede ver que la información se transmite desde la Entidad Origen (ENTIDAD A), a la Entidad Destino (ENTIDAD B) y la Unidad Administradora asegura y certifica el envío, pero no tiene acceso al detalle de la información. Se registran *logs* en los diferentes estados del envío.

- Acuerdos bilaterales establecidos por la República de Corea.

El *National Pension Service* de la República de Corea, acordó en 2002 un memorando de entendimiento (MOU) con las instituciones de Seguridad Social de siete países. Dicho acuerdo permite el intercambio de datos relacionado con beneficiarios que residen en esos Estados (NPS, 2016). El objetivo principal es compartir la información sobre cambios en la situación de los beneficiarios (como fallecimiento, divorcio, etc.) que pudieran dar lugar a modificaciones en las pensiones recibidas. Por ejemplo, el intercambio de datos entre el NPS y la *Social Security Administration* (USA) son realizados a través del sitio web *Government Services Online* (GSO), mientras que los intercambios con el Centrelink de la Seguridad Social Australiana son realizados a través del sitio web EXCpert. NPS ha minimizado las exigencias para la identificación personal de los ciudadanos de las instituciones extranjeras. Actualmente, esta información consiste en el nombre, la fecha de nacimiento y el número de Seguridad Social.

- Acuerdos bilaterales establecidos por Australia

Australia es uno de los países que ha puesto en práctica un mayor número de acuerdos internacionales en materia de Seguridad Social. Los correspondientes intercambios de datos se realizan mediante el sistema EXCpert basado en el “*the Australian Government Locator Services*” (AGLS, 2010) que es un estándar australiano de metadatos alineado con el *Dublin Core Metadata Initiative* (DCMI). Se utiliza para describir los recursos online de la Administración. La utilización del estándar AGLS es obligatorio para las Agencias del Gobierno australiano.

- *Death Data Community* (DDC)

En el año 2013, a iniciativa del *Deutsche Post AG* y del *Deutsche Rentenversicherung* (institución competente alemana en materia de pensiones) se estableció la *Death Data Community* o Comunidad de Datos de Defunción (DDC, 2013). Su objetivo principal es constituir un marco de referencia para facilitar el intercambio de datos de defunción de una forma utilizable, segura y eficiente.

El origen de estos intercambios está en las pensiones que cada uno de los países abona en un país diferente. El control de vivencia de los perceptores de estas pensiones resulta caro e inseguro para las instituciones responsables de abonar las prestaciones y la experiencia muestra un alto índice de fraude y error en estos abonos.

Si bien existen convenios bilaterales con este fin, los intercambios reales y eficientes son muy inferiores a lo que sería deseable. Con la creación de esta comunidad, Alemania pretende facilitar la estandarización y la agilización de los intercambios.

En septiembre de 2013, se le celebró en Colonia la que se anunció como primera conferencia sobre este tema (*Conference on the International Exchange of Death Data*), si bien no se ha celebrado ninguna más hasta el momento. Representantes de 24 países de la Unión Europea, Estados Unidos y otras partes del mundo participaron en el encuentro.

En primer lugar, la conferencia pretendía poner de relieve las ventajas de una plena automatización de los intercambios de datos por defunción, impidiendo un pago indebido de las pensiones. Según se decía, comparado con un procedimiento manual, una solución automatizada ahorraría tiempo y mejoraría la eficiencia. En el futuro, la intención era que todos los participantes involucrados se beneficien de la implementación de un procedimiento estandarizado. Con ese fin, se propuso un *dataset* (conjunto de datos o diseño de registro) para el intercambio de datos, basado en los intercambios que ya estaban entonces en producción entre diversos países.

También se trataron otros detalles sobre los principios básicos de los procedimientos, así como los desarrollos que actualmente se están llevando a cabo.

En el caso de Alemania y por su legislación concreta, es el *Deutsche Post AG* quien asume la competencia de los pagos de pensiones en el extranjero y también de verificar que se están haciendo correctamente, de ahí su papel en estos intercambios de datos.

Según los convocantes de la conferencia, los beneficios esperados de un intercambio automatizado de datos son:

- Prevención anticipada de pagos indebidos.
- Hace innecesario el trámite de los certificados de vivencia.
- Mejora de la eficiencia y reducción de los costes de proceso.
- Establece procedimientos automatizados, estandarizados y por lotes (no por casos individuales).
- Ofrece formatos sencillos para los intercambios de datos.
- Reduce los esfuerzos de implementación.
- Asegura la protección de la privacidad.

Uno de los objetivos principales es conseguir fomentar el uso de un formato acordado para el conjunto de datos que sea práctico y facilite el intercambio entre las partes implicadas vía solicitud y respuesta. Como vía de transmisión, el organismo alemán proponía sPad (*secure Platform for the electronic Access to Data*), herramienta de tipo web desarrollada por la institución alemana. Medios alternativos serían también admisibles. En cuanto a la frecuencia, como regla general, los intercambios se realizarían una vez al mes.

1.2 Contexto tecnológico

1.2.1 Situación de los intercambios en materia de Seguridad Social

El número de trabajadores migrantes en el mundo sigue creciendo y hay actualmente más de mil millones de personas que pueden considerarse migrantes internos o internacionales (AISS, 2014). Tal como se ha visto en la sección anterior, la implementación de acuerdos internacionales en materia de Seguridad Social constituye un desafío, incluso para los países con sistemas de administración electrónica altamente desarrollados, como son el caso de Australia y Corea. A pesar de que un importante número de países han establecido acuerdos bilaterales y multilaterales en un esfuerzo por proteger al trabajador migrante, la parte operacional de estos acuerdos sigue siendo un desafío. Los principales problemas tienen su origen tanto en los conceptos de la Seguridad Social como en la terminología y las prácticas en los diversos países, así como en el gran número de instituciones involucradas en los procedimientos necesarios para poner en marcha los acuerdos. Otros problemas frecuentemente citados son las barreras políticas y legales, la accesibilidad de los datos y los costes tecnológicos.

Por otra parte, la falta de estándares predominantes para los procesos y datos de protección social, ha hecho que la adopción de intercambio de datos entre los países haya sido extremadamente lenta, llevando a retrasos en los procesos, errores y altos costes de administración (Lee-Archer et al., 2007).

Como consecuencia, los intercambios electrónicos de datos para la aplicación de los acuerdos internacionales en materia de Seguridad Social todavía se basan principalmente en procesos manuales, con una automatización muy limitada. Hay soluciones basadas en mecanismos específicos para el acuerdo, pero obligan a las instituciones a implementar diferentes procedimientos para interactuar con cada una de las instituciones.

Las experiencias con éxito, a pesar de la diversidad de los escenarios y las implementaciones, comparten ciertas características:

- Se centran en los mecanismos técnicos para el intercambio de datos. Utilizan lenguaje XML para el formato de datos y servicios web para la interconexión de sistemas.
- Utilizan herramientas básicas de seguridad, como el control de *login* como sistema de acceso, seguridad de las comunicaciones, encriptación de los paquetes de datos y autenticación básica de los mensajes.
- Realizan una aplicación básica y no automatizada de las leyes de protección de datos, con un uso muy bajo de firmas electrónicas y certificados digitales.
- Cuentan con diferentes implementaciones para diferentes acuerdos, con un bajo nivel de reutilización de las soluciones.
- Los sistemas de intercambio de datos no están conectados a los principales sistemas de las instituciones de Seguridad Social.

Aún en el caso de que estén operativas, la mayoría de las soluciones existentes no serían capaces de hacer frente al desafío del crecimiento en el número de acuerdos y de casos individuales, haciéndose aún más necesaria una estandarización y modelos comunes con el fin de implementar mecanismos basados en las TIC. Incluso en el caso de la Unión Europea, con distintas experiencias desarrolladas desde hace más de 30 años, la situación dista mucho de un intercambio 100 % informatizado, a pesar de estar así previsto en los vigentes Reglamentos Comunitarios.

En definitiva, en un escenario creciente y de mayor complejidad de los acuerdos y los consiguientes intercambios de datos, las soluciones actuales basadas en las TIC no son capaces de satisfacer estos requerimientos. Tales limitaciones en la implementación pueden poner en peligro los esfuerzos realizados por los países para mejorar la cobertura de la Seguridad Social de los trabajadores migrantes. Si se pretende mejorar esta situación, se deberá profundizar en los conceptos que tratan la capacidad de los sistemas de trabajar conjuntamente independientemente de su diseño y entorno de ejecución, es decir, se debe observar el problema desde el punto de vista de la interoperabilidad.

1.2.2 Interoperabilidad entre sistemas de información de Seguridad Social

Establecer intercambios regulares de datos requiere que las instituciones de los distintos países compartan un marco común de datos y procesos de interacción. Eso es lo que se conoce normalmente como interoperabilidad, concepto que trata de la interconexión de diferentes sistemas de información.

La aplicación de las técnicas de interoperabilidad ha mostrado su eficacia para facilitar la implementación de programas integrados de acuerdo con las actuales tendencias de la Seguridad Social (ISSA, 2012; Durán-Valverde, 2012). Su aplicación requiere contar con un fuerte apoyo político y cooperación inter-institucional; así como el establecimiento de metas institucionales compatibles y enfoques operacionales para implementar programas de Seguridad Social coordinados (Delgado et al., 2013). Llevar a cabo una integración internacional supone un desafío adicional para instituciones en las que a menudo no es posible aprovechar las infraestructuras existentes para los procesos nacionales. Además, la aplicación de las técnicas de interoperabilidad para implementar los intercambios de información internacionales tiene que ser flexible y capaces de integrar partes con diferentes tecnologías, diferentes niveles de desarrollo y distintas políticas de interoperabilidad.

La aplicación de la interoperabilidad, especialmente en contextos organizacionales complejos, requiere tener en cuenta diferentes niveles o áreas involucradas en la interacción. De esta forma, el concepto general se subdivide en un conjunto de dimensiones. Se suelen distinguir las dimensiones siguientes, ya sea en la interoperabilidad aplicada al gobierno electrónico, en las bibliotecas y servicios de información digitales, o en otros ámbitos diferentes (EIF, 2017):

- **Política:** las partes en colaboración deben tener visiones compatibles, prioridades alineadas y centrarse en los mismos objetivos.
- **Legal:** la sincronización adecuada de la legislación entre los países que colaboren incluye el hecho de que los datos electrónicos generados en un país tengan el debido reconocimiento en el momento de ser utilizados en el país receptor.
- **Organizativa:** hace referencia a la definición de objetivos de negocio, modelado de procesos de negocio, así como tratar de facilitar la colaboración entre administraciones o instituciones que desean intercambiar información y pueden tener diferentes estructuras y procesos internos.
- **Semántica:** se preocupa de asegurar que el significado preciso de la información intercambiada sea entendible por cualquier otra aplicación que no haya sido inicialmente desarrollada para este propósito. La interoperabilidad semántica permite a los sistemas combinar la información recibida con otros recursos de información y procesarla de manera coherente.
- **Técnica:** abarca los aspectos críticos de enlazar sistemas de ordenadores y servicios. Incluye aspectos clave como interfaces abiertas, interconexión de servicios, integración de datos y middleware, presentación e intercambio de datos, seguridad y accesibilidad de los servicios.

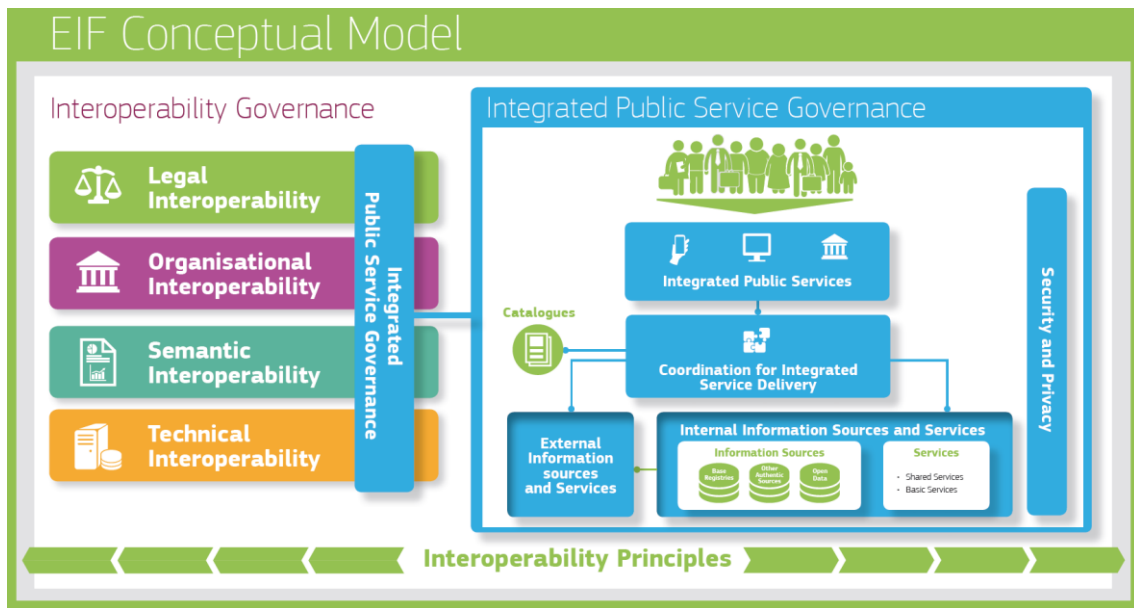


Figura 3. Modelo conceptual de interoperabilidad. EIF (2007)

Esta es la visión del programa europeo ISA (*the Interoperability Solutions for European Public Administrations program*), y, por tanto, el punto de vista de la Administración Pública electrónica. Dicha visión queda plasmada en el documento de referencia *European Interoperability Framework v 2.0* (EIF) desde sus versiones publicadas en 2004 (EIF, 2004) y 2010 (EIF, 2010), hasta la más reciente publicada en 2017 (EIF, 2017). Su modelo conceptual puede verse en la figura 3.

Otros autores distinguen incluso más dimensiones. Así por ejemplo, desde una perspectiva de la defensa y estrategia, Munk (2002) considera que la interoperabilidad tiene sentido en distintos dominios como el físico, el de la información, el cognitivo y el social. Al centrarse en la interoperabilidad de la información, Munk distingue un “nivel sintáctico” relativo a la representación de los datos, en particular incluyendo lenguajes, mensajes y formatos de datos empleados durante el intercambio de información.

Por otro lado, la organización británica de investigación UKOLN especialmente enfocada a informar prácticas e influenciar políticas en el área de las bibliotecas digitales, sistemas de información, gestión bibliográfica y tecnologías web, distingue un conjunto más amplio de dimensiones de la interoperabilidad (tecnológica, semántica, política y humana, intercomunitaria, legal, e incluso internacional) que abarcan distintas cuestiones necesarias para el intercambio de información entre sistemas (UKOLN, 2013).

Todas las dimensiones de interoperabilidad son relevantes en el escenario de los intercambios de datos de Seguridad Social. Es más, el personal de negocio debería tener una participación activa en la definición de los aspectos relacionados con las dimensiones política, legal, organizacional y semántica.

Siguiendo el modelo de EIF, en el caso de la Seguridad Social, la **interoperabilidad política** está basada en el hecho de que los gobiernos y las instituciones implementan sistemas de Seguridad Social ratifican las recomendaciones internacionales y firman acuerdos. Implementar proyectos de interoperabilidad a gran escala encaja más en una agenda política de cambio que en resolver un típico proyecto (Hellberg, 2013).

La **interoperabilidad legal** debe ser utilizada donde sea posible dar una estructura formal a los enfoques políticos que afectan a la aplicación de la interoperabilidad en la implementación de los sistemas de Seguridad Social.

Como los programas de Seguridad Social están fuertemente basados en leyes y reglamentos, tratar la dimensión legal significa contar con una adecuada sincronización de estándares en los que se debe basar la cooperación para asegurar que las jurisdicciones, competencias y responsabilidades de los organismos que participan en proyectos y que requieren de interoperabilidad están claramente definidos, así como que los datos electrónicos intercambiados tengan entidad y reconocimiento legal.

En el contexto de los intercambios de datos de Seguridad Social, la interoperabilidad legal consiste en los acuerdos por sí mismos, así como en los procedimientos administrativos formalizados sobre cómo será intercambiada la información.

La **interoperabilidad organizacional** se refiere a la definición de modelos de proceso y administrativos compatibles, independientemente de la estructura organizacional, con el fin de construir una colaboración efectiva entre las partes buscando el intercambio de información y servicios, o bien compartirlos.

En el ámbito de la Seguridad Social, debido a la falta de estandarización en los procesos de negocio, funciones administrativas similares pueden realizarse de diferentes formas. Es más, no hay todavía un modelo formal generalizado en la práctica. En definitiva, la dimensión organizacional debería especificar el proceso de la institución que debería interoperar, definiendo los correspondientes servicios y acuerdos de nivel de servicio.

En cuanto a la **interoperabilidad semántica**, asegura que el significado exacto de la información intercambiada es comprensible para cualquier otra aplicación no desarrollada inicialmente para el mismo propósito. La interoperabilidad semántica facilita a los sistemas combinar de forma correcta la información recibida con la información de otras fuentes. Se trata de asegurar que el sistema receptor va a interpretar la información de la misma forma que el remitente, uno de los problemas principales a abordar cuando se implementa un intercambio de datos automatizado entre sistemas. Conseguirlo requiere acordar, por ejemplo, la forma en que la información y su contexto deben ser representados. Eso es lo que permitirán las herramientas automáticas para compartir y procesar la información, incluso cuando los datos se han almacenado de forma independiente.

Algunos autores distinguen entre interoperabilidad semántica para la visualización e interoperabilidad semántica para el procesamiento (Indarte, 2011). La primera se refiere a la posibilidad de ver toda la información y su contexto. La segunda a la capacidad de procesar la información.

Efectivamente, el objetivo de la interoperabilidad semántica no es sólo facilitar la interconexión de recursos de información, sino también permitirles entenderse automáticamente y, como resultado, ser reutilizados por aplicaciones de ordenador no relacionadas con su creación (Delgado et al., 2013b).

Otros autores distinguen entre la interoperabilidad técnico-semántica, como aquella que permite la identidad y el intercambio de información y, por otra parte, la interoperabilidad idiomática-semántica, como aquella que permite el soporte multilingüe (Criado et al., 2010).

La interoperabilidad semántica se destaca como la dimensión más débil en los intercambios de datos sobre Seguridad Social y las soluciones propuestas para solucionar los problemas que presenta tienen todavía un nivel muy bajo de desarrollo. En el capítulo 2 se analizará su situación actual.

Por último, la **interoperabilidad técnica** cubre los aspectos de conexión y comunicación entre sistemas de información. Incluye aspectos clave como interfaces abiertas, servicios de interconexión, la integración de datos y el middleware, la presentación e intercambio de datos, localización y recuperación de recursos, accesibilidad, seguridad e integración de aplicaciones y servicios. Distintos estándares y un uso extendido de especificaciones pueden ser identificados en diferentes áreas.

La madurez de Internet y las tecnologías web asociadas, como es el caso de los servicios web, han hecho en los últimos tiempos mucho más asequible la construcción de aplicaciones y procesos que puedan ser utilizados más allá de las fronteras. En particular, la adopción de arquitecturas orientadas a servicios (SOA) se está acelerando gracias a la innata flexibilidad que proporciona.

La interoperabilidad técnica para los intercambios de información sobre Seguridad Social se consigue principalmente mediante la utilización de estándares como:

- La mencionada arquitectura SOA, que proporciona un modelo común de interacción.
- Los modelos de datos que facilitan los intercambios de datos.
- Las tecnologías de interconexión y *middleware* que permiten una interacción general.

1.2.3 Problemas semánticos en los intercambios de datos de Seguridad Social

En el contexto de los sistemas de Seguridad Social que dan apoyo a los acuerdos internacionales y los consiguientes intercambios de datos, las prácticas de interoperabilidad semántica son esenciales para desarrollar definiciones comunes e interpretaciones de datos que tienen que ser procesados por diferentes organizaciones. Las operaciones de Seguridad Social utilizan un amplio rango de términos, que, a pesar de tener el mismo o parecido nombre, pueden ser interpretados de forma diferente. Son ejemplos, las expresiones idiomáticas de difícil traducción e interpretación, tales como “*family group*”, “*members of the same household*”, “*unemployed person*”, “*old age pension*”, “*health benefits*”, “*social security contributions*”, etc.

Dada la heterogeneidad lingüística en los países implicados en los intercambios de datos de Seguridad Social, así como la falta de estandarización terminológica en esta área, la interoperabilidad semántica presenta importantes deficiencias.

Una de las causas de estos problemas es la falta de política de nomenclatura, que lleva a disparidades y conflictos en las interpretaciones de los mismos términos. Un término debería referirse siempre a un solo concepto; de otra forma las relaciones entre conceptos y definiciones podrían llegar a ser muy difíciles de mantener. Si el mismo término es utilizado para definir múltiples conceptos en diferentes contextos, debería llevar un prefijo que hiciera las veces de identificador (técnicamente un espacio de nombre) que establecen claramente al contexto.

Si nos remitimos a los primeros antecedentes de intercambios electrónicos en el seno de la Unión Europea, el informe del Comité Consultivo para la Seguridad Social de los Trabajadores Migrantes de 1994 (EC, 1997), señala entre las principales dificultades a la hora de tratar la información “*los problemas de interpretación debido a la falta de un glosario administrativo*”. La nota del Comité Consultivo para la Seguridad Social de los Trabajadores Migrantes de 7-11-1994 decía textualmente:

“(...) Una interpretación incorrecta de un término por parte de una institución de Seguridad Social puede causar malentendidos, lo que finalmente puede hacer que el trabajador migrante no reciba la prestación o que, al menos, la reciba con un retraso considerable.

(...)

El glosario terminológico puede ayudar a las instituciones de Seguridad Social, por ejemplo, a la hora de informar a instituciones de otros Estados miembros sobre los ingresos de una persona. Una descripción del término “ingresos” en el glosario terminológico podría indicar si la cantidad es bruta o neta y si se han deducido las cotizaciones a la Seguridad Social de los ingresos.”

Posteriormente, el Plan General (EC, 1997) recogía en su página 15 las mismas dificultades, nombrando textualmente los “*problemas de interpretación debidos a la falta de un glosario semántico*”. Como ejemplo se decía:

“(...) se dan problemas de traducción conceptual: Al preguntar por el lugar de nacimiento para identificar a una persona, ¿qué grado de precisión se está pidiendo: el Estado, la provincia, la ciudad, el municipio o la parroquia? ¿Significa “parroquia” lo mismo en Portugal que en España? (...)”

“Cada solicitud requiere cuatro fases de traducción conceptual:

- *La solicitud pasa del nivel nacional de la institución remitente al nivel europeo;*
- *Pasa entonces del nivel europeo al nivel nacional de la institución receptora;*
- *Después la solicitud pasa del nivel nacional de la institución receptora al nivel europeo;*
- *Por último, pasa del nivel europeo al nivel nacional de la institución emisora de la solicitud.*

Por consiguiente, para que una solicitud siga su curso correctamente, es necesario conocer bien tanto los reglamentos europeos como la legislación nacional de Estado miembro receptor. El riesgo de que se produzcan errores de comprensión o de interpretación es muy real y puede dar lugar a demoras o, peor aún, a la aplicación incorrecta de los reglamentos de coordinación y de la legislación nacional.”

Como se menciona anteriormente en este capítulo, dentro del desarrollo del programa TESS se desarrolló un sub-proyecto, Build 6, que incluía un modelo de datos conceptual y el desarrollo de un glosario de términos. De ambos se hablará en profundidad en el capítulo 2.

Más recientemente, en el caso de EESSI, a pesar de las validaciones basadas en el modelo de negocio, los conflictos semánticos han surgido desde los comienzos del proyecto. Por ejemplo, en el año 2018 no hay todavía un uso consistente de la terminología para la descripción de los datos utilizados en los Documentos Electrónicos Estructurados (SED) y muchos de estos documentos pueden ser normalmente mal interpretados o incluso rechazados.

Esa fue la causa de que, en pleno desarrollo del proyecto, se creara el *Data Modelling Expert Group* (DMEG). De este grupo es miembro e impulsor el autor de esta tesis. Entre las tareas que le son asignadas al grupo se encuentra el asegurar la corrección semántica de los datos intercambiados.

En noviembre de 2015, el grupo produce un primer documento con recomendaciones de las que conviene recoger las realizadas en su introducción (AC, 2015):

“Dadas las diferencias lingüísticas, culturales, legales, técnicas y entornos administrativos, hay desafíos significativos como para asegurar que el significado preciso y los formatos de la información intercambiada en EESSI son entendidos y preservados.

Analizando el contenido de los formularios en papel actualmente en uso se ve que parecidos o idénticos conceptos son usados en todos los sectores de la Seguridad Social. Lo que se ve en el análisis mencionado es que hay diferencias en la convención de nombres, en la estructura del contenido y en los principios del contenido en el contexto de negocio en que se aplica. La consecuencia de esta diversidad es un enorme esfuerzo para procesar y evaluar la información basada en estas entidades especialmente si la arquitectura de negocio depende de un intercambio electrónico de datos. Además, esta diversidad introduce el riesgo de malas interpretaciones y manejo erróneo de la información.”

Como solución, el DMEG propuso la creación de un modelo de datos basado en la definición de un conjunto de bloques comunes que asegurasen el mismo significado y formato para todos los formularios y sectores. Este modelo de datos debe entenderse como una pieza dentro de la arquitectura de información del sistema (figura 4).

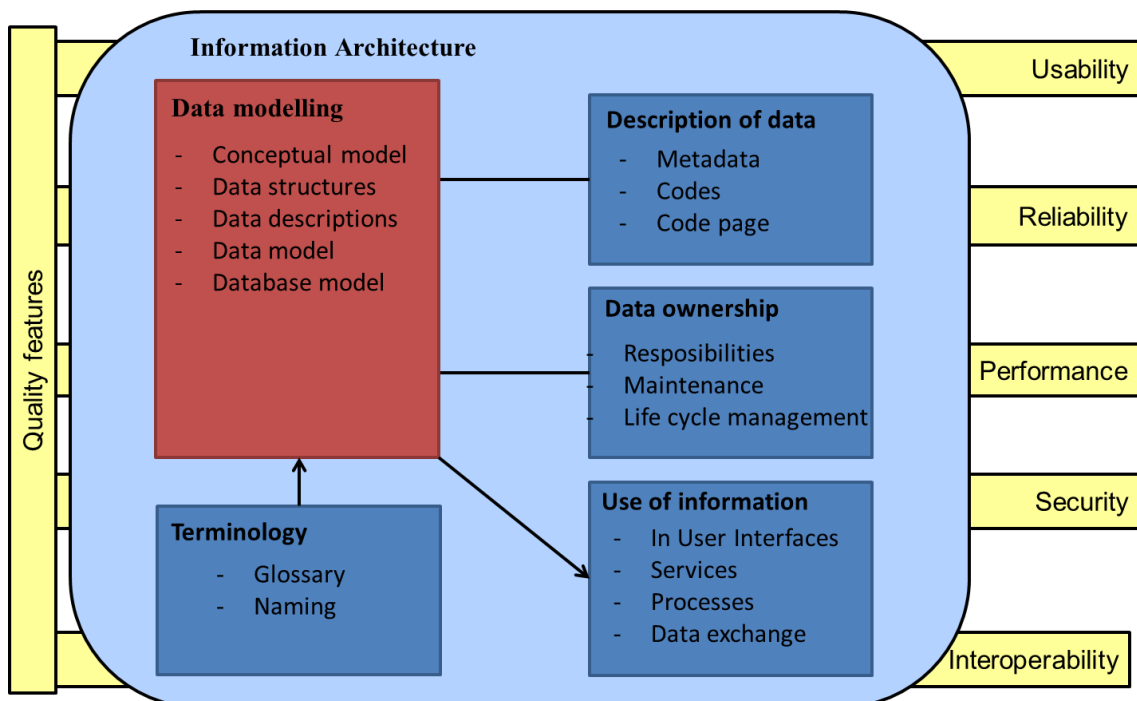


Figura 4. Arquitectura de información propuesta por el DMG (Pekka, 2013)

Hay que recordar que la arquitectura de información de un sistema consiste en las especificaciones de su estructura, clasificación y uso de la información (Pekka, 2013). Cuando se trata de un intercambio de datos entre sistemas, también se incluye la forma en que la información es intercambiada.

La parte más importante de la arquitectura de información es el modelo de datos. El propósito del modelado de datos es describir la información de una forma estructurada y formal (modelo conceptual). Esta tarea se realiza habitualmente siguiendo un enfoque *top-down*: desde el más alto nivel del modelo conceptual hasta el más bajo nivel del modelo de datos y finalmente hasta el modelo de base de datos orientado a la implementación. El diagrama de la figura 4 define las capas dentro de la arquitectura de información.

El modelo incluye el uso de glosarios. En ese sentido, los glosarios son utilizados comúnmente para orientar las políticas y relacionar términos y conceptos, (Fraunhofer, 2009). Los glosarios también permiten detectar conflictos de homonimia en cualquier sitio donde se proporcione una representación que produzca ese problema. En consecuencia, el conflicto de homonimia puede ser detectado consultando el glosario con el contexto como parámetro y obteniendo los diferentes calificadores.

Los glosarios pueden ser un punto de partida para resolver los problemas semánticos en los intercambios de datos de Seguridad Social, pero una solución más sofisticada sería deseable. A largo plazo podría ser en un mecanismo que definiera:

- Los términos básicos y las relaciones comprendidas en tema de que se trata.
- Las reglas para combinar los términos y las relaciones para definir extensiones del vocabulario.

Aunque eso podría hacernos pensar en taxonomías u ontologías, las características de los intercambios y los datos intercambiados van a hacer descartar estas opciones y escoger otras más sencillas y más acordes con los estudios actuales.

En escenarios similares, la Estrategia Europea de Interoperabilidad propone una solución de compromiso que consiste en una definición de metadatos (EIS, 2012). Un esquema de metadatos de este tipo debería especificar la información a intercambiar en términos conceptuales, sus propiedades y las relaciones entre los conceptos. Tal como apunta Christian (2001), las características del metadato necesario en cada caso dependen en última instancia de los intereses y los recursos de las comunidades que los utilizan.

1.3 Hipótesis y objetivos de la tesis doctoral

El objetivo general de esta tesis se enmarca dentro de los intercambios de datos entre las instituciones de Seguridad Social de distintos países. En ese contexto, tomando como referencia la definición del concepto de interoperabilidad del EIF ya mencionada en la sección 1.2.2, se va a partir de una perspectiva general de la problemática y se profundizará en las distintas dimensiones.

La hipótesis de partida de esta tesis se enuncia de la siguiente forma: *“es posible solucionar una parte importante de los problemas de la interoperabilidad semántica existente en los intercambios de datos internacionales entre instituciones de Seguridad Social mediante la utilización de una especificación de metadatos, vocabularios controlados y una arquitectura de intercambios basada en tecnologías web”*.

Se establecen los siguientes objetivos para demostrar la hipótesis:

1. Analizar la situación de los intercambios electrónicos de información entre instituciones de Seguridad Social de diferentes países y detectar los problemas existentes en el campo de la interoperabilidad que inciden en el grado de éxito de los intercambios.
2. Dado que la dimensión semántica aparece como la más débil y una de las principales causas de que los intercambios en el contexto estudiado no estén teniendo el éxito esperado, estudiar que glosarios, taxonomías, tesauros y ontologías existen, tanto del ámbito del *e-Government* (Administración Pública en general), como de la Seguridad Social, y pueden contribuir a paliar los problemas existentes.
3. Proponer un tipo de vocabulario que permita dotar semánticamente los intercambios con metadatos para garantizar la correcta interpretación de cada dato y definir la correspondiente especificación que en el futuro pueda convertirse en un estándar internacional.
4. Proponer un esquema de arquitectura de intercambios entre instituciones de Seguridad Social, basada en servicios web, que demuestre la viabilidad y aplicabilidad de la especificación propuesta.

Dado el carácter internacional de la propuesta, en la tesis será importante tener en cuenta en todo momento el uso de estándares. Por ese motivo, se va a analizar los estándares existentes en otros dominios, sus diferencias y similitudes, así como su posible utilidad para el dominio de la Seguridad Social. Como conclusión, se va a proponer una especificación de metadatos capaz de eliminar todas las ambigüedades en la interpretación de los datos intercambiados de forma automática, de tal forma que pueda ser propuesta como estándar ISO y puesta a disposición de las instituciones de Seguridad Social de todo el mundo.

La especificación desarrollada soluciona problemas que aparecen en los intercambios ordenador-ordenador que se realizan al amparo de los acuerdos internacionales, como son la compatibilidad de estructuras y la incompatibilidad de representaciones en los modelos de los conceptos, las incompatibilidades lingüísticas, así como incompatibilidades conceptuales producidas por el hecho de que distintos sistemas de información utilicen conceptos con significados difusos.

La condición del autor de miembro del DMEG, la Comisión Técnica para el tratamiento de la información de la Unión Europea, la Comisión Técnico-Informática del Convenio Iberoamericano de Seguridad Social y la red de expertos en Tecnologías de la Información y las Comunicaciones de la AISS, podrá contribuir a su aceptación. Por otra parte, la circunstancia de contar como director de la tesis a un miembro de la AISS con un alto nivel de responsabilidad, ayudará a su difusión entre los afiliados de esta asociación en todo el mundo.

1.4 Organización del trabajo

El capítulo 2 se va a dedicar principalmente a los objetivos uno y dos, tratando de establecer cuál es la situación actual en cuestión de interoperabilidad en general y en el ámbito de la Seguridad Social en particular. En concreto:

- Definiciones y modelos de madurez de interoperabilidad.
- Interoperabilidad en las Administraciones Públicas.
- Interoperabilidad entre organizaciones de Seguridad Social.
- Vocabularios.
- Metadatos.
- Estándares.

Termina el capítulo con un análisis crítico de la situación, especialmente a lo que se refiere a la interoperabilidad en los intercambios de información entre instituciones de Seguridad Social en el ámbito internacional.

En el capítulo 3 se va a desarrollar la propuesta apuntada en el tercer objetivo y que es el núcleo central de esta tesis: una nueva especificación de metadatos basada en los elementos de *Dublin Core*, centrándose en los problemas de los sistemas de información de la Seguridad Social y apoyándose en vocabularios controlados.

El capítulo 4 se va a dedicar al último de los objetivos propuestos: mostrar la implementación de una nueva especificación de metadatos para los intercambios entre instituciones de Seguridad Social, utilizando tecnologías de web semántica.

La implementación se muestra mediante dos pruebas. La primera de ellas se centró en la creación y validación de los metadatos propuestos en capítulo 3. La segunda consistió en crear una arquitectura simplificada para un caso de estudio, basada en servicios web.

En el capítulo 5 se van a presentar las conclusiones del trabajo realizado y se van a proponer nuevas líneas de investigación que desarrollen las soluciones apuntadas.

Capítulo 2. Estado del Arte

En el capítulo 1 se ha descrito el contexto general en el que se desarrolla esta tesis: la migración de personas entre distintos países en todas partes del mundo, y los consiguientes intercambios de datos entre instituciones de Seguridad Social que se producen para dar cumplimiento a los acuerdos internacionales establecidos para garantizar los derechos de las personas que se desplazan, sea por motivos laborales o sea por cualquier otro motivo. Estos intercambios de datos se mueven a su vez en un contexto tecnológico como es el de la interoperabilidad entre sistemas de información.

En este segundo capítulo de la tesis se va a analizar el estado del arte de estas técnicas de interoperabilidad con sucesivas aproximaciones al problema que se pretende analizar.

En primer lugar, desde un punto de vista general, se recogen las principales definiciones encontradas en la literatura, así como los modelos de madurez de carácter más general.

A continuación, se van a ver los principales modelos de referencia adaptados a la problemática concreta de las administraciones públicas, así como los principales casos y referencias encontrados de la utilización de las técnicas de interoperabilidad en este ámbito concreto.

Dado que la Seguridad Social presenta importantes particularidades dentro del mundo de las Administraciones Públicas, se ha incluido el principal modelo de referencia de interoperabilidad para, a continuación, recoger las experiencias existentes en este ámbito agrupadas por los posibles escenarios.

A partir de este punto se va a profundizar en las distintas dimensiones de interoperabilidad, su situación general y las posibles vías de mejora. Dado que el fin principal de esta tesis es estudiar los aspectos semánticos, se va a recoger a continuación las principales herramientas utilizadas en esta dimensión: vocabularios y metadatos.

El capítulo termina con un análisis crítico de la información recopilada hasta el momento y la situación de la interoperabilidad semántica en el ámbito de la Seguridad Social.

2.1 Interoperabilidad

2.1.1 Definiciones

Si bien existen múltiples definiciones de la interoperabilidad, una de las más citadas, es la que propuso el IEEE en 1990 en el *IEEE Standard Computer Dictionary*: “la capacidad de dos o más sistemas, redes de comunicación, aplicaciones o componentes para intercambiar información entre ellos y para usar la información que ha sido intercambiada” (IEEE, 1990). Define la interoperabilidad de la información a un alto nivel conceptual.

Por su parte, en la *Interoperability Technical Framework* (ITF, 2003), el gobierno australiano enmarca la interoperabilidad en el ámbito de las tecnologías de la información, definiéndola como “la capacidad de transferir y utilizar información de una manera uniforme y eficiente a través de múltiples organizaciones y sistemas de tecnologías de la información. Permite asegurar el nivel de beneficios que recaudan las empresas, gobierno y la economía en general a través del comercio electrónico”.

En la misma línea, el “*European Interoperability Framework for Pan-European eGovernment Services*” (EIF, 2004), documento de referencia para el desarrollo de estándares e infraestructuras comunes necesarias para la implementación de la interoperabilidad en el área de la administración electrónica europea, publicado por el programa europeo IDABC (*Interoperable Delivery of European eGovernment Services to public Administrations, Businesses and Citizens*), considera que “la interoperabilidad es la capacidad de los sistemas de tecnologías de la información y la comunicación (TIC) y de los procesos de negocio que éstas soportan, para intercambiar datos y para ser capaces de compartir información y conocimiento”.

El comité CEN-TC251 del Comité Europeo de Normalización define la interoperabilidad no como una capacidad o habilidad como los anteriores, sino como un estado “que existe entre dos entidades de aplicaciones, en relación con una tarea específica, cuando una aplicación puede aceptar datos de la otra y llevar a cabo una tarea de forma apropiada y satisfactoria sin necesidad de intervención externa” (CEN, 2012).

En definitiva, desde una perspectiva puramente técnica, y como se desprende de las definiciones recogidas, la interoperabilidad hace referencia a la relación entre dos o más recursos o sistemas (dispositivos hardware y de comunicación o componentes de software) TIC que precisan trabajar conjuntamente de forma fácil o automática. Asimismo, las definiciones coinciden en señalar dos capacidades con las que deben contar los sistemas implicados: comunicarse entre ellos (para poder transferir información) y entender la estructura de la información que se transfiere entre las entidades (para poder utilizarla). En cuanto a la información objeto de la transferencia, puede ser de cualquier tipo: texto, video, audio, aplicaciones software, datos sin elaborar, etc.; en general recursos y su meta-información asociada.

Todos los agentes implicados, incluidos los gobiernos, la industria, los usuarios y consumidores, los interlocutores sociales y otros, han reconocido la necesidad de la interoperabilidad y los beneficios que puede aportar su consecución. Sin embargo, aunque todo el mundo está de acuerdo sobre su necesidad, el alcance y la aplicación de la interoperabilidad, así como los incentivos que la alientan o las barreras técnicas o políticas que la dificultan, siguen siendo cuestiones controvertidas.

Una de las principales cuestiones de preocupación se refiere a la aplicación de la interoperabilidad y la definición de estándares abiertos que la hacen posible. El desarrollo de estándares, y su adopción a gran escala, es uno de los principales mecanismos o vías para lograr la interoperabilidad, al proporcionar guía y consejo a los implicados en la creación de información digital o servicios y establecer unas normas mínimas que los sistemas deben cumplir para poder comunicarse e intercambiar información.

La Conferencia Internacional de Interoperabilidad se presentaba con la siguiente declaración (IC, 2008): *“En una sociedad globalizada y dado el aumento de la diversidad de sistemas y aplicaciones, la interoperabilidad hace posible el desarrollo de un mercado de masas y evita los efectos no deseables de la fragmentación. En este contexto, los estándares son una pieza clave para alcanzar un nivel satisfactorio de interoperabilidad en este mercado global, lo cual explica la razón de que la interoperabilidad sea un negocio a nivel internacional.”*

Por ese motivo, los principales actores de la industria de las TIC han trabajado activamente a través de asociaciones y organismos de estandarización para avanzar en el desarrollo y adopción de estándares abiertos. Además de invertir importantes sumas de dinero en el desarrollo de estándares, y de hacer que el software y el hardware que producen sean compatibles con dichos estándares, estos actores están cooperando a niveles nunca antes vistos para alinear sus tecnologías de manera que puedan interoperar, consiguiendo progresos significativos.

Existen una serie de principios que guían el desarrollo de las TIC en las administraciones y proporcionan los criterios de selección de estándares que ayudan a desarrollar marcos de interoperabilidad (GIF, 2007). A continuación, se presentan siete principios clave:

- **Interoperabilidad:** Garantizar un medio de comunicación consistente para crear un flujo de información entre los ciudadanos, las empresas, los gobiernos y sus socios. Se deben seleccionar sólo aquellas especificaciones que son relevantes para la interconectividad de los sistemas, integración de datos, servicios de acceso y contenido.
- **Escalabilidad:** Asegurar la facilidad de uso, capacidad de adaptación y capacidad de respuesta de las aplicaciones cuando los requisitos cambian y las peticiones fluctúan.

- **Reusabilidad:** Establecer procesos y estándares para procedimientos similares tratando de desarrollar servicios y estructuras de datos reutilizables.
- **Apertura:** Centrarse en estándares abiertos, es decir, todos los estándares y directrices deberán ajustarse a los principios de los estándares abiertos. Siempre que sea posible, se adoptarán los estándares abiertos al mismo tiempo que se establecen las especificaciones técnicas. Los estándares que son independientes del vendedor y el producto deben ser tenidos en cuenta en favor de sus alternativas propietarias.
- **De apoyo al mercado:** Se deben seleccionar aquellos proveedores TIC que basan sus desarrollos en estándares.
- **Seguridad:** Asegurar el intercambio fiable de información que puede llevarse a cabo de conformidad con una política de seguridad establecida.
- **Privacidad:** Garantizar la privacidad de la información en cuanto a ciudadanos, empresas y organizaciones gubernamentales, y respetar y hacer cumplir las restricciones definidas legalmente en el acceso y difusión de la información, y así como asegurar que los servicios tienen que soportar un nivel uniforme de protección de datos personales.

2.1.2 Modelos de madurez de interoperabilidad

La interoperabilidad puede ser vista desde distintas perspectivas, que pretenden determinar los tipos, vertientes, enfoques o dimensiones de la interoperabilidad. La aplicación de la interoperabilidad, especialmente en contextos organizacionales complejos, requiere tener en cuenta los diferentes niveles o áreas organizacionales involucradas en la interacción. Como se ha visto en el capítulo 1, el concepto general de interoperabilidad se subdivide en un conjunto de dimensiones: política, legal, organizativa, semántica y técnica.

Por otra parte, la organización británica de investigación UKOLN, especialmente enfocada en informar prácticas e influenciar políticas en el área de las bibliotecas digitales, sistemas de información, gestión bibliográfica y tecnologías web, distingue un conjunto más amplio de dimensiones de la interoperabilidad (tecnológica, semántica, política y humana, intercomunitaria, legal, e incluso internacional) que abarcan distintas cuestiones necesarias para el intercambio de información entre sistemas (UKOLN, 2013).

En cualquier caso, no se trata de niveles aislados, sino que guardan una relación entre sí que es conveniente tener en cuenta. Si consideramos las dimensiones técnicas, semántica y organizativa como las básicas en cualquier esquema de interoperabilidad nos encontramos con las siguientes relaciones:

- La **interoperabilidad técnica** puede ser considerada un prerrequisito para alcanzar la interoperabilidad semántica y organizativa. Los sistemas y componentes deben conectarse primero en lo físico, en el nivel de protocolos, antes de que se pueda establecer una interoperabilidad a nivel de datos y semántica.
- Por su parte, la **interoperabilidad semántica** precisa de acuerdos a nivel técnico y organizativo. En primer lugar, cualquier medida tomada tiene que tener en cuenta aspectos de la interoperabilidad técnica, y ser construida en base a estándares, guías y soluciones dadas. En segundo lugar, como el significado de los datos depende de la finalidad o contexto en el que es empleado, las medidas en el contexto de la interoperabilidad semántica están estrechamente vinculadas, incluso requieren e implican medidas en el contexto de la interoperabilidad organizativa.
- De la misma manera, la **interoperabilidad organizativa** precisa de las dos anteriores. Por un lado, es necesario que se proporcionen unos mínimos niveles de seguridad en el nivel técnico para poder iniciar la interoperabilidad a nivel organizativo, definiendo políticas, estrategias y procedimientos. Por otro lado, el entendimiento común y el procesado e intercambio adecuado de los conjuntos de datos se basa en acuerdos sobre conceptos o su relación mutua. Estos acuerdos y el proceso de alcanzarlos sólo pueden ser afrontados desde un nivel organizativo.

Estas relaciones posibles, pero no únicas, entre los niveles o dimensiones de interoperabilidad, son los que dan lugar al diseño de modelos concretos de interoperabilidad, generalmente ajustados al contexto de que se trate.

Modelo LISI

Un aspecto importante en el ámbito de la interoperabilidad es la posibilidad de poder medir el nivel de madurez alcanzado por un conjunto de sistemas de información. El modelo desarrollado con este propósito es el LISI (*Levels of Informations Systems Interoperability*) (LISI, 2018).

Su desarrollo comenzó en 1993 por un encargo a la *Corporación MITRE*, y fue publicado finalmente en 1998 por el *DoD CAISR Architecture Working Group (AWG)* del IEEE. Su objetivo es evaluar las capacidades e implementaciones en el contexto de un sistema de información con respecto al nivel de interoperabilidad requerida.

Para lograr la interoperabilidad técnica se pueden establecer varios niveles progresivos (figura 5): en el nivel técnico el objetivo es lograr la compatibilidad entre los sistemas que están "conectados"; en el nivel de sistemas es necesario obedecer a los mismos protocolos como requisito para lograr la interconexión de los sistemas; y, por último, el nivel operativo se centra en el modelo de datos y objetos que permita el intercambio de información.

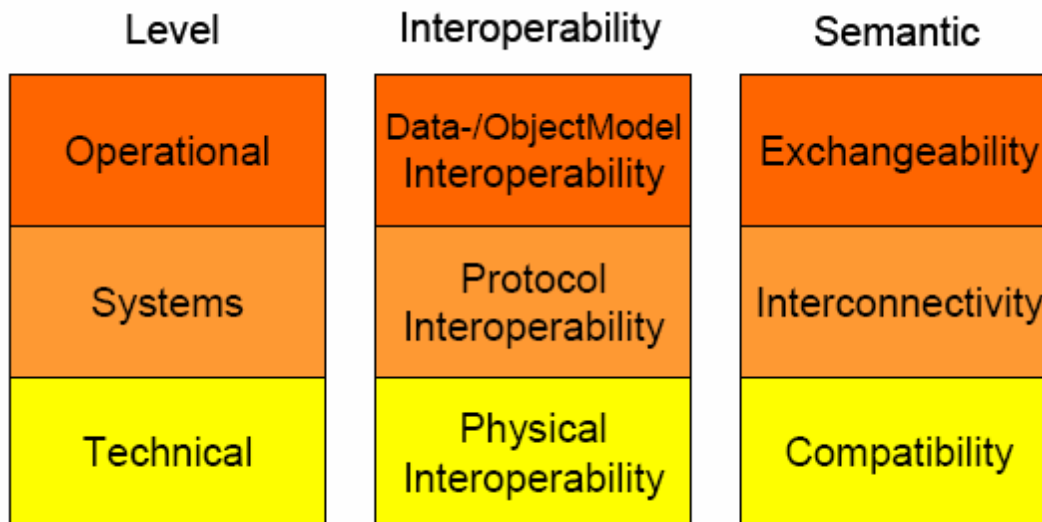


Figura 5. Bases del modelo de interoperabilidad LISI (Bueno de la Fuente, 2018)

A partir de esta base, LISI establece cinco niveles de madurez de la interoperabilidad entre sistemas de información, asociando una serie de atributos a esos niveles que abarcan un amplio rango de consideraciones, principalmente técnicas y organizativas, en torno a la interoperabilidad. Estos atributos se describen mediante el acrónimo PAID (*Procedures, Applications, Infraestructure y Data*). El Modelo de Referencia LISI sitúa los cinco niveles como filas de una matriz y los cuatro atributos como columnas. Las celdas de las intersecciones describen las capacidades necesarias para alcanzar un nivel específico de interoperabilidad para un determinado atributo (figura 6).

Nature of Operational Information Interaction	Corresponding Interoperability Level	Implications			
		P	A	I	D
Cross-Domain Interactive Manipulation	Enterprise 4	Enterprise Level	Interactive	Multiple Topologies	Enterprise Model
Shared Applications & Databases	Domain 3	Domain Level	Groupware	World Wide Networks	Domain Model
Complex Media Exchange	Functional 2	Program Level	Desktop Automation	Local Networks	Program Model
Simple Electronic Exchange	Connected 1	Local/Site Level	Standard System Drivers	Simple Connection	Local
Manual Gateway	Isolated 0	Access Control	N/A	Independent	Private

Figura 6. Modelo de referencia LISI (LISI, 2018)

Los cinco niveles de interoperabilidad de LISI son los siguientes:

- Nivel 0. *Isolated* (Aislada): sin conexión.
- Nivel 1. *Connected* (Conectada): conexión electrónica, datos y aplicaciones separadas.
- Nivel 2. *Functional* (Funcional): funciones comunes mínimas, datos y aplicaciones separadas.
- Nivel 3. *Domain* (de Dominio): datos compartidos, aplicaciones separadas.
- Nivel 4. *Enterprise* (Empresarial): manipulación interactiva, datos y aplicaciones compartidas.

Según estos niveles de madurez hay que tener en cuenta varios aspectos que permiten que los sistemas alcancen la interoperabilidad:

- Una infraestructura para que las aplicaciones intercambien estructuras de datos e información.
- Las políticas y procedimientos que definen las reglas para la interoperabilidad e intercambio.

A cada nivel, el modelo LISI le atribuye un perfil que se compone de diversos atributos. Estos atributos los define en su paradigma PAID (*Procedures, Applications, Infrastructure y Data*):

- Procedimientos (*Procedures*): se centran en las distintas formas de orientación que tienen algún impacto en la interoperabilidad de sistemas, incluyendo doctrina, misión, arquitecturas y estándares.
- Aplicaciones (*Applications*): representan los aspectos funcionales de un sistema. Estas funciones se manifiestan en los componentes software de los sistemas, desde procesos individuales a conjuntos de aplicaciones integradas.
- Infraestructura (*Infrastructure*): define el rango de componentes que hacen posible las interacciones entre sistemas, incluyendo hardware, comunicaciones, servicios de sistemas y seguridad. Por ejemplo, las infraestructuras tienen en cuenta los protocolos, los servicios software, y otras estructuras de datos de soporte al flujo de información entre las aplicaciones y los datos.

- Datos (*Data*): incluyen los formatos de datos y estándares que soportan la interoperabilidad a todos los niveles. Esto incluye todo el rango de estilos y formatos desde texto simple a modelos de datos de empresa.

Cada nivel de interoperabilidad se caracteriza por la forma de conexión entre los sistemas y el perfil establecido en relación a los atributos definidos en el Paradigma PAID (figura 7):

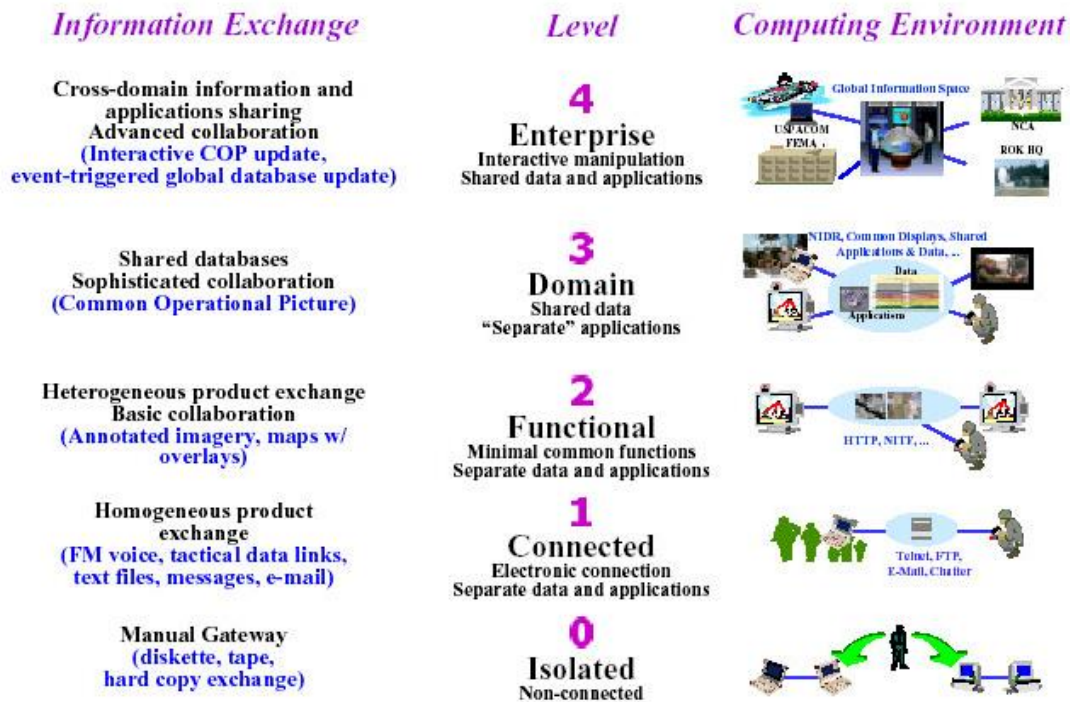


Figura 7. Niveles de interoperabilidad LISI (Morris et al., 2004)

- Nivel 0. Aislada: sin conexión.

Los sistemas de Nivel 0 no tienen conexión electrónica directa. El intercambio de datos entre estos sistemas se produce mediante una entrada manual con el teclado o con un sistema de almacenamiento común.

Los atributos PAID a nivel 0, son:

- Procedimientos: los sistemas establecen procedimientos locales para gestionar el control de acceso. El usuario debe acceder al sistema directamente para compartir información con otros sistemas.

- Aplicaciones: funcionalidades independientes en sistemas aislados. Los datos resultantes son importantes, pero no se tiene en cuenta la capacidad de gestionar esos datos de forma consistente.
 - Infraestructura: principalmente independiente entre sistemas. La mayor parte del intercambio de información se lleva a cabo mediante acceso físico directo. El intercambio de datos entre los sistemas aislados se lleva a cabo mediante el uso de medios de almacenamiento como disquetes, etc.
 - Datos: modelos de datos propietarios.
- Nivel 1. Conectada: conexión electrónica, datos y aplicaciones separados.

Los sistemas de Nivel 1 están conectados de forma electrónica. Estos sistemas permiten el intercambio peer-to-peer de tipos homogéneos, tales como texto, ficheros gráficos (por ejemplo, imágenes GIF, TIFF, etc.). Generalmente los sistemas de Nivel 1 permiten el intercambio de ficheros de forma simple.

Los atributos PAID a nivel 1, son:

- Procedimientos: más allá del simple control de acceso, la mayoría siguen refiriéndose principalmente a políticas a nivel local.
 - Aplicaciones: independientes entre sistemas, pero con mecanismos e interfaces comunes.
 - Infraestructura: soportan conexiones simples tipo peer-to-peer para permitir la transferencia de datos locales en consonancia con los procedimientos establecidos localmente.
 - Datos: pueden existir modelos de datos locales, pero por lo general son específicos de un programa en particular. Los informes simples y los gráficos son un ejemplo.
- Nivel 2. Funcional: funciones comunes mínimas, datos y aplicaciones separados.

Los sistemas de Nivel 2 son sistemas distribuidos que residen en redes locales que permiten transferir conjuntos de datos complejos y heterogéneos (imágenes anotadas, mapas, documentos multimedia o hiper-enlazados, etc.) de un sistema a otro. Un aspecto clave de los sistemas o aplicaciones de este nivel es su capacidad para proporcionar acceso vía Web a los datos.

En estos sistemas están presentes formatos de modelos de datos (lógicos y físicos), pero generalmente existe un modelo de datos aceptado entre los distintos programas. Después, cada programa define su modelo de datos propio.

Generalmente, los usuarios son capaces de compartir información integrada entre sistemas o funciones.

Los atributos PAID a nivel 2, son:

- Procedimientos: centrándose en el nivel de programas individuales, los centros de excelencia estarían encargados de especificar muchas de las implementaciones que los programas deben soportar.
 - Aplicaciones: las funciones incluyen automatización de escritorio y la capacidad de intercambiar algunos datos estructurados. Los programas de ofimática son un ejemplo. Las interfaces web son importantes en este nivel.
 - Infraestructura: los sistemas interactúan con otros sistemas en un área local mediante las LAN. Estas LAN pueden usar protocolos como TCP/IP para soportar redes de área amplia.
 - Datos: pueden existir estructuras avanzadas de datos, aunque se centran principalmente en aplicaciones individuales (modelos de datos de programas). El uso de formatos de datos comunes entre distintos programas cada vez es mayor.
- Nivel 3. Dominio: datos compartidos, aplicaciones separadas.

El nivel 3 se caracteriza por las múltiples interacciones entre aplicaciones. Los sistemas y aplicaciones de nivel 3 son sistemas integrados, capaces de interconectarse mediante redes de área amplia (WAN) que permiten a múltiples usuarios acceder a los datos. En este nivel las implementaciones generalmente tienen una visión localizada de un espacio de información distribuido y a través de un dominio operativo o funcional, de manera que la información es compartida por aplicaciones independientes.

Está presente un modelo de datos (lógico y físico) basado en el dominio, que es entendido, aceptado e implementado en un área funcional o grupo de organizaciones que componen un dominio. Los sistemas son capaces de implementar reglas de negocio y procesos para facilitar las interacciones directas entre bases de datos, como aquellas requeridas para soportar la réplica de bases de datos en distintos servidores.

En este nivel, las aplicaciones individuales pueden compartir repositorios de datos centrales o distribuidos. Los sistemas soportan el trabajo en grupo y colaboración basada en productos de información agrupados (fusionados). Generalmente, se facilita la toma de decisiones sobre problemas localizados en un dominio específico.

Los atributos PAID a nivel 3, son:

- Procedimientos: se centran en la interacción de dominio, donde un dominio puede abarcar varias áreas geográficas, pero se centra en un área funcional (C2, inteligencia, logística).
 - Aplicaciones: avanzadas más allá de los programas individuales, se soportan capacidades básicas de trabajo en grupo, como el seguimiento de las revisiones de documentos o de gestión del flujo de trabajo.
 - Infraestructura: las redes son globales. En este nivel la interacción tiene lugar en el espacio global de información, aunque no todo.
 - Datos: existen modelos de datos definidos que son entendidos entre las aplicaciones, aunque sólo representan un dominio particular.
- Nivel 4. Empresarial: manipulación interactiva, datos y aplicaciones compartidas.

El nivel 4 es el objetivo último de los sistemas de información que persiguen la interoperabilidad, en el que los sistemas son capaces de operar empleando un espacio de información global distribuido a través de distintos dominios. Múltiples usuarios son capaces de interactuar con datos complejos de forma simultánea. Los datos y las aplicaciones son completamente independientes y pueden ser distribuidas en este espacio para soportar la integración o agrupación de información. Este nivel permite formas avanzadas de colaboración (concepto de oficina virtual). Los datos tienen una interpretación común sin tener en cuenta su forma, y se aplican de igual forma en toda la organización. La necesidad de aplicaciones redundantes, con funcionalidades equivalentes, se ve disminuida, dado que las aplicaciones pueden ser compartidas al igual que los datos. La toma de decisiones tiene lugar en contexto, y es facilitada por la información global de la organización que se encuentra en este espacio de información común.

Los atributos PAID a nivel 4 son:

- Procedimientos: procedimientos en el ámbito de la empresa.
- Aplicaciones: integradas en el espacio de información común distribuido. Múltiples usuarios pueden acceder a las mismas instancias de datos de toda la empresa u organización.
- Infraestructura: redes globales que soportan topologías multidimensionales. Estas redes pueden tener diferentes áreas basadas en seguridad o control de acceso, pero que están apropiadamente integradas para soportar las necesidades de los usuarios.
- Datos: modelos de datos de empresa que soportan la integración entre aplicaciones. Existe un entendimiento común de los datos a través de la organización.

2.2 La interoperabilidad en las Administraciones Públicas

2.2.1 Modelos de interoperabilidad en las Administraciones Públicas

Es posible diferenciar entornos de aplicación de la interoperabilidad por el sector de las tecnologías al que pertenecen: al sector privado, al sector público o a ambos. En particular, los entornos se caracterizan por el alcance político y geográfico y la cantidad de sistemas y servicios implicados, puesto que la interoperabilidad no es aplicable solamente a un dominio específico, a un sector o a nivel local o nacional, sino que puede y debería tener un alcance global.

Cuando hablamos de interoperabilidad entre administraciones públicas surgen una serie de necesidades y beneficios. Entre las necesidades podemos encontrar:

- Cooperar entre las administraciones tendentes al establecimiento de los servicios públicos.
- Intercambiar información entre las administraciones para cumplir con los requisitos legales o compromisos políticos.
- Compartir y reutilizar información, tanto internamente, como entre las administraciones para aumentar la eficiencia administrativa y reducir las cargas administrativas para ciudadanos y empresas.

Cuando se afrontan estas necesidades conllevan a los siguientes beneficios:

- Mejorar la prestación de servicios públicos a ciudadanos y empresas, facilitando la entrega de ventanilla única de servicios públicos.
- Reducir los costes para las administraciones públicas, empresas y ciudadanos a través de la entrega eficiente y eficaz de los servicios públicos.

En este sentido, existen varios intentos de clasificar y definir los modelos aplicables en el ámbito de la interoperabilidad, que tienen como objetivo categorizar e incluso medir el alcance de los distintos entornos en los que se puede establecer la interoperabilidad.

Modelo Carnegie Mellon

El Instituto de Ingeniería de Software de la Universidad Carnegie Mellon propone un modelo para entender la interoperabilidad en el ámbito del *e-Government*

(Novakouski, 2012). Se basa en el modelo LISI visto en la sección anterior, pero también en modelos tales como OIMM, LCIM, NATO C3, EIF y GIF. Todos ellos se centran en términos de objetivos, tipos y niveles de interoperabilidad. El modelo finalmente propuesto por la Universidad Carnegie Mellon parte de la definición de los objetivos generales de interoperabilidad, para a continuación, situarlos en determinados niveles de interoperabilidad y finalizar añadiendo una serie de factores de influencia.

Los objetivos de interoperabilidad que propone son los siguientes:

- **Intercambio de datos:** Como primer objetivo de interoperabilidad tenemos el intercambio básico de datos, no desde la perspectiva de su significado sino de la mera transmisión de los mismos.
- **Intercambio de significado:** En el intercambio de significado todos los participantes en una comunicación asignan el mismo significado a la información que intercambian.
- **Acuerdo en los procesos:** En este objetivo de interoperabilidad se llega a un acuerdo sobre cómo actuar con la información intercambiada, es decir, cada participante de la comunicación tiene el mismo conocimiento de las acciones a realizar una vez recibida la información.

Los niveles de interoperabilidad introducidos en el capítulo 1 guardan una serie de relaciones entre ellos. En este modelo, la relación es la que se refleja en la figura 8:

- **Interoperabilidad técnica:** Relacionada con el objetivo de intercambio de datos, base de cualquier comunicación. En algunos modelos técnicos de interoperabilidad, este nivel se divide en sub-niveles que se asignan a modos específicos de comunicación y que separan los datos del canal de comunicación. El enfoque adoptado en los actuales modelos de interoperabilidad de *e-Government* es abstraer los detalles de la comunicación y tener un único nivel, que es el enfoque seleccionado.
- **Interoperabilidad semántica:** Relacionada con el objetivo de intercambio de significado, situado por encima del técnico debido a que es necesario que la información haya sido intercambiada.
- **Interoperabilidad organizativa:** Relacionada con el objetivo de acuerdo en los procesos, situado en el nivel más alto debido a la necesidad de cumplir los otros objetivos: intercambio y significado de la información. De esta forma establecemos la comunicación y los procesos para que el destinatario actúe en consecuencia.

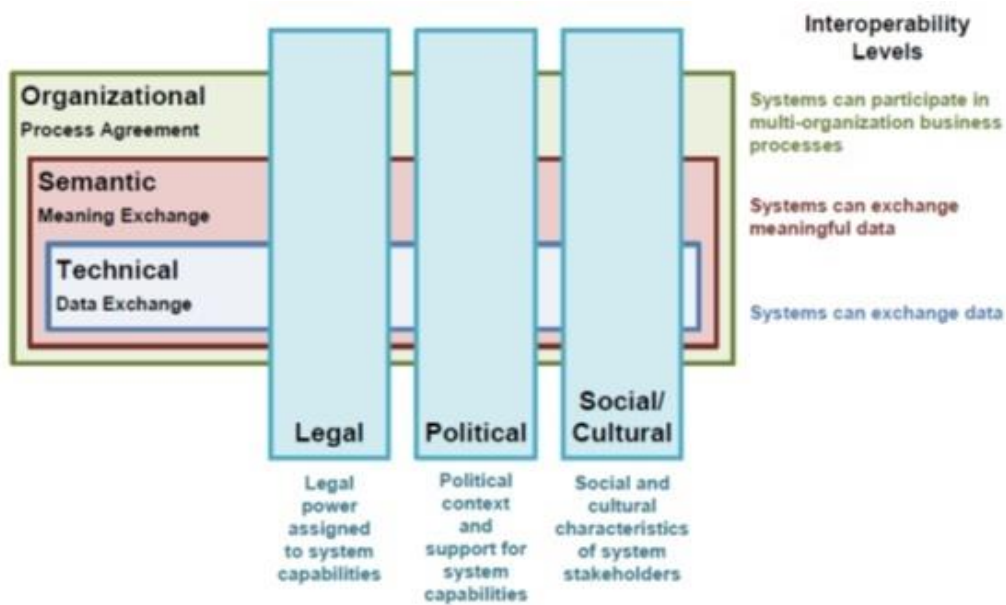


Figura 8. Relación entre los niveles de interoperabilidad LISI (LISI, 2018)

Para finalizar, dicho documento plantea una serie de factores de influencia:

- **Factores legales.**

Una de las principales preocupaciones desde el punto de vista legal es la asignación de valor legal a las salidas de un determinado sistema de *e-Government* para que el sistema pueda apoyar o sustituir a los servicios públicos existentes. La identificación de cuestiones legales inherentes a los servicios públicos es fundamental para permitir identificar qué servicios son los adecuados en el gobierno electrónico. Otra área importante en los factores legales es la asignación de la responsabilidad. Todos los sistemas de administración electrónica deben cumplir plenamente con las leyes, normas y regulaciones. Se pueden incluir: normativas en materia de derecho administrativo, derechos de propiedad intelectual, así como cuestiones de privacidad y protección de datos.

- **Factores políticos.**

Una administración pública que tenga la voluntad política y el poder de organizar, administrar y financiar un proyecto de interoperabilidad de gobierno electrónico tendrá una mejor oportunidad de éxito. Sin este elemento fundamental, puede ser difícil alcanzar el nivel de cooperación y coordinación entre los departamentos gubernamentales que participan en el proyecto para hacer frente a los niveles técnicos, semánticos y organizativos.

- **Factores socioculturales.**

Las fuentes potenciales de impacto en este ámbito varían ampliamente, dependiendo del contexto de aplicación del proyecto. Por ejemplo, en diversas regiones puede ser necesario prestar servicios públicos a través de diferentes canales o canales personalizables. El modelo EIF (EC, 2004) propone un "acceso multicanal" para compensar las desigualdades económicas y el multilingüismo como un factor clave para la creación de interfaces de servicios públicos. Los diseñadores pueden desarrollar servicios que respondan a los objetivos del *e-Government* y de las necesidades de los ciudadanos, pero si los usuarios finales no consideran que son accesibles y usables, no podrán adoptar estos servicios, desaprovechando su potencial.

Modelo LCI

El Ministerio de Defensa australiano ha extendido el modelo LISI hacia el concepto de interoperabilidad organizativa (Clark, 1999). Este Ministerio concibe una dimensión diferente para la interoperabilidad de la de datos, proponiendo una estructura paralela a la LISI, pero con el énfasis en la interacción humana. Está formado por cinco niveles (figura 9):

1. **Nivel 0 – Independiente:** Describe la interacción entre organizaciones independientes. Se trata de organizaciones que normalmente trabajan sin ningún tipo de interacción distinta a la prevista por el contacto personal. Lo más probable es que las organizaciones normalmente no compartan los mismos objetivos o fines, pero puede ser necesario interactuar en algunos escenarios.
2. **Nivel 1 - Ad hoc:** En este nivel de interoperabilidad los *frameworks* de la organización son muy limitados. pero pueden dar lugar a acuerdos ad hoc. Se contará con algunas pautas para describir cómo se producirá la interoperabilidad, pero en esencia los acuerdos específicos no se han planificado. Se habrá fijado algún objetivo general compartido, pero las aspiraciones de las organizaciones siguen siendo individuales.
3. **Nivel 2 - Colaboración:** Los *frameworks* de las organizaciones son reconocidos y están presentes para apoyar la interoperabilidad de forma que se reconocen los objetivos y los roles y las responsabilidades se asignan como parte de las actuales responsabilidades; sin embargo, las organizaciones siguen siendo independientes. Se producirá una comunicación e interconexión significativa entre las organizaciones con intercambio de conocimiento, pero los *frameworks* de las propias organizaciones todavía tienen una influencia significativa.
4. **Nivel 3 - Integrado:** Se comparten los sistemas de valor y las metas, un entendimiento común y la disposición a interactuar. *Los frameworks* están en su

lugar y se practican, sin embargo, aún hay archivos residuales a organizaciones de origen.

5. **Nivel 4 - Unificado:** Una organización unificada es aquella en la que los objetivos de la organización, los sistemas de valores, la estructura de mando y las bases de conocimientos son compartidos por el sistema. Este es realmente el nivel ideal donde no existe ningún impedimento en el marco organizativo para la interoperabilidad plena y completa.

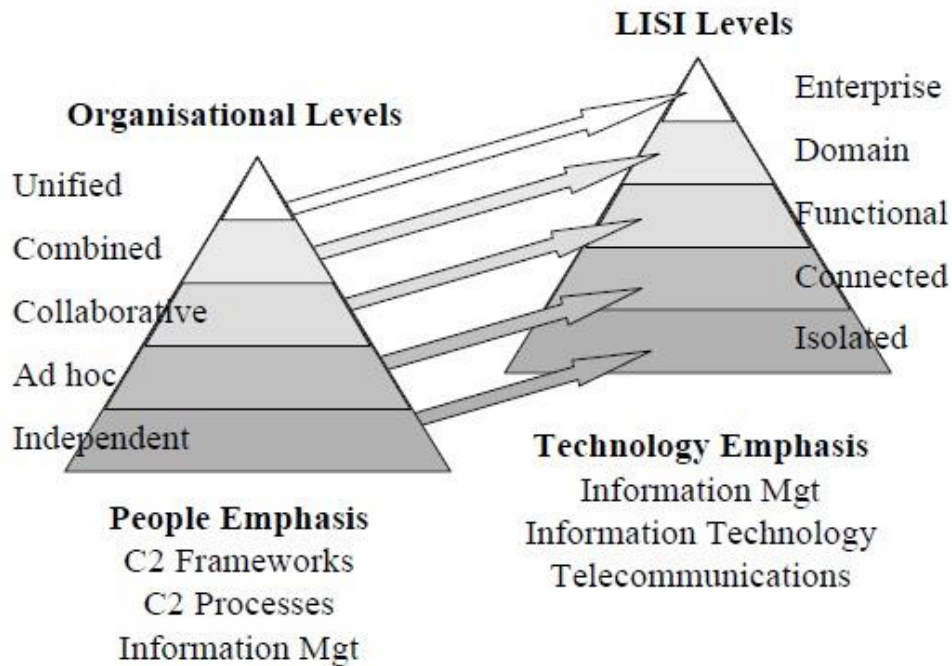


Figura 9. Relación entre los modelos LCI y LISI (Clark, 1999)

El modelo LCI abarca un amplio rango de características técnicas y organizativas que deben estar alineadas para que los sistemas de coalición puedan interoperar (Tolk, 2003). No pretende ser un sustituto de otros modelos. Los cuatro niveles más bajos del modelo pertenecen a la interoperabilidad técnica. El nivel de conocimiento proporciona una transición entre la interoperabilidad técnica y la interoperabilidad organizativa, que está representado por los cuatro niveles superiores (figura 10).

Para alcanzar los valores máximos de interoperabilidad del modelo habría que alcanzarlos en cada uno de los niveles en que se divide. Con el fin de evaluar cada uno de los niveles, el modelo propone una serie de cuestiones que deben responderse para comprobar su nivel de interoperabilidad.

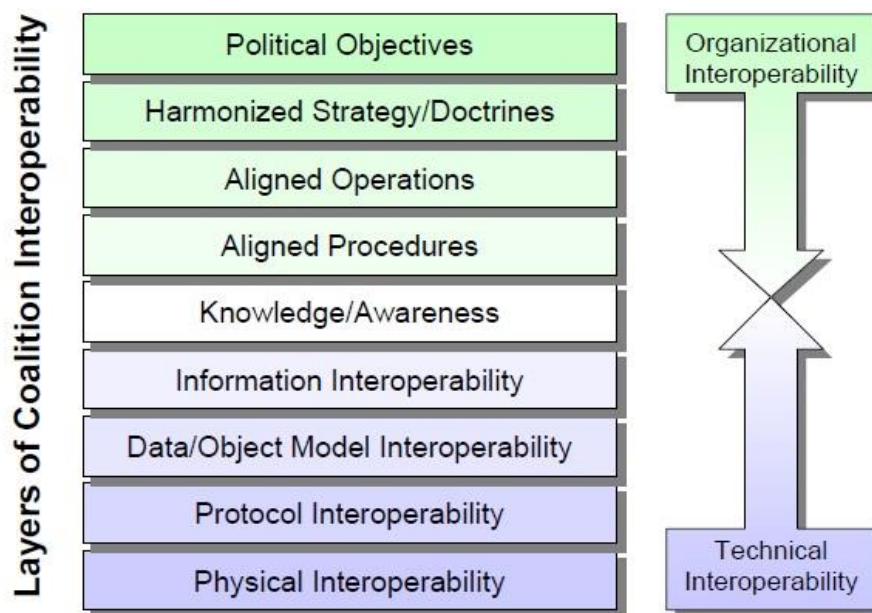


Figura 10. Capas de *Coalition Interoperability* (LCI, 2003)

2.2.2 La interoperabilidad en las Administraciones Públicas: casos y referencias

El interés del sector público por la interoperabilidad se remonta a los años ochenta (Guijarro, 2007). En un principio este interés se relacionaba con los problemas que planteaban los sistemas propietarios. La estandarización fue la primera propuesta para afrontar estos problemas. Así podríamos mencionar que en 1984 la Organización Internacional de Normalización (*International Organization for Standardization - ISO*) produjo el modelo de referencia básico de la interconexión de sistemas abiertos, conocido como "*Open Systems Interconnection (OSI) Reference Model and Standards*". Hasta ese momento, la información existente en sistemas de red estaba generalmente desarrollada en sistemas propietarios y la interoperabilidad entre estos sistemas no existía. Por otra parte, había una demanda de estándares para facilitar los procesos y aplicaciones de cooperación independientemente de las plataformas.

En 1988 el Instituto Nacional de Normalización y Tecnología (*National Institute of Standards and Technology (NIST)*) en Estados Unidos, aprobó el Perfil de interconexión de sistemas abiertos del Gobierno (*Government Open Systems Interconnection Profile*), (GOSIP, 1991), basándose en un reconocimiento tanto por parte del sector público como del sector privado de la necesidad de desarrollar un conjunto común de protocolos basados en los estándares OSI, haciendo referencia a los procesamientos

automáticos de datos intercambiados entre agencias gubernamentales como una de las razones para establecer estos protocolos.

Efectivamente, el gobierno federal de los Estados Unidos mostraba un fuerte compromiso con OSI y GOSIP. GOSIP iba a ser utilizado por las agencias federales listas para adquirir redes OSI. Incluso el Departamento de Defensa (*Department of Defense - DoD*) tomó la iniciativa de solicitar GOSIP en su futura adquisición de red. En 1987 el DoD marcó un hito al cambiar del protocolo TCP/IP (*Transmission Control Protocol/Internet Protocol*) a OSI. TCP/IP era el protocolo base de Internet. Durante dos años TCP/IP y OSI fueron co-estándares. Después de dos años, los protocolos OSI serían utilizados en las adquisiciones. Algunos años después, sin embargo, en 1995 el NIST aprobó la revisión de GOSIP con el título "*Profiles for Open Systems Internetworking Technologies*" (POSIT, 1995). En POSIT eran revisados los estándares ampliándolos para agencias, facilitándoles la adquisición de productos de red que implementasen estándares abiertos. Tales estándares incluían los desarrollados no sólo por ISO, sino también por el "*Internet Engineering Task Force*" (IETF).

Se intentó que GOSIP fuera una guía de compras para los departamentos del gobierno para asegurar que los sistemas adquiridos por separado fueran capaces de interactuar. En el Reino Unido, su versión de GOSIP se orientaba más hacia las aplicaciones de usuario y a proporcionar asistencia técnica que auxiliase a los usuarios en los procesos de compra de las aplicaciones de escritorio.

Otro hito importante lo marca la Unión Europea al desarrollar en 1993 el Manual Europeo de Compras de Sistemas Abiertos (EU, 1993). Se basó fundamentalmente en las especificaciones del GOSIP inglés con contribuciones de Francia y Alemania.

Por regla general, en Europa se puso mucho énfasis en la definición de perfiles estandarizados, de forma similar al GOSIP de Estados Unidos, especialmente porque el mercado europeo se caracterizaba por un mayor número de fabricantes de ordenadores que Norte América, lo que se traducía en mayores dificultades de interconexión. El otro gran impulso era necesario para la interconexión entre Estados miembros (Hartmann, 1990).

El surgimiento de la administración electrónica (*e-Government*) ha colocado de nuevo todas estas preocupaciones en la agenda de las administraciones públicas. Podría decirse que en estos momentos el problema tiene dimensión mundial. Las nuevas vías de prestación de servicios públicos incluyen la aproximación *customer-centric*, intentando ocultar la complejidad de los procedimientos administrativos e incluyendo un alto grado de interacción entre las administraciones local, regional, nacional e internacional. En este escenario, la interoperabilidad es claramente un aspecto clave y tiene que ser puesta de manifiesto como un principio para la concepción y el desarrollo de las iniciativas del gobierno.

A continuación, se recogen las iniciativas importantes de interoperabilidad en el terreno de la administración electrónica.

- En el Reino Unido, la Unidad de *e-Government* publicó en el año 2000 su guía técnica en el “*e-Government interoperability Framework*” (e-GIF), actualizada posteriormente en 2005 (e-GIF, 2005). e-GIF establece un conjunto de especificaciones y políticas para la colaboración entre agencias y la prestación de servicios. Cubre cuatro áreas: interconexión, integración de datos, acceso a los servicios y gestión del contenido. Asimismo, incluye un catálogo de estándares técnicos que es revisado y actualizado cada seis meses. Varios países siguieron este modelo, estableciendo catálogos de estándares para su uso en los respectivos sistemas de administración electrónica. Estos catálogos permitieron establecer pautas comunes en las tecnologías utilizadas por diferentes agencias o ministerios.
- En Francia, la Agencia para el Desarrollo de la Administración Electrónica (*Agence pour le développement de l’administration électronique* (ADAE)) publicó “*Le cadre commun d’interopérabilité*” (CCI) en enero de 2001 y su última versión (2.1) en septiembre de 2003, ya descatalogado. CCI comprendía las recomendaciones para reforzar la coherencia de los sistemas públicos electrónicos y para facilitar la prestación de servicios electrónicos multi-agencia. La ADAE desapareció en enero de 2006 al crearse la “*Direction Générale de la modernisation de l’Etat*”, en el seno del “*Ministère de l’économie, des finances et de l’industrie*”, que reagrupaba varias estructuras que se ocupaban de temas relacionados con la modernización del Estado en general y la administración electrónica en particular.
- En la Unión Europea hay que destacar el programa IDABC. Por lo que se refiere al marco de interoperabilidad, el programa IDABC publicó su Guía de Arquitecturas (versión 4.1) en marzo de 1999, y la herramienta de apoyo para la Decisión del Parlamento y el Consejo 1720/1999/EC “*Interoperability and Access to Trans-European Networks for the electronic Interchange of Data between Administrations*”. La versión 7.1 fue publicada en septiembre de 2004. Estas guías proporcionan conceptos y referencias para una óptima interoperabilidad entre las instituciones europeas, las agencias y los gobiernos de los estados miembros. Más adelante, IDABC publicó la versión final 1.0 del “*European Interoperability Framework*” (EIF, 2004). Su versión 2 fue aprobada por la Comisión Europea en la comunicación “*Towards interoperability for European public services*” el 16 de diciembre de 2010 (EIF, 2010).

El programa “Soluciones de interoperabilidad para las administraciones públicas europeas” (ISA) (2010-2015), y su sucesor, el programa ISA (2016-2020) , son los principales instrumentos mediante los cuales se han aplicado la Estrategia Europea de Interoperabilidad y el Marco Europeo de Interoperabilidad actuales (COM, 2017).

- En Estados Unidos, la *Federal Chief Information Officers Council* (CIOOC) publicó el “*Federal Enterprise Architecture Framework*” (FEAF) en septiembre de 1999. Para aumentar la influencia de la guía FEAF en las implementaciones de *e-Government*,

la CIOC incluyó en julio de 2002 en la guía de proyectos del gobierno federal. La versión más reciente es de 2013 (FEAF, 2013).

- En España se ha desarrollado el “Esquema Nacional de Interoperabilidad” (ENI) (ENI, 2007). Su objeto es abarcar el conjunto de criterios y recomendaciones en materia de seguridad, conservación y normalización de la información, de los formatos y de las aplicaciones que deberán ser tenidos en cuenta por las Administraciones Públicas para la toma de decisiones tecnológicas que garanticen la interoperabilidad.
- En Nueva Zelanda, el “*New Zealand e-Government Interoperability Framework*” (NZ e-GIF, 2008) establece un conjunto de políticas, estándares técnicos y directrices. Cubre las distintas formas de conseguir la interoperabilidad de datos y recursos de información del sector público, tecnologías de la información y de las comunicaciones y procesos electrónicos. Permite a cualquier agencia unir o compartir sus recursos informáticos con los de cualquier otra agencia utilizando un marco predeterminado basado estándares internacionales abiertos.
- Honk Kong también cuenta con un marco de interoperabilidad que define un conjunto de especificaciones orientadas a facilitar la interoperabilidad entre sistemas y servicios que sirven de referencia para el desarrollo de proyectos en el campo del *e-Government*. Las especificaciones técnicas se revisan cada seis o doce meses. La versión más reciente (v16.0) es de 2017 (IF, 2017).
- En septiembre de 2007, la Comisión para América Latina y el Caribe de las Naciones Unidas (CEPAL), publicó el “Libro blanco de interoperabilidad de gobierno electrónico para América Latina y el Caribe” (CEPAL, 2007). El libro se alimentaba de las experiencias y proyectos de interoperabilidad en Brasil, Chile, Colombia, México y Trinidad y Tobago, así como en la experiencia en ese campo de la Unión Europea, a la que cita como principal referente.
- También en el ámbito Iberoamericano, merece destacarse el documento “Bases para una Estrategia Iberoamericana de Interoperabilidad”, presentado en la “XII Conferencia Iberoamericana de Ministros de Administración Pública y Reforma del Estado”. En la IX Conferencia se adoptó la Carta Iberoamericana de Gobierno Electrónico (Criado et al., 2010). Significó un hito en la construcción de una perspectiva regional para el desarrollo del Gobierno electrónico, citándose explícitamente la interoperabilidad como uno de los aspectos claves a tener en consideración.

El documento presenta las Bases para una Estrategia Iberoamericana de Interoperabilidad. Define unos fundamentos básicos para facilitar a los gobiernos y administraciones públicas de la región el uso de las tecnologías de la información y las comunicaciones para mejorar su desempeño, ofrecer nuevos servicios públicos

o reforzar los procesos de gobernanza de las políticas públicas que implican a dos o más Estados.

- Por último, los Estados Unidos y Canadá han tomado distintas iniciativas sobre interoperabilidad en el ámbito del e-Government, apoyando un marco arquitectura empresarial como el “*Zachman Framework*” (Zachman, 2008) y el “*Open Group Architecture Framework*” (TOGAF, 2008). El gobierno de los Estados Unidos también proporciona una serie de guías sobre marcos de arquitecturas actualmente en uso en estos proyectos (OMB, 2011).

2.3 La interoperabilidad entre las organizaciones de Seguridad Social

La Seguridad Social se considera por regla general como parte de la Administración Pública en todos los países. Sin embargo, tanto su repercusión social como las características propias de los procesos que maneja, hace que merezcan una atención especial y tratamiento diferenciado tanto a nivel nacional como internacional.

2.3.1 Modelo de aplicación de interoperabilidad para la Seguridad Social

La AISS presentó en la Conferencia de Brasilia (ISSA, 2012) un extenso informe sobre la aplicación de la Interoperabilidad en la Seguridad Social. El informe era parte del proyecto “Las TIC como facilitadoras de la integración de políticas y programas de Seguridad Social”, que llevó adelante la Comisión Técnica de la AISS sobre las Tecnologías de la Información y las Comunicaciones durante el trienio 2011-2013. En ese informe, la AISS propuso un modelo de aplicación de interoperabilidad genérico para sugerir formas de organizar la creación y el funcionamiento de servicios públicos en instituciones de Seguridad Social. El modelo se basa en el propuesto en el EIF v2.0 en el marco de la Unión Europea donde se hace un estudio de la aplicación de los servicios públicos en los estados miembros, con la intención de ayudar a desarrollar un modelo de referencia, así como un vocabulario común y la comprensión de todas las organizaciones implicadas sobre los principales elementos para implementar un servicio público inter-institucional.

El modelo utiliza el concepto de “bloques de construcción” para la creación de los servicios públicos, utilizando componentes inter-conectables y reutilizables en la construcción de nuevos servicios. Dado que está fuertemente orientado hacia la implementación de servicios futuros, no todos los servicios existentes encajan en este modelo.

Su carácter es genérico, en el sentido de que es aplicable en cualquier nivel de prestación de servicios públicos, desde el local hasta el internacional. También puede ser utilizado para clarificar y racionalizar las relaciones entre las entidades que colaboran para ofrecer mejores servicios públicos, en los diferentes niveles de gobierno o en los diferentes sectores.

En síntesis, el objetivo del modelo es lograr beneficios prácticos para el establecimiento de servicios públicos de forma que se definan claramente las funciones de los servicios, sus interfaces y se permita su reutilización. De esta forma se evita duplicar esfuerzos tratando de utilizar servicios existentes y automatizando una gran variedad de prácticas básicas en el uso de estos servicios.

Conceptos clave

El modelo promueve la reutilización de la información, conceptos, modelos, patrones, soluciones y especificaciones contenidas en cada una de las instituciones de la Seguridad Social. Sus servicios públicos se basan en información de diversas fuentes ubicadas en los diferentes niveles de la administración y en distintos estados. Los servicios públicos básicos construidos de forma independiente por las instituciones de la Seguridad Social en los diferentes estados se deben poder combinar.

Asimismo, pone de relieve la necesidad de una infraestructura modular, con componentes débilmente acoplados que permitan la cooperación para la prestación de los servicios. Se propone la adopción de arquitecturas orientadas a servicios (SOA) para el diseño y desarrollo de los sistemas de las instituciones de forma que se empaqueten los procesos de negocio en servicios.

Según esta propuesta, las instituciones de la Seguridad Social deberían reconfigurar sus sistemas y aplicaciones para aumentar la reutilización y satisfacer las nuevas necesidades de sus usuarios. Además, deberían integrar acuerdos a nivel de servicio y políticas de operación a nivel de componentes. También necesitarían un acuerdo sobre un régimen común para interconectar los componentes de servicios débilmente acoplados y poner en marcha la infraestructura necesaria a la hora de establecer sus servicios públicos.

Capas del Modelo

El modelo se ha dividido en tres capas: los servicios públicos básicos, el intercambio seguro de datos y la agrupación de los servicios públicos, que se pueden apreciar en la figura 11.

Capa 1: servicios públicos básicos

La capa más baja del modelo se refiere a los componentes de los servicios más básicos con los cuales los servicios públicos de la Seguridad Social se pueden construir. Agrupa tres tipos de componentes: servicios basados en registros base, facilitadores de interoperabilidad y servicios externos.

Algunos de los servicios públicos básicos se han desarrollado principalmente para su uso particular en la administración pública que los creó, o por sus clientes directos, es decir, las empresas y los ciudadanos; pero se ponen a disposición para su reutilización en otros lugares con el fin de proporcionar servicios públicos globales. Otros son genéricos o de infraestructura por naturaleza, mientras que el resto representan los servicios externos, es decir, servicios prestados por terceros. A continuación, se describen con más detalle cada tipo de servicio público básico.

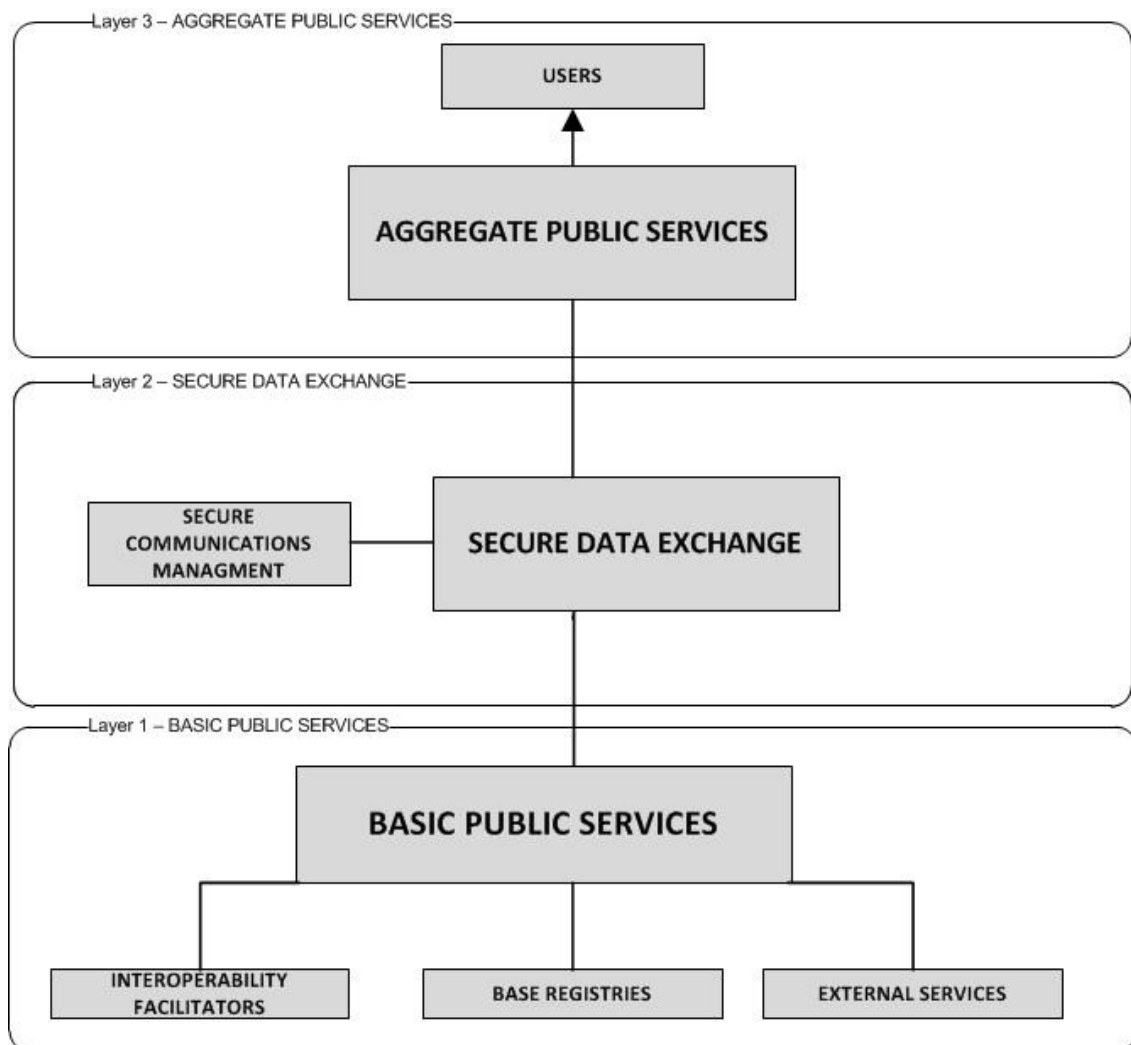


Figura 11. Modelo conceptual de Interoperabilidad para la Seguridad Social (ISSA, 2012)

- **Registros básicos o recursos de datos compartidos.**

Los registros básicos o recursos de información compartidos, (fuentes confiables de información básica) representan datos centrales (“core”) al área de Seguridad Social. Estos registros están bajo el control legal de las instituciones de la Seguridad Social, pero podrían estar disponibles para su reutilización aplicando mecanismos de seguridad y privacidad apropiados.

Los registros básicos pueden presentarse de formas variadas, pero la característica principal es que están validados y autorizados por la administración correspondiente. Esto significa que tales repositorios de datos contienen la información pertinente, de referencia, (no necesariamente toda la información) sobre los ciudadanos, empresas, beneficios, etc., que se considera correcta. En términos generales, el contenido no es estático, se debe administrar el ciclo de vida de la información, lo cual, en un contexto de datos compartidos, supone un desafío.

El acceso a fuentes de datos auténticas entre diversas administraciones se verá facilitado si las interfaces con estas fuentes se publican y armonizan, tanto a nivel semántico como técnico. Estos datos pueden provenir de sistemas legados (*legacy systems*), lo que puede representar un obstáculo en la adopción del modelo ya que estos sistemas y sus repositorios de datos tienen unas características concretas que limitan sus posibilidades y pueden requerir de una amplia reingeniería con el fin de hacer su información disponible para los servicios públicos.

A los efectos de construir este tipo de recursos de información, las instituciones deben evaluar sus servicios existentes y los procesos de negocio asociados, con el fin de identificar y documentar los elementos de los servicios o procesos de negocio, incluyendo las funciones de negocio llevadas a cabo por las aplicaciones, los datos utilizados por los procesos, los servicios prestados por los sistemas y las tareas realizadas por personas.

Siguiendo con el enfoque modular planteado en el modelo para la construcción de servicios, se deben implementar bloques de construcción que impliquen el establecimiento de ciertos niveles de confianza para actividades específicas de interacción, acompañado por la introducción de acuerdos negociados entre las administraciones. Estos pueden adoptar la forma de memorandos de entendimiento y acuerdos de nivel de servicio, asociados con la idea de interfaces bien definidas en las que otros componentes (los usuarios del servicio) pueden confiar.

- **Facilitadores de interoperabilidad**

Estos son los mecanismos básicos sobre los cuales se construyen los servicios y su interacción de acuerdo a las necesidades de interoperabilidad. Los más extendidos son: los traductores de protocolos (*gateways*), los traductores de formato (p.ej.

XSLT) y las plataformas de integración (p.ej. *Integration Brokers* y *Enterprise Service Bus*).

- **Servicios externos**

Estos incluyen los servicios prestados por entidades externas, como empresas u organismos que pueden ser incluidos en el alcance de los servicios. Pueden ser:

- A nivel de información: datos sobre estado civil de personas proporcionados por otras oficinas del Estado.
- A nivel de negocio: los servicios de pago proporcionados por las instituciones financieras.
- A nivel de infraestructuras: servicios de conectividad proporcionados por los proveedores de telecomunicaciones.

Capa 2: intercambio seguro de datos

Esta capa es fundamental para el modelo conceptual ya que todo el acceso a los servicios públicos básicos pasa a través de ella.

Desde el punto de vista empresarial, las administraciones y otras entidades intercambian información oficial que puede implicar el acceso a los registros base. Esto debe pasar por una capa segura, armonizada y controlada que permita el intercambio de información entre administraciones, empresas y ciudadanos, que establece:

- Firmado y certificado: tanto el emisor como el receptor han sido identificados y autenticados por mecanismos seguros.
- Cifrado: la confidencialidad de los datos intercambiados está asegurada.
- Registrado: los registros electrónicos son archivados para garantizar una auditoría legal.

En el modelo conceptual propuesto, estas funciones se agrupan en la capa "intercambio seguro de datos". Esta capa debe permitir el intercambio seguro de mensajes certificados, registros, formularios y otro tipo de información entre los diferentes sistemas. Además del transporte de datos, en esta capa también se deben manejar los requisitos específicos de seguridad, como la firma electrónica, la certificación, el cifrado y sellado de tiempo (*timestamping*).

La seguridad es potencialmente uno de los principales obstáculos a la interoperabilidad, si no se aplica de un modo armonizado y acordado entre las organizaciones.

El modelo conceptual pide a todos los proveedores de servicios: examinar las cuestiones de seguridad; cooperar en un marco común para satisfacer sus necesidades de seguridad a través de mecanismos compatibles y especificaciones acordadas en común; así como llegar a un entendimiento común sobre las características esenciales, tales como los niveles de protección, niveles de autorización y autenticación. Por lo tanto, las administraciones públicas deben ponerse de acuerdo sobre un marco común de seguridad cuando se establece un servicio público.

Uno de los requisitos clave para la implementación de la funcionalidad que se espera en el intercambio seguro de datos consiste en el aprovechamiento de la identificación nacional y las infraestructuras de autenticación en las administraciones. Este plan debe establecer que las arquitecturas de las TIC y los datos son necesarios en un contexto de interoperabilidad para hacer las infraestructuras existentes de identidad electrónica interoperables.

La provisión de seguridad para el intercambio de datos (es decir, firmado, certificado, cifrado y registro) también requiere de varias funciones de gestión, incluyendo:

- Servicio de gestión: para supervisar todas las comunicaciones en materia de identificación, autenticación, autorización de transporte de datos, etc., incluyendo los derechos de acceso, revocación, y de auditoría.
- Servicio de registro: para ofrecer (sujeto a la autorización apropiada) el acceso a los servicios disponibles a través de su previa localización y verificación de autenticación.
- Servicio de registro: para asegurar que todos los intercambios de datos se registran para pruebas futuras, y se archivan cuando sea necesario.

Capa 3: agrupación de los servicios públicos

Los servicios compuestos (o agrupados) se construyen mediante la composición de servicios básicos a los que se pueden acceder de manera segura y controlada. Estos pueden ser proporcionados por varias administraciones a cualquier nivel, es decir, local, regional, nacional o internacional.

Un servicio agrupado típico debe aparecer a sus usuarios (administraciones, empresas y ciudadanos) como un único servicio, mientras que la ejecución de transacciones puede residir en diferentes unidades administrativas. La agrupación se lleva a cabo a través de mecanismos adaptados a las necesidades específicas del negocio, por ejemplo, a través de orquestaciones o flujo de trabajo.

Hoy en día, los usuarios esperan acceder a los servicios públicos no sólo a través de portales o sitios web de la administración, sino también a través de intermediarios con los que están en contacto sobre una base regulada. Por lo tanto, los servicios públicos deben desarrollarse de tal manera que se puedan integrar fácilmente en sitios web de intermediarios a través de mecanismos tales como *mash-ups* y *widgets*, sin perder la responsabilidad del gobierno sobre los servicios y con claros indicios de que los usuarios puedan distinguir la diferencia entre servicios públicos y privados.

Si la agrupación de servicios públicos es prestada por los intermediarios, las administraciones públicas deben establecer:

- Un proceso de autorización para determinar qué servicios públicos básicos pueden ser adjudicados a los intermediarios.
- Un proceso para la certificación de los intermediarios para establecer la confianza necesaria entre los usuarios y los proveedores de servicios.

2.3.2 Casos de interoperabilidad en la Seguridad Social

A pesar de que las peculiaridades de los sistemas de Seguridad Social son lo suficientemente significativas como para que se considere de forma independiente dentro de la Administración Pública, tienen una serie de características comunes que permiten pensar en un tratamiento sistemático y estandarizado.

Como se ha mencionado en el apartado anterior, la AISS presentó en la Conferencia de Brasilia un extenso informe sobre la aplicación de la Interoperabilidad en la Seguridad Social (ISSA, 2012). Este informe se adjuntó a la documentación de las Directrices de la AISS sobre las Tecnologías de la Información, presentadas en el FORO mundial de Doha (Catar) en noviembre de 2013. En todas estas actividades tuvo participación el autor de la tesis.

Tal como se recogía en dicho informe, en el ámbito de la Seguridad Social, las técnicas de interoperabilidad han sido aplicadas en escenarios diversos, que pueden agruparse en las siguientes categorías (ISSA, 2012):

- Programas de protección social y de servicios sociales.
- Sistemas integrados de seguro de salud.
- Recaudación de contribuciones de seguridad social y control del cumplimiento de las obligaciones.
- Mejora de la eficiencia institucional, de la calidad de la información y los servicios.

- Acuerdos internacionales de Seguridad Social.

Cabe aclarar que estas categorías de escenarios no son disjuntas, dado que hay casos que pueden ubicarse en más de una de ellas.

2.3.2.1 Programas de protección social y servicios sociales

Los programas de protección social, o también las redes de protección social, tienen como objetivo proveer beneficios económicos a quienes no contribuyen al sistema, principalmente los más pobres. Para la CEPAL, se consideran “mecanismos de articulación y ejecución de diversos programas sociales, que coordinan los alcances de la seguridad social tradicional, la provisión de servicios sociales y la oferta de programas asistenciales” (CEPAL, 2015). Si bien las características de los programas de protección social varían, es posible distinguir ciertos elementos comunes (Gómez, 2011):

- Se focalizan en la población más pobre, más desprotegida, vulnerable o críticamente pobre.
- Cubren una amplia gama de iniciativas, ya sean transferencias monetarias, seguridad alimentaria, acceso a salud básica, acceso a educación, vivienda u otros.
- Las intervenciones tienden a estructurarse en forma de red o programas sociales integrados, de forma que las intervenciones no constituyan acciones aisladas, sino que siguen en un conjunto de políticas y estrategias de protección social.
- Los programas de protección social definen incentivos, requisitos, condiciones o contraprestaciones por parte de los beneficiarios, quienes deben asumir alguna cuota de compromiso con los mismos, lo cual corresponde al concepto de programas de transferencia condicionada.
- Las iniciativas propician la participación de los beneficiarios y muchas de ellas definen mecanismos de control social por parte de los mismos

La implementación de sistemas o redes de protección social con estas características requiere de la integración de información y servicios, frecuentemente involucrando múltiples instituciones. Algunas de estas implementaciones se han realizado en Argentina, Brasil, Chile, Colombia, México y Uruguay (Gómez, 2011; ILO, 2011).

Por otra parte, los programas de servicios sociales, generalizados en los países desarrollados, proveen diferentes tipos de beneficios al conjunto de la población, lo cual resulta en una diferencia importante con los programas de protección social mencionados antes. En este contexto, también se ha identificado la necesidad de utilizar datos integrados de los beneficiarios para mejorar la calidad de los servicios,

especialmente cuando se realizan los casos para el otorgamiento de beneficios (también conocido como *Case Management*). En particular, la situación presentada en Carter (2009) expone las consecuencias trágicas de la falta de coordinación entre programas sociales del Distrito de Columbia en los Estados Unidos. Estos problemas se han abordado, por ejemplo, en los servicios sociales de la ciudad de Nueva York y de un grupo de condados del estado de California. En ambos proyectos, se construyeron sistemas de información integrados con datos de diferentes áreas de beneficios como apoyo al reconocimiento de derecho de prestaciones. También, en Francia, con el objetivo de mejorar la calidad de los servicios y la detección de fraudes, se creó el “*Répertoire national commun de la protection sociale*” (RNCPS) (RNCPS, 2006). Este sistema registra los datos de identificación de beneficiarios, el organismo de afiliación, los riesgos cubiertos, y otros datos que forman parte de la declaración que realiza en la solicitud de los beneficios.

En síntesis, las características más sobresalientes de este tipo de escenario son la creación de sistemas de información sobre personas (potenciales beneficiarios), integrando datos de diferentes instituciones o programas sociales. Estos sistemas se utilizan tanto para identificar beneficiarios y servir como ayuda para caracterizar los casos, como para el seguimiento de los programas y el análisis de impacto.

2.3.2.1 Sistemas integrados de seguro de salud

Los sistemas de seguros de salud con proveedores de servicios públicos y privados constituyen uno de los modelos más comunes. En varios países, a los efectos de mejorar la cobertura y sustentabilidad del sistema, así como la calidad y eficiencia de los servicios al público, se han implementado sistemas integrados de seguro de salud que realizan, en diferente grado, la integración de los diferentes seguros, así como la vinculación entre los mismos y con las entidades prestadoras de servicios médicos.

Los sistemas integrados de seguro de salud requieren fundamentalmente de información unificada sobre las personas aseguradas, de forma que permitan a los prestadores de servicios de salud la comprobación de los derechos de personas solicitantes de servicios. Esto implica la integración de fuentes de datos con información que permita reconocer los derechos de las personas a los servicios de salud, así como también proveer mecanismo de acceso a tal información a las entidades prestadoras de servicios médicos. Este tipo de sistemas se ha implementado en Chile, Colombia, Tailandia, Uruguay, entre otros países.

En síntesis, si bien hay diferencias entre los sistemas implementados en cada país, algunas características comunes son las siguientes:

- Los sistemas se basan en bases de datos de referencia sobre personas cubiertas por el seguro de salud y sus derechos, los cuales pueden basarse en diferentes situaciones (trabajador, hijos, cónyuge, etc.).

- Se realiza la integración de diferentes sistemas para obtener información que permite caracterizar los derechos (por ejemplo, otros seguros de salud, sistemas de recaudación de contribuciones, etc.).
- Se trata de implementar una conectividad fluida e intercambio de datos con instituciones prestadoras de Servicios de Salud; así como entidades políticas y de control (por ejemplo, Ministerio de Salud).

2.3.2.2 Recaudación de cotizaciones de Seguridad Social y control del cumplimiento de las obligaciones

En la recaudación de cotizaciones y el control del cumplimiento de las obligaciones, una o varias instituciones administradoras de programas de Seguridad Social recaudan las contribuciones de los afiliados y relevan la información necesaria para la generación de los derechos correspondientes. A los efectos de generar economía de escala, las instituciones, que pueden ser públicas o privadas, usualmente interactúan con los empleadores, quienes retienen las contribuciones de los trabajadores y las aportan conjuntamente, junto con la información para la generación de los derechos de los dependientes.

En este contexto, la integración entre el conjunto de los actores involucrados tiene como objetivo aumentar la eficacia y eficiencia del sistema en su globalidad, así como facilitar la afiliación de nuevos cotizantes y también la realización de controles de cumplimiento de las obligaciones.

Los tipos más comunes de integración se basan en el intercambio de datos:

- Entre instituciones recaudadoras y contribuyentes, en particular con empleadores que acumulan contribuciones de los trabajadores dependientes para luego aportarlas conjuntamente.
- Entre varias instituciones recaudadoras, incluyendo oficinas responsables de impuestos generales.
- Con bancos y otras entidades financieras con el objetivo de facilitar el pago de las contribuciones.

Por otro lado, se han implementado sistemas integrados de recaudación de cotizaciones, que integran operaciones de recaudación de diferentes instituciones, tanto de Seguridad Social como también de impuestos, pero sin que esto implique la unificación de la función completa de recaudación en una única institución. Como ejemplo, se podría citar los casos de Corea, Francia y Uruguay.

2.3.2.3 Mejora de la eficiencia institucional, de la calidad de la información y los servicios

La integración de información y servicios permite realizar mejoras importantes en la gestión, así como lograr un mayor aprovechamiento de la complementariedad de funciones provistas por instituciones o departamentos diferentes. Este tipo de integración se ha puesto en práctica internamente en instituciones, a través de la implementación de sistemas de información únicos para datos centrales (o registros únicos), así como también de la implementación de sistemas informáticos siguiendo Arquitecturas Orientadas a Servicios.

Los sistemas de información unificados o de registros únicos dentro de las instituciones permiten evitar inconsistencias derivadas de multiplicaciones de los datos, así como reducir costos de su recolección. Asimismo, permite implementar comunicaciones ágiles de cambios en los datos, por ejemplo, referidas a la situación personal, familiar y laboral. Este tipo de sistemas de información ha sido implementado en numerosas instituciones.

En algunos países se han implementado *frameworks* y plataformas de interoperabilidad especializadas en Seguridad Social, que permiten la interconexión de diferentes instituciones. Es el caso del *Social Security Crossroads Bank* de Bélgica, que cumple un rol de “integrador” en la Seguridad Social de Bélgica (CBSS, 2018).

En el mismo sentido hay que mencionar el estándar de interoperabilidad de la Seguridad Social en Francia (INTEROPS) (AISS, 2013; INTEROPS, 2011). Se trata de una norma utilizada por los organismos de protección social para proyectos de intercambio de información. Respeto las normas técnicas internacionales (SOAP, SAML, XML, SSLV3, etc.) y sus especificaciones son públicas. Son claros ejemplos de interoperabilidad técnica.

Paralelamente a la elaboración de especificaciones, se llevó a cabo un proyecto de herramientas que facilitó la adopción de la norma y aceleró su aplicación, fundamentalmente con objeto de abordar los aspectos técnicos más complejos. Esto condujo al suministro de un programa informático competitivo e industrial para aplicar la norma.

2.3.2.4 Acuerdos internacionales de Seguridad Social

En el capítulo 1, al describir el contexto de esta tesis, se ha hablado ampliamente de los acuerdos internacionales en materia de Seguridad Social. Tal como se ha mencionado sus objetivos principales son:

- Permitir la acumulación de períodos de actividad generados en diferentes países para cumplir con los requisitos de obtención de una pensión.

- Establecer condiciones para que trabajadores en “desplazamiento temporal” estén cubiertos por el sistema de Seguridad Social habitual y no deban contribuir al régimen del país destino.
- Facilitar las gestiones y pago de beneficios independientemente del país de residencia del beneficiario.

La implementación de estos acuerdos se basa en el intercambio de datos entre las instituciones firmantes, por ejemplo, información de periodos trabajados en un país, datos sobre un traslado temporal, etc. Corresponden, en cierta medida, a operaciones en un sistema de Seguridad Social internacional regulado por el acuerdo firmado.

Asimismo, se requiere la gestión de un registro de operaciones que aporte transparencia y posibilidad de seguimiento a las transacciones realizadas. Las tecnologías de interoperabilidad a aplicar serán principalmente los estándares para formateo de datos (XML) e interconexión de sistemas (servicios web) en sus versiones que permiten incluir firmas electrónicas y encriptar los datos (WS-Security, 2004).

2.3.3 Dimensiones de interoperabilidad en la Seguridad Social: situación actual

Si bien el objeto principal de esta tesis es la interoperabilidad semántica, las interdependencias entre los distintos niveles o dimensiones hacen necesario un estudio de situación de todos ellos.

2.3.3.1 Interoperabilidad política y legal

La aplicación de la interoperabilidad requiere de un apoyo político claro, que promueva los objetivos de cooperación, y que genere visiones institucionales compatibles y centradas en los objetivos similares de implementación de programas de Seguridad Social. En términos prácticos, esto implica alinear visiones políticas y prioridades, apoyar al más alto nivel la aplicación de estándares de información y tecnologías, y transmitir directivas a las Gerencias para que coordinen los planes y proyectos de forma que hagan posible las implementaciones basadas en interoperabilidad, tanto dentro de la institución como con otras instituciones.

Por su parte, la interoperabilidad inter-institucional, requiere tanto de acuerdos entre las instituciones involucradas, como de alineamientos políticos en el ámbito del gobierno, que promuevan la implementación conjunta y coordinada de programas sociales, teniendo en cuenta las competencias de cada institución y aprovechando sus capacidades operativas.

En el caso de la interoperabilidad internacional, asociada a los intercambios de datos para la implementación de acuerdos internacionales de Seguridad Social, se requiere

además la aprobación por parte de los parlamentos de los países de los acuerdos y sus procedimientos administrativos correspondientes. Idealmente, dado que la interoperabilidad legal es impulsada por los objetivos decididos al más alto nivel político, a partir de los acuerdos políticos se debería llegar a un marco legislativo o normativo común aplicable a determinados dominios de interoperabilidad entre las entidades participantes.

Tanto en el capítulo 1, sección 1.2, como en los apartados anteriores de esta sección, se ha visto la situación de los acuerdos internacionales en materia de Seguridad Social e intercambio de datos en este ámbito.

No hay que olvidar tampoco otro de los aspectos legales fundamentales, como es el cumplimiento de las normativas de protección de datos en particular en el tratamiento de datos administrados por otras instituciones. La implementación suele conllevar la asignación de una o más autoridades responsables de dicha protección, así como del establecimiento de las correspondientes estructuras colaborativas y los mecanismos asociados. Sin embargo, no es un aspecto especialmente característico de los intercambios de Seguridad Social, por lo que no se va a profundizar en él.

Un aspecto que puede ayudar a la interoperabilidad legal es la utilización de especificaciones y estándares abiertos en la definición de las directivas de colaboración.

2.3.3.2 Dimensión organizativa

Situación general

La interoperabilidad organizativa está relacionada con la definición modelos administrativos y procesos compatibles, con el fin de lograr la colaboración de las partes que se proponen intercambiar o compartir información y servicios, independientemente de su organización y estructura interna. Se entiende por “aspectos de interoperabilidad organizacional” aquellos que posibilitan conocer y entender las políticas de acceso y uso de datos o servicios, las responsabilidades personales o institucionales, los objetivos y los fines perseguidos por la organización al crear un dato o proporcionar un servicio. Además, la interoperabilidad organizativa tiene por objeto atender las necesidades de la comunidad de usuarios proporcionando servicios disponibles, localizables, accesibles y orientados al usuario.

En el contexto de los sistemas de Seguridad Social, como procesos potencialmente involucrados se tienen los siguientes:

- Gestión de datos comunes, por ejemplo, sobre personas y empresas.
- Interacción con beneficiarios o contribuyentes, empresas y trabajadores.
- Realización de cálculos de prestaciones y cotizaciones.

- Validación de condiciones relativas a la elegibilidad para el otorgamiento de beneficios o el cálculo de cotizaciones.
- Emisión de constancias variadas, registro de empresas y sus trabajadores, etc.

La notificación de cambios de estado y de eventos en general sobre personas u entidades que interactúan con los sistemas de Seguridad Social también debe considerarse como procesos que requieran de interoperabilidad, ya que tales cambios de estado pueden repercutir en diferentes servicios. Son un ejemplo los matrimonios, fallecimientos, enfermedad, situación de desempleo, etc.

Según EIF v2.0 (EIF, 2010) la interoperabilidad organizativa se refiere a un amplio conjunto de elementos de interacción, incluyendo los procesos de negocio, interfaces de negocio tales como correo electrónico, portales web, etc. y eventos de negocios dentro y entre las administraciones con la participación de las partes externas (empresas y ciudadanos). Como enfoque metodológico, suelen utilizarse las técnicas de *"life events"* (Sanati, 2010) para los ciudadanos (todo lo que le acontece a un ciudadano desde su nacimiento hasta su fallecimiento), y los *"business episodes"* (desde la creación de una empresa hasta su desaparición) para empresas. De esta forma, ciudadanos y empresas pueden centrarse en sus necesidades en lugar de tener que conocer las competencias concretas de las organizaciones del sector público. El servicio prestado está orientado al cliente, es transparente y sigue el enfoque *"ventanilla única"*. Este aspecto de la interoperabilidad se refiere a cómo las diferentes organizaciones, como por ejemplo las administraciones de los Estados miembros de la Unión Europea, colaboran para lograr un beneficio mutuo y alcanzan acuerdos en servicios relacionados con los objetivos de *e-Government*.

De acuerdo con las experiencias estudiadas, la interoperabilidad organizativa en la práctica implica la coordinación de procesos de negocio y el intercambio de información entre las instituciones o departamentos de las mismas. Para que tenga éxito, los participantes deben estar de acuerdo en el qué y el cuándo en el acceso a servicios compartidos y en el intercambio de información, sobre las normas comunes para que la interoperación se produzca de forma segura, con una sobrecarga mínima y de manera continuada. Asimismo, resulta necesario utilizar los mismos estándares para describir los procesos de negocio, que deben estar adecuadamente documentados. Según las directrices de la AISS sobre Interoperabilidad (AISS, 2013), resulta conveniente establecer un repositorio de procesos de negocio y mejores prácticas entre las entidades participantes, facilitando la reutilización de las mismas. Asimismo, resulta importante el uso de instrumentos como los acuerdos de nivel de servicio para la formalización de aspectos específicos a los servicios provistos a través de la interoperación, o también de *"Memoranda of Understanding"* que detallan los acuerdos entre instituciones o países sobre las acciones conjuntas y la cooperación.

Intercambio internacional de datos: directorio de instituciones

En el ámbito internacional, según la forma en que cada administración de Seguridad Social esté organizada, los sistemas de información adoptan las arquitecturas necesarias para darles servicio. En consecuencia, los intercambios no se realizan entre sistemas análogos, sino requieren ciertas “conversiones” de modelos organizativos. Como ejemplo, podría citarse el caso del desempleo en España: mientras que en la mayoría de los países se incluye como un sector más dentro de la Seguridad Social, en España tiene una consideración diferente, encuadrándose en una organización totalmente independiente tanto desde el punto de vista lógico como físico (redes de comunicación, sistemas de información, etc.). Otro caso similar es la organización de la asistencia sanitaria. Podemos encontrar estructuras totalmente distintas según el sistema nacional responda a un modelo nacional contributivo o un modelo residencial o el de múltiples cajas. En este último caso, estaríamos en una situación que fuerza la aparición de los organismos de enlace (desde el punto de vista organizacional, no técnico). Los organismos de enlace tienen que reunir información heterogénea en formato, incontrolada y sobre la que no tienen capacidad de decisión.

La solución más avanzada para resolver estos problemas consiste en utilizar un directorio de instituciones encargado de centralizar tanto la información básica sobre todas las instituciones que van a intervenir en los intercambios, como la información sobre los procesos de negocio en los que puede y debe intervenir cada una de las instituciones, y los flujos de información asociados a cada documento electrónico para prestación sobre la que tienen competencia (Delgado et al., 2012).

En el marco de la Unión Europea, el problema ha estado presente mucho antes de la aparición de los intercambios telemáticos, hasta el punto de que se plasmó en los anexos de dos reglamentos comunitarios de los años 1971 y 1972 (DOUE, 1971; 1972). La solución práctica propuesta pasaba, en primer lugar, por la figura de los organismos de enlace, algo parecido a un centro de intercambio de ámbito nacional responsable de hacer llegar cada documento a la institución competente correspondiente. Por otra parte, la labor del organismo de enlace se debía basar en las relaciones de instituciones competentes incluidas como anexo en los citados reglamentos.

La llegada del tratamiento electrónico de la información hizo conveniente buscar un mecanismo que facilitara la búsqueda de forma sistemática. La primera aproximación rigurosa coincide con el arranque del proyecto de la Tarjeta Sanitaria Europea. Se trata de una iniciativa del Consejo Europeo de Barcelona de marzo de 2002, desarrollada a partir de 2003 (DOUE, 2003). La Tarjeta empezó a emitirse el primero de junio de 2004 en algunos Estados miembros de la UE, entre ellos España, y terminó su implantación en diciembre de 2005.

El hecho de que dicha tarjeta generase unos derechos que tenían que ser reconocidos por otras instituciones, hizo imprescindible el control de las instituciones con capacidad de emitir este documento. Con ese fin principal se creó la base de datos de instituciones (*Code List Database*, CLD), en principio restringida al ámbito de la asistencia sanitaria internacional (figura 12).

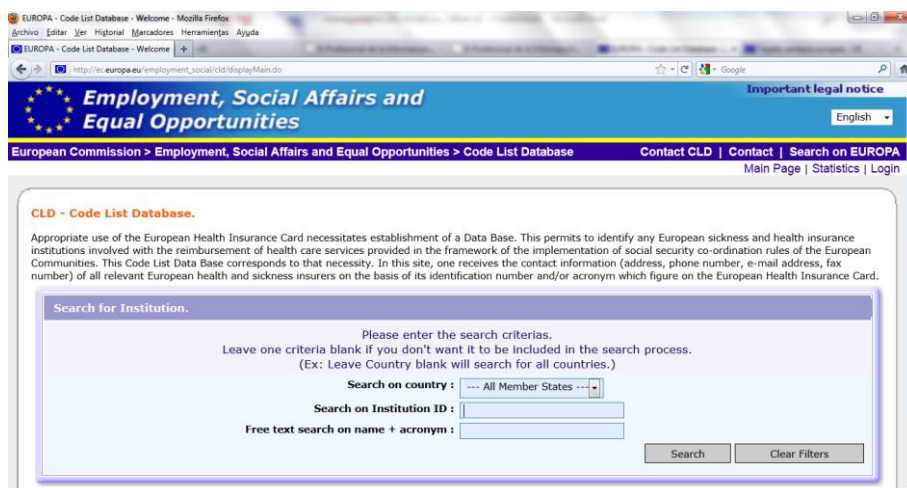


Figura 12. *Code List Database* (PAI, 2015)

Su objetivo era proporcionar a los Estados miembros de la UE un punto centralizado donde poder encontrar la información sobre las instituciones de los otros Estados. Por “información de las instituciones” se entendía información de contacto, así como códigos y calificadores de las instituciones competentes. Los datos disponibles podían dividirse en cinco categorías:

- Identificación de la institución
- Dirección de la institución
- Información de contacto
- Papel de la institución (*rol*)
- Información para archivo

Se accedía a la CLD a través de una dirección de internet sin ningún tipo de seguridad, pero la dirección no era de conocimiento del público en general. En 2005 se decidió incluir en la CLD todos los sectores, creándose los códigos correspondientes para representarlos. En la figura 13 se muestra el listado obtenido en una consulta al CLD.

CLD - Code List Database.

Search for Institution.

Search on country: Spain

Search on Institution ID:

Free text search on name + acronym:

Search Clear Filters

79 Institutions found, displaying 1 to 20.
[First/Prev] 1, 2, 3, 4 [Next/Last]

ID Number	Acronym	Name	Country	Valid From	Valid To	Reason for change	New ID Number
0000	INSS ORG. ENLACE	DIRECCION GENERAL DEL INSS	ES	01/01/1986			
0100	INSS ALAVIA	DIRECCION PROVINCIAL DEL INSS DE ALAVIA	ES	01/01/1986			
0200	INSS ALBACETE	DIRECCION PROVINCIAL DEL INSS DE ALBACETE	ES	01/01/1986			
0300	INSS ALICANTE	DIRECCION PROVINCIAL DEL INSS DE ALICANTE	ES	01/01/1986			
0400	INSS ALMERIA	DIRECCION PROVINCIAL DEL INSS DE ALMERIA	ES	01/01/1986			
0500	INSS AVILA	DIRECCION PROVINCIAL DEL INSS DE AVILA	ES	01/01/1986			
0600	INSS BADAJOZ	DIRECCION PROVINCIAL DEL INSS DE BADAJOZ	ES	01/01/1986			
0700	INSS ILLES BALEARS	DIRECCION PROVINCIAL DEL INSS DE ILLES BALEARS	ES	01/01/1986			
0800	INSS BARCELONA	DIRECCION PROVINCIAL DEL INSS DE BARCELONA	ES	01/01/1986			
0900	INSS BURGOS	DIRECCION PROVINCIAL DEL INSS DE BURGOS	ES	01/01/1986			
1000	INSS CACERES	DIRECCION PROVINCIAL DEL INSS DE CACERES	ES	01/01/1986			
1100	INSS CAZIS	DIRECCION PROVINCIAL DEL INSS DE CAZIS	ES	01/01/1986			
1200	INSS CASTELLON	DIRECCION PROVINCIAL DEL INSS DE CASTELLON	ES	01/01/1986			
1300	INSS CIUDAD REAL	DIRECCION PROVINCIAL DEL INSS DE CIUDAD REAL	ES	01/01/1986			
1400	INSS CORDOBA	DIRECCION PROVINCIAL DEL INSS DE CORDOBA	ES	01/01/1986			
1500	INSS A CORUÑA	DIRECCION PROVINCIAL DEL INSS DE A CORUÑA	ES	01/01/1986			
1600	INSS CUBENCA	DIRECCION PROVINCIAL DEL INSS DE CUBENCA	ES	01/01/1986			
1700	INSS GIRONA	DIRECCION PROVINCIAL DEL INSS DE GIRONA	ES	01/01/1986			
1800	INSS GRANADA	DIRECCION PROVINCIAL DEL INSS DE GRANADA	ES	01/01/1986			
1900	INSS GUADALAJARA	DIRECCION PROVINCIAL DEL INSS DE GUADALAJARA	ES	01/01/1986			

Click to export: [CSV \(ISO 8859-1\)](#) | [CSV \(UTF-8\)](#) | [Excel \(UTF-8\)](#)

Figura 13. Ejemplo de consulta a la *Code List Database*

Con la aprobación de los Reglamentos comunitarios 883/2004 (OJEU, 2004) y 987/2009 (OJEU, 2009), el panorama de los intercambios electrónicos en el ámbito de la Unión Europea cambió radicalmente. Según dichos reglamentos, tal como se ha dicho en la sección 1.1.6, desaparecían los documentos en papel reconocidos hasta el momento y aparecían los Documentos Electrónicos Estructurados (*Structured Electronic Documents, SED*), únicos legalmente reconocidos a partir de mayo de 2012. En consecuencia, la única vía de comunicación para las instituciones debía ser el intercambio electrónico. El apoyo técnico a este nuevo panorama lo aportaba el proyecto EESSI, mencionado en el capítulo 1, sección 1.1.6.

El primer componente de EESSI en entrar en producción fue el entonces denominado "Directorio Maestro" (*EESSI Master Directory*); en funcionamiento desde primero de mayo de 2010 (EESSI, 2010). Se trataba de una evolución natural de la CLD, de forma que se pudiera tener conocimiento de todas las instituciones que iban a intervenir en los intercambios, así como de sus detalles asociados. Pero este directorio no tenía un carácter meramente informativo, sino que iba a jugar un papel clave en el funcionamiento del intercambio electrónico, tanto desde el punto de vista técnico como desde el punto de vista de negocio.

En lo que a funcionalidades de negocio se refiere, el directorio de EESSI debía proporcionar, tanto a los ciudadanos como a las instituciones, un punto de referencia electrónico donde buscar información referida a las instituciones de Seguridad Social (servicios proporcionados, información de contacto, etc.). El fin es ayudar a los residentes en la UE a localizar y acceder a los servicios de Seguridad Social.

En cuanto a las funcionalidades técnicas, el directorio de EESSI debía mantener información que sirviera de base a los intercambios electrónicos entre instituciones. Por ejemplo, era necesario mantener una lista de certificados de seguridad para facilitar la confidencialidad, integridad y no repudio de los mensajes intercambiados. Es más, era necesario almacenar la información de encaminamiento que dirigiera la transferencia de mensajes al destino apropiado.

Físicamente, el directorio estaba alojado en el nodo central (Comisión Europea) como un componente más de EESSI, siendo ésta la versión de referencia (figura 14). Había una réplica en cada punto de acceso y había asimismo posibilidad de realizar más réplicas o extracción de datos para su utilización por parte de las aplicaciones nacionales. Se dio acceso tanto a los funcionarios como al público en general.

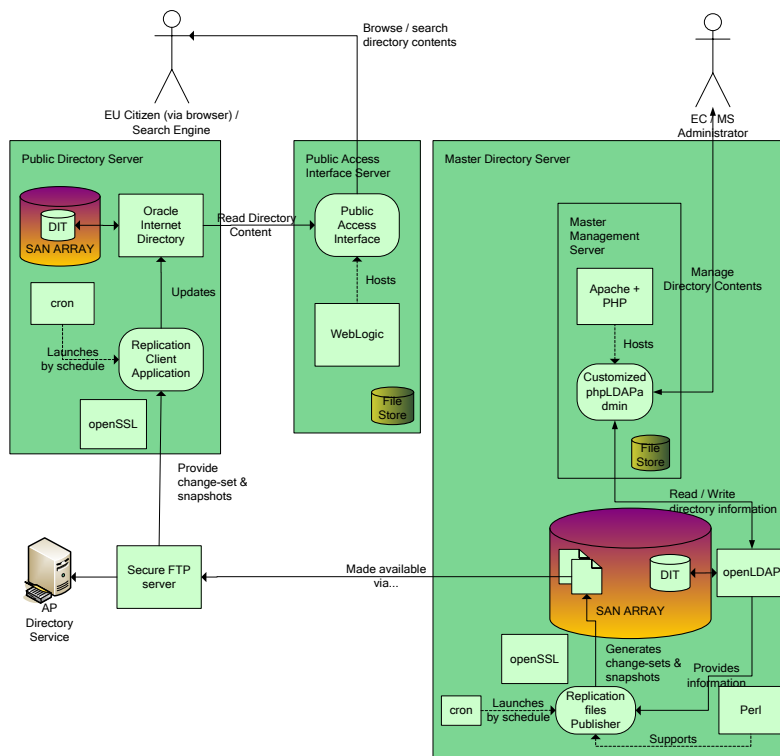


Figura 14. Arquitectura del directorio maestro EESSI (Delgado et al., 2012)

Técnicamente, el sistema de almacenamiento del directorio estaba ubicado en una plataforma basada en una versión de código abierto (*OpenLdap*) de el “Protocolo de Acceso Ligero a Directorio” (LDAP) basado en una simplificación del estándar X.500. Para el acceso externo se utiliza el producto phLDAPAdmin, escrito en lenguaje PHP.

1	Identification Number of the Institution
2	Acronym of the Institution
3	Full name of the Institution
4	Full name of the institution (lang 2)
5	Full name of the institution (lang 3)
6	Full name of the institution (ENGLISH)
7	Full name of the institution in Latin (if applicable)
8	street and number
9	street and number (Language 2)
10	street and number (Language 3)
11	street and number (ENGLISH)
12	street and number (in Latin if applicable)
13	Visitors address
14	Town
15	Town (Lang 2)
16	Town (Lang 3)
17	Town (ENGLISH)
18	Town (in Latin if applicable)
19	Zip code
20	Country code
21	Phone
22	Fax
23	Electronic Address (address where SED will be sent)
24	Public Contact Internet address
25	Public Key
26	URL
27	Function(s)
28	Coverage of social security matters
29	Types of benefits & obligations
30	Personal coverage
31	Geographical coverage
32	Territorial coverage
33	EC Regulations
34	Elements of EC Regulations
35	Grouping element of function
36	Relationship with other institution
37	Validity period: from
38	Validity period: to
39	Reason for change
40	Comments
41	Issues EHIC

Tabla 4. Campos (atributos) del directorio maestro en su versión inicial

Los Estados miembros eran responsables de recoger y depurar la información necesaria de las instituciones y hacerla llegar a la Comisión Europea, tanto para la carga inicial como las futuras modificaciones. Para recopilar dicha información, fue necesario empezar por solicitar una relación de instituciones por cada Estado miembro y los puntos de acceso (mínimo uno y máximo cinco) designados al efecto. Posteriormente, a cada institución de las que aparecen en el directorio se le solicitó que determinase qué tipo de prestaciones cubría. Se definió un proceso para la actualización de los datos contenidos en el directorio (DOUE, 2010).

Los campos (atributos) de esta primera versión del directorio maestro son los derivados del artículo 83.4 del reglamento 883, que establecía que el directorio debe registrar (ver campos en tabla 4).

Para la versión definitiva de EESSI, oficialmente en producción desde julio de 2017, la primera versión del Directorio Maestro se consideró insuficiente. En noviembre de 2018 se está trabajando en una nueva versión que se denominará *Institution Repository*, técnicamente en la línea con la arquitectura general de EESSI pero conceptualmente siguiendo los mismos principios que la versión anterior.

2.3.3.3 Dimensión semántica

Consideraciones generales

El objetivo de la interoperabilidad semántica es permitir que los recursos de información estén conectados y sean automáticamente entendibles y, en consecuencia, reusables por las aplicaciones de ordenador que no intervinieron en su creación. El hecho de que la información generada por un ordenador tenga que ser procesada por otro sistema interpretando su significado de forma correcta, hace que surjan una serie de complicaciones adicionales que afectan tanto a la fuente de la información como a su receptor. Para conseguirlo se necesitan acuerdos en la forma de representar y dar contexto a la información. De esta forma, las herramientas automáticas podrán compartir y procesar la información, incluso cuando se haya diseñado de forma independiente.

Como ocurre con la interoperabilidad organizacional, el enfoque sistemático para afrontar las dificultades producidas por la interoperabilidad semántica es una actividad que debe ser realizada a nivel sectorial, teniendo en cuenta los *life events* o *business episodes* afectados en cada caso. Un conjunto de datos, como identificadores nacionales de empresas, ciudadanos y administraciones, necesitan ser identificados a nivel nacional, internacional o el que corresponda al tipo de interoperabilidad. Eso exigirá la preparación de infraestructuras organizacionales y técnicas para dar solución al nivel semántico.

Desde un punto de vista organizativo, para cada servicio que requiera interoperabilidad es necesario:

- Que la administración responsable publique la información sobre los correspondientes elementos de datos involucrados.
- Que la administración responsable proponga un acuerdo sobre los datos y los diccionarios de datos requeridos para el nivel de los intercambios.
- Que la administración responsable presente una propuesta para un acuerdo multilateral, asociando tablas entre los distintos niveles.

Desde un punto de vista técnico, un requerimiento esencial para el intercambio de información es contar con lenguaje que permita la descripción del significado y estructura de los datos subyacentes. Es el caso, por ejemplo, de los lenguajes de marcas (*mark-up language*). En el contexto actual tecnológico y de mercado, el más ampliamente difundido es XML. Sin embargo, XML por sí sólo no garantiza ni proporciona la interoperabilidad semántica. Eso se consigue mediante iniciativas de desarrollos semánticos comunes basados en XML. La subsecuente introducción de esquemas XML y los artefactos relacionados (metadatos, ontologías, etc.) hacen posible la integración de servicios desarrollados con diferentes vocabularios y con diferentes perspectivas de datos.

Conflictos semánticos

Hay diferentes clasificaciones de los posibles problemas semánticos que pueden aparecer en los intercambios de información. La literatura especializada distingue entre los que aparecen en el ámbito de los datos (*data-level conflicts*) y en los esquemas o estructuras (*schema-level conflicts*) (Fraunhofer IST, 2009). Estos conflictos están relacionados con las metodologías y las técnicas para evitarlos, resolverlos y detectarlos.

Los conflictos en el ámbito de los datos son diferencias producidas por múltiples representaciones e interpretaciones de datos similares (Park, 2004).

Dentro de su ámbito, a su vez puede distinguirse entre:

- **Conflictos de valores de datos.**

Tienen lugar cuando se dan diferentes significados para el mismo nombre de dato. Normalmente se los conoce como homónimos (mismo nombre, distintos significados). A menudo el origen de estos problemas está en otros conflictos, como en la convención de nombres que se haya adoptado.

También entran en esta categoría los sinónimos (distinto nombre, mismo significado). Sería el caso de Holanda-Países Bajos.

Estos conflictos se pueden evitar mediante vocabularios controlados, con significado definido con precisión.

- **Conflictos de representación de datos.**

Una representación diferente impide la correcta interpretación. El caso más habitual son las fechas (Tambouris et al., 2007). Distintas representaciones sintácticas, como separadores en listas o diferentes conjuntos de caracteres, puede también dar lugar a estos conflictos.

La forma de detectarlos puede ser la utilización de “*schema matching*”, término que podríamos traducir como coincidencia de esquemas. La resolución está generalmente ligada a técnicas de conversión.

- **Conflictos de unidades.**

Están relacionados con distintos sistemas de unidades. Ejemplo, kilómetros y millas. Se resuelven mediante conversiones con procesos de coincidencia de esquemas.

- **Precisión de datos.**

Se deben a la utilización de diferentes escalas en los intercambios numéricos, con diferente número de decimales.

Se resuelven con una comparación entre lo remitido y lo recibido, con una eventual conversión.

- **Conflictos de lenguaje de datos.**

Se refieren a los casos de multilingüismo. En muchas ocasiones es difícil la traducción de la denominación de un concepto determinado. La solución más apropiada es la utilización de vocabularios controlados que definan con precisión los términos, como es el ejemplo del glosario de términos de Build 6 (proyecto TESS).

- **Conflictos de esquema o estructura.**

Este tipo de conflictos aparecen por la utilización de diferentes nombres para información equivalente, diferentes conceptos para un dominio, así como diferentes formas de estructurar un esquema (Fraunhofer, 2008). Similar a los homónimos, pero a nivel de estructura.

- **Conflictos de nombre y concepto.**

Aparecen asociados a la relación entre símbolos o nombre y conceptos:

- Sinónimos: dos símbolos o nombres son sinónimos si dan denominación al mismo concepto.
- Homónimos: son los símbolos o nombres que pueden activar diferentes conceptos, normalmente dependiendo del contexto en el que son usados.
- Hiperónimos: un símbolo es un hiperónimo de otro si es un término más amplio que lo comprende. Ejemplo: persona es un hiperónimo de hombre y mujer.

- Solapamiento: si dos símbolos o nombres diferentes se refieren a diferentes conjuntos de cosas, pero la intersección de estos conjuntos no es vacía.

Este tipo de conflictos se detectan con la técnica de *schema matching*. Su resolución se consigue aplicando coincidencia de esquemas, pero su resolución no es siempre posible. En particular la hiperonimia y el solapamiento.

- **Isomorfismos de esquema.**

Son conflictos que ocurren cuando el remitente describe el mismo concepto con diferentes conjuntos de atributos. En términos de conjuntos de operaciones, estos conjuntos de entidades no son compatibles (Sheth et al., 1993).

La resolución de este tipo de conflictos se consigue aplicando técnicas de coincidencia de esquemas; por ejemplo, con las transformaciones apropiadas, aunque no siempre es posible.

- **Conflictos de pérdida de entidades.**

Este tipo de conflictos ocurren cuando las entidades aparecen bien sea en la fuente o bien en el destino, sin entidades equivalentes en la contraparte.

Detectar estos conflictos necesita de una coincidencia de esquemas. Su resolución puede conseguirse aplicando también estas técnicas, pero depende de la posibilidad de aplicar las funciones de conversión necesarias. Por ejemplo, un número entero puede transformarse en un racional pero no a la inversa.

- **Generalización.**

Tienen lugar cuando se utilizan diferentes elecciones de generalización de conceptos son utilizadas. Por ejemplo, cuando un esquema contiene el concepto de una persona, mientras otro distingue entre hombre y mujer.

También estos conflictos se detectan con coincidencia de esquemas utilizando vocabularios controlados. Su resolución no siempre es posible.

- **Discrepancias en los esquemas.**

Tienen lugar cuando los dos esquemas tienen distintas estructuras lógicas y un término está modelado como un elemento en un esquema, mientras en otro está modelado como un atributo.

Se resuelven aplicando las técnicas de coincidencia de esquemas. Consisten en comparar unos esquemas con otros y establecer reglas de transformación entre ambos.

- **Conflictos jerárquicos.**

Se deben a que dos esquemas tienen diferentes niveles de jerarquía. Por ejemplo, un esquema origen es plano y el destino está estructurado o viceversa.

Se resuelven aplicando las técnicas de coincidencia de esquemas.

- **Claves externas.**

Estos conflictos ocurren cuando un esquema en la fuente utiliza claves externas mientras que no se utilizan en destino.

Se resuelven aplicando las técnicas de coincidencia de esquemas.

- **Conflictos de restricciones.**

Tienen lugar cuando un esquema en la fuente y en el destino difieren en sus limitaciones. Por ejemplo, en la cardinalidad impuesta a un atributo.

Se detectan mediante coincidencia de esquemas. Su resolución puede ser imposible si la cardinalidad de la fuente es mayor que la del destino.

Avances semánticos en los intercambios de Seguridad Social.

El mayor avance en la creación de un modelo de datos para los intercambios de Seguridad Social tuvo lugar con motivo del proyecto TESS, antecedente de EESSI, concretamente en el subproyecto conocido como Build 6, para lo que hay que remontarse a los años 90.

Tal como se decía en el capítulo 1, Build 6 era un proyecto piloto en los sectores de Pensiones y Enfermedad. Su objetivo era el desarrollo de los mensajes electrónicos que reemplazarán a los entonces existentes formularios denominados "E", utilizados para el intercambio de datos entre los Estados miembros. Se adoptaba un enfoque intersectorial para identificar las partes comunes entre los diferentes formularios en papel, identificar los conceptos básicos que aparecen en distintos formularios, racionalizando de esta manera las estructuras de datos. El resultado debían ser unos mensajes electrónicos en formato EDIFACT que tendrían que ser homologados por el correspondiente organismo de estandarización a nivel mundial. Todo esto debe encuadrarse dentro de la aplicación del Reglamento 1408/71 (DOUE, 1971), entonces en vigor, y el más amplio programa de intercambio de datos entonces existente en Europa, como era el caso del EDI (*Electronic Data Interchange*) (Hípola, 1992).

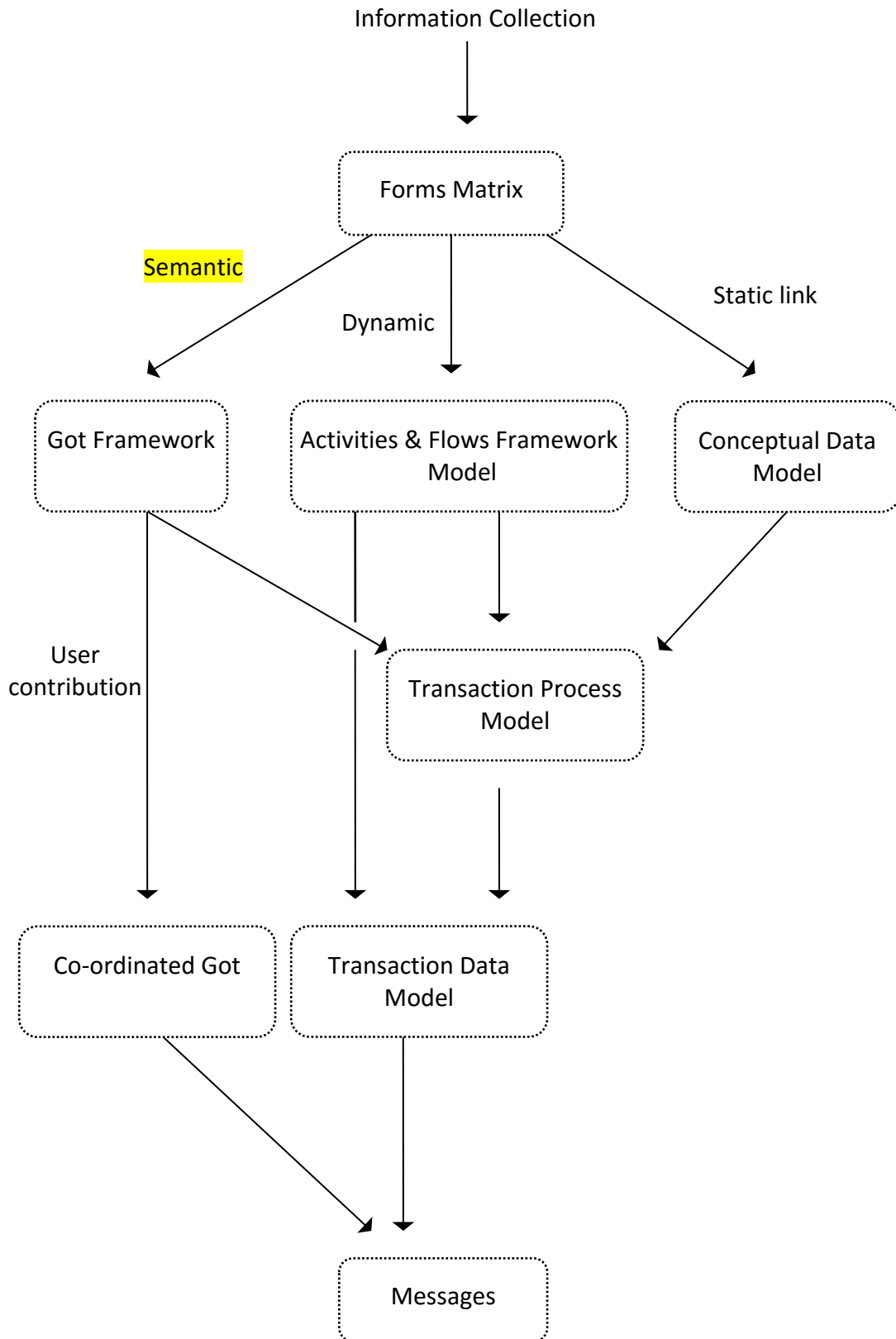


Figura 15. Diagrama de flujo de documentos del proyecto (TF, 2007)

En el diagrama de la figura 15 (TF, 2007), puede verse la relación entre las distintas partes del proyecto Build 6 y su incidencia en los temas semánticos.

Como paso previo se debía construir un Modelo Conceptual de Datos (CDM) y un Glosario de Términos (GOT) que describiera todos los términos utilizados en los procesos de intercambio.

En primer lugar, se identificaron los datos como parte de los campos de los formularios, deduciendo de estos el concepto al que se refieren. En lo posible, se trataba de describir el dato en términos de su tipo, posibles valores y utilización. Se trataba de establecer después una relación entre los datos y los formularios en los que aparecía.

El paso siguiente fue modelar los datos (conceptos de información, información conceptual) de manera que se definieran entidades, atributos y relaciones manejadas en el dominio de la Seguridad Social. Cada una de las “informaciones conceptuales” era definida para obtener todos los posibles diferentes significados de acuerdo con los diferentes actores, países, reglamentos, etc.

Se trataba de definir las partes comunes de todos los formularios, estructurar la información y conseguir una reusabilidad en el momento de transformarlos en mensajes electrónicos.

Posteriormente se estudiaban los aspectos dinámicos, identificando las actividades ejecutadas por las instituciones de Seguridad Social en relación con los formularios.

Todos los intercambios de datos eran comparados para identificar un número limitado de funciones de negocio que los abarcaran. El objetivo era una posterior racionalización de los flujos de información.

Partiendo de los formularios existentes en 1995, Build 6 identificó todos los conceptos utilizados, creando un repositorio de información. A continuación, se reproduce un extracto de dicho repositorio en el que puede verse un formulario en papel, el E-106 del sector enfermedad (figura 16), así como el inventario de campos de este formulario (tabla 5) y los conceptos de información a los que corresponde (tabla 6).

ADMINISTRATIVE COMMISSION
ON SOCIAL SECURITY
FOR MIGRANT WORKERS

E 106 (1)

CERTIFICATE OF ENTITLEMENT TO SICKNESS AND MATERNITY INSURANCE BENEFITS IN KIND FOR PERSONS RESIDING IN A COUNTRY OTHER THAN THE COMPETENT COUNTRY

Employed and self-employed persons and members of their families residing with them; members of the family of unemployed persons

*Regulation (EEC) No 1408/71: Article 19.1.a.; Article 19.2.; Article 25.3.i
Regulation (EEC) No 574/72: Article 17.1 and 4, Article 27 (first sentence)*

The competent institution should complete Part A of the form and send two copies to the insured person, or send them - where necessary through the liaison body - to the institution in the place of residence if the form is drawn up at that institution's request. As soon as it has received the two copies, the latter institution should complete Part B and return one copy to the competent institution.

Please complete this form in block letters, writing on the dotted lines only. It consists of 4 pages, none of which may be left out.

A. Notification of entitlement

1.	Institution of the place of residence ⁽²⁾	
1.1	Name:	
1.2	Identification number of the institution:	
1.3	Address:	
1.4	Reference: your E 107 form of	
2.	the insured person	
2.1	Surname(s) ⁽³⁾ :	
2.2	Forename(s) ⁽⁴⁾ :	Date of birth:
2.3	Previous name(s):	
2.4	Address in the country of residence:	
2.5	Personal identification number:	
2.6	The person <input type="checkbox"/> is an employed person	

Full Screen
Close Full Screen

Figura 16. Formulario E-106 en su versión de 2007

IDCode	Description	FFGroup	Values	Data Element	DataElement 2	DataElement 3
E106-1.1-1	Name	E106-1		IC031		
E106-1.1-2	Code number	E106-1		IC030		
E106-1.2	Address (street, number, post code, town, country)	E106-1		IC032		
E106-1.3	Reference: your E107 form of	E106-1		IC083		
E106-10.1-1	Surname	E106-10		IC001		
E106-10.1-2	Forenames	E106-10		IC003		
E106-10.1-3	Sex	E106-10	F – M	IC005		
E106-10.1-4	Previous names	E106-10		IC004		
E106-10.1-5	Date of birth	E106-10		IC006		
E106-11.1	Name	E106-11		IC031		
E106-11.2	Address (street, number, post code, town, country)	E106-11		IC032		
E106-11.3	Stamp	E106-11		IC089		
E106-11.4	Date	E106-11		IC088		
E106-11.5	Signature	E106-11		IC089		

Tabla 5. Campos del formulario E-106

ID Code	Name	Description	Type	Values	Usage
IC001	Surname	official name of a person	Text	anything, E: both names at birth	M/I
IC002	Surname at birth	official name of a person at birth, maiden name	Text	Anything	O/M/I
IC003	Forenames	first name(s) associated with surname to identify a person	Text	Anything	M/I
IC004	Previous names	different names a person had during his life	Text	Anything	O/M/I
IC005	Sex		Single selection	male - female	O/M/I
IC006	Date of birth		Date	-	M/I
IC007	Place of birth	locality where a person is born	Structure	place, province, department or county, country	O/M
IC008	Address of person		Structure	street, number, post code, town, country. D: no tel#	M/I
IC009	Nationality	nationality and date of naturalisation	Text	a valid nationality	M
IC010	Father	father's surname and forenames	Text	Anything	
IC011	Mother	mother's surname and forenames	Text	Anything	
IC012	Registration/Insurance number	unique id of insured person within institution in country	Text	A/D: Vers# B: No.Nat. DK: CPR F: pop.reg. S/N: pers# L: AHVins#	O/M/I
IC013	Pension scheme, dead was receiving	survivor's pension claim	Single selection	none - employed - self-employed - all residents	O/M/I
IC014	Relationship	to the members of the family	Text	Anything	
IC030	Code number	institution number	unknown	Unknown	O/M
IC031	Name of institution		Text	Anything	M
IC032	Address of institution		Structure	street, number, post code, town, country, tel	O/M
IC033	Type of insurance cover (Pension)	risk, reason justifying the right to social security benefit	Single selection	death/survivor - invalidity - old-age	
IC034	Type of benefit (Health)	for benefits in kind	Single selection	medical treatment - dental treatment - medicaments - hospitalisation - examination - other	

ID Code	Name	Description	Type	Values	Usage
IC035	Address of employer		Structure	street, number, post code, town, country	N/M
IC036	Description of benefit		Text	Anything	
IC037	Type of employment		Single selection	employed - self-employed	O/M

Tabla 6. Conceptos de información

Los valores de las columnas de la tabla 6 se explican a continuación:

- Identification code (ID): código numérico único que identifica el término.
- Name: palabra que represente lo más posible el significado del término; se obtiene del Modelo Conceptual de datos y del Repositorio de Información.
- Description: explicación del significado del término.
- Type: calificador del término (texto, dato, elección simple, etc.).
- Values: valores concretos que puede tener este término, si los hubiere.
- Usage: dominios o sectores de la Seguridad Social en los que se utiliza.

En el momento de desarrollarse Build 6, los dominios reconocidos en los Reglamentos eran:

- HC = *“health care area and determination of legislation applicable”* (asistencia sanitaria y determinación de la legislación aplicable). Corresponde a las series de formularios E100 y E500.
- OAP = *“pension area, survivors, orphans and invalidity, accident at work, occupational disease”*. Corresponde a pensiones tanto por contingencias comunes como derivadas de enfermedades profesionales. Formularios de la serie E200.
- UN = *“unemployment area”* (sector de desempleo). Formularios de la serie E300.
- CB = *“family benefits área”* (sector de Protección Familiar). Formularios de la serie E400.

En el documento que presentaba el repositorio se decía, entre otras cosas que se intentaba: *“Resolver los problemas de interoperabilidad semántica producidos porque*

el mismo dato puede tener diferentes interpretaciones en diferentes Estados miembros (por ejemplo, el nombre de familia de una mujer casada puede ser su nombre de nacimiento o el nombre de su marido o la combinación de ambos)” (TF, 2007).

A título de ejemplo, el modelo de datos de Build 6 y sus cuatro diagramas modelaban todos los datos relacionados con la persona y su institución (figura 17) pudiendo incluso ser reutilizados cuando se crease los bloques intersectoriales de datos para la creación de los formularios electrónicos.

El diagrama de la figura 17 es el principal de los cuatro desarrollados y muestra las entidades que son la parte central del modelo: *“Person and Institution”*, así como sus relaciones y partes. Los otros tres diagramas reflejan los periodos, las prestaciones para un sector determinado y las solicitudes de prestaciones.

Debe destacarse que el concepto de Persona como cliente de una institución de Seguridad Social es la parte central del modelo por su carácter común a todas las áreas. Hay que distinguir entre *“Person concerned”*, referido a la persona asegurada, y *“Other persons”*, que hace referencia a los que tiene que ser considerados en función de la relación de dependencia con la persona asegurada, como es el caso del cónyuge o ex cónyuge de una persona asegurada, los hijos, etc.

Si la persona asegurada fallece o desaparece, la información adicional estará disponible en las entidades *“Death res. Missing person”* que es una de las situaciones especiales.

Puede haber también información relacionada con otros fines, como formación académica o datos de salud. Este último caso es especialmente complejo por poder relacionarse a su vez con periodos de tiempo, datos de afiliación o prestaciones, etc. Para reflejar estos casos es necesario un diagrama adicional que no se ha incluido.

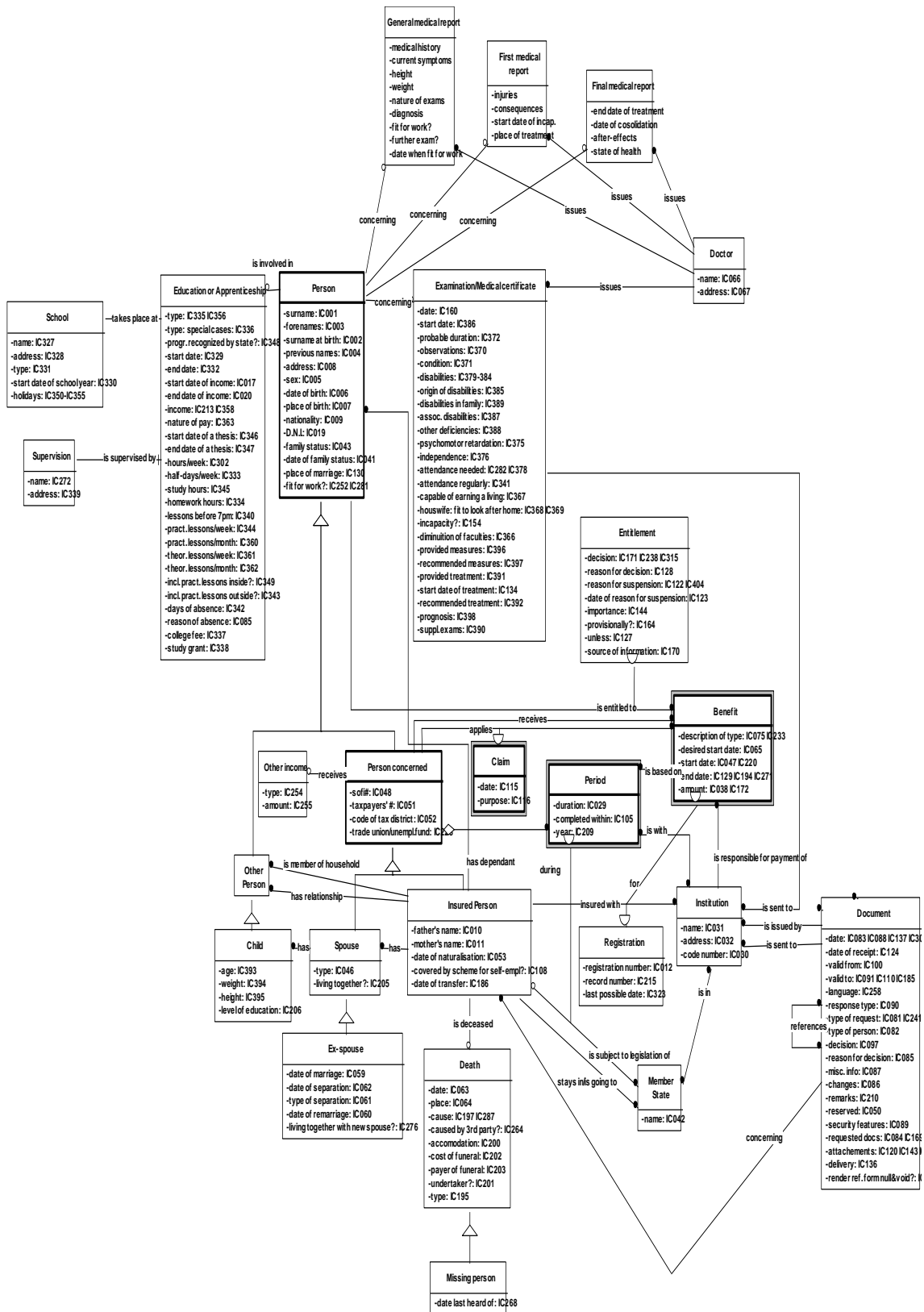


Figura 17. Modelo de datos relacionado con la persona y su institución (TF, 2007)

Cada uno de los datos del repositorio se definió en un Glosario de Términos. En primer lugar, mediante una información genérica, a continuación, con la información específica sobre ese término para cada estado miembro. En la sección correspondiente a vocabularios se profundizará en este glosario, pieza clave desde el punto de vista semántico.

2.3.3.4 Dimensión técnica

En general, la interoperabilidad técnica cubre los aspectos de la conexión y comunicación entre equipos, dispositivos informáticos, aplicaciones y servicios. Esto incluye aspectos claves como interfaces abiertos, servicios de interconexión, integración de datos y mediadores, presentación e intercambio de datos, localización y recuperación de recursos, accesibilidad, seguridad e integración de aplicaciones y servicios. En cada una de estas áreas se pueden identificar distintos estándares o especificaciones de uso extendido, si bien algunos de ellos aparecen también en otras dimensiones. Cabría mencionar los siguientes:

- **Interconexión.**

Precisa de políticas y tecnologías para conectar sistemas mediante protocolos de comunicación, como pueden ser TIC/IP, HTTP, FTP, SMTP, SOAP, CORBA y otros protocolos de amplia aplicación en Internet.

- **Intercambio de datos.**

Se basa en tecnologías y estándares para la descripción de la estructura y codificación de los datos para permitir el intercambio de la información entre sistemas que pueden tener distintas formas de representación interna de los datos. Algunos de los estándares más utilizados en la actualidad son: sistemas universales de codificación de caracteres como Unicode, el lenguaje de marcado XML (*eXtensible Mark-Up Language*), mecanismos para la transformación y presentación de documentos como el lenguaje XSL (*eXtensible Stylesheet Language*), el uso de extensiones consistentes como S/MIME (*Secure/Multipurpose Internet Mail Extensions*), formatos de texto ligero como JSON (*JavaScript Object Notation*), entre otros.

- **Presentación de la información.**

Mediante el uso de formatos de ficheros estandarizados como TXT, PDF, JPEG, PNG, HTML, XHTML, XML, etc., que puedan ser entendidos y utilizados convenientemente por sistemas heterogéneos para representar con exactitud información proveniente de múltiples fuentes.

- **Metadatos para la descripción de procesos y datos.**

Siguiendo mecanismos de descripción como el Modelo Entidad Relación empleado para estructurar la información en las bases de datos relacionales, o esquemas XML (*XML Schema*) que también permiten dotar de una determinada estructura y contenidos a los distintos tipos de documentos.

- **Localización y recuperación de información.**

Hacen referencia a los mecanismos de búsqueda y localización de recursos (servicio de directorio como DNS, o protocolos para la consulta de redes como LDAP), así como a los estándares de metadatos y vocabularios controlados que permiten la descripción consistente de recursos, como es el caso del marco de descripción RDF (*Resource Description Framework*) o de la definición de ontologías como OWL (*Ontology Web Language*).

- **Servicios de negocio.**

Cada proceso de negocio se compone de un conjunto de servicios integrados y coordinados, para su implementación se recomienda el uso de servicios web. Estos servicios se apoyan en un conjunto de estándares (WSDL, 2007; UDDI, 2004; y SOAP, 2007) basados en XML. Además de los estándares anteriores es necesario utilizar los llamados estándares de servicios web de segunda generación conocidos como "WS-*". También debemos tener en cuenta las propuestas del organismo encargado de definir y asegurar la interoperabilidad para los servicios web llamado WS-I (*Web Services interoperability Organization*). Para la coordinación de los servicios web se utilizan mecanismos de orquestación de servicios tales como *Web Services Business Process Execution Language* (WSBPEL, 2007).

- **Integración de aplicaciones.**

Permite lograr la interoperabilidad y organización del flujo de información entre aplicaciones heterogéneas, es decir, asegurar la comunicación entre las distintas aplicaciones. Tecnologías destacadas en este aspecto son EAI (*Enterprise Application Integration*) (Frantz et al., 2015) y ESB (*Enterprise Service Bus*) (Chappell, 2004).

- **Integración de servicios.**

Para la integración de servicios o procesos de negocio se utilizan las llamadas arquitecturas orientadas a servicios (SOA). Es un concepto de arquitectura de software que define la utilización de servicios para dar soporte a los procesos de negocio. Permite la creación de sistemas de información altamente escalables que reflejan el negocio de la organización, a su vez brinda una forma bien definida de exposición e invocación de servicios (comúnmente servicios web), lo cual facilita la

interacción entre diferentes sistemas propios o de terceros. Una versión más evolucionada la constituyen los servicios web RESTful, diseñados de acuerdo con las restricciones REST (Mehta, B., 2014). Se tratarán más en detalle en la sección 4.2.3.

Otros aspectos técnicos destacados son los dirigidos a la identificación única de recursos (como los URI), o de elementos y atributos empleados en esquemas XML y otros esquemas de metadatos (como los espacios de nombres XML). Y, por último, en relación con la seguridad, es importante tener en cuenta otras tecnologías y estándares para el encriptado de datos, firmas digitales y otros protocolos de transmisión segura de información.

En el “*New Zealand E-Government Interoperability Framework*” (NZ e-GIF, 2008) se establece un modelo que ayuda a entender la funcionalidad que debe tener un sistema en cada una de las capas en que se divide el modelo (figura 18).



Figura 18. Modelo en capas de NZ e-GIF para la interoperabilidad técnica

Las cuatro capas básicas del modelo son las siguientes:

- **Red (*Network*).**

Cubre los detalles del transporte de datos, tales como protocolos de red. Esta es un área crucial para la interoperabilidad. Sin un acuerdo sobre estándares de red, es difícil o imposible hacer que los sistemas se comuniquen.

- **Integración de datos (*Data integration*).**

Facilita la interoperabilidad de datos y su procesamiento. Los estándares utilizados permiten el intercambio de datos entre distintos sistemas y el análisis de los datos recibidos.

- **Servicios empresariales (*Business services*).**

Permite el intercambio de datos en las aplicaciones de negocio y contextos de información. Algunos de los estándares en esta capa son de carácter genérico, que abarca múltiples contextos de información de negocios. Otros trabajan con los estándares de integración de datos para definir el significado de los datos para asociarlos con las aplicaciones de negocio.

- **Acceso y presentación (*Access and presentation*).**

Cubre cómo los usuarios acceden y utilizan los sistemas de negocio.

Además de estas capas básicas, el modelo propone que todas las capas deben tener en cuenta los aspectos de seguridad y también las mejores prácticas en la utilización de estándares. También recomienda la utilización de servicios de *e-Government* disponibles para los organismos del sector público y la utilización de servicios web para la interconexión de aplicaciones.

2.4 Vocabularios

2.4.1 Conceptos generales

Según Criado et al. (2010), la interoperabilidad semántica requiere de sistemas de clasificación e indización de la información y el conocimiento que utilicen relaciones de modelos de datos comunes de intercambio, que estén actualizados, definidos y codificados según estándares. Estos aspectos adquieren un papel destacado para la reutilización de la información.

Herramientas clave en la construcción de estos sistemas y, por tanto, en la referencia y organización del conocimiento, las constituyen los **metadatos** y los **vocabularios controlados**, listas de términos que, entre otros aspectos, reducen las ambigüedades semánticas, incrementan la consistencia en la representación y facilitan las búsquedas.

Los vocabularios controlados se consideran la metodología fundamental para abordar y solucionar los problemas semánticos (Fraunhofer, 2009). Su contenido consiste en palabras estandarizadas y frases utilizadas para referirse a ideas, características físicas, personas, lugares, eventos, temas y otros conceptos. Permiten una categorización, indexación y recuperación de la información.

En definitiva, la utilización de vocabularios controlados proviene de la necesidad de tener una interpretación común sin posible ambigüedad. Esta circunstancia se considera un prerequisite cuando varios socios necesitan comunicarse con precisión y eficacia para producir resultados correctos. Para asegurar la consistencia y la alta calidad del vocabulario, debe controlarse de acuerdo con Prytherch (1999), que lo

define como *“una lista concreta de términos con un significado fijo e inalterable (...) el control intenta evitar la dispersión de temas relacionados bajo diferentes cabeceras (...). La lista solo puede ser modificada o extendida por su publicador o la agencia responsable”*.

Conviene distinguir entre vocabularios controlados y formatos controlados. Los vocabularios controlados se organizan en conjuntos de valores con una terminología controlada, mientras que los formatos controlados se refieren a reglas a los tipos de datos posibles y el formato de la información.

Los formatos controlados son apropiados para almacenar medidas, coordenadas geográficas y otros tipos de información en campos en los que los números o códigos son utilizados. Las restricciones pueden referirse a los campos de los registros para regular el número de dígitos permitidos, la forma de expresar los decimales o números negativos y, en un caso ideal, la adecuación a las normas ISO, NISO (2004) u otro estándar apropiado cuando sea posible.

También es frecuente su utilización para fechas de todo tipo. En este caso, se combina su utilización con un campo de visualización de datos.

2.4.2 Tipos de Vocabularios Controlados

El tipo de vocabulario controlado a elegir depende de la clase de problemas que se presenten para el caso concreto. En función de su complejidad y características será apropiado para manejar determinados tipos de conflictos.

Si bien no hay un consenso general sobre estos términos, la bibliografía distingue entre los siguientes tipos de vocabularios controlados (Hilera, et al., 2010):

- Glosarios (diccionarios) y listas de códigos.

Son la forma más sencilla de vocabularios controlados y también la más extendida. Su propósito general es proporcionar una comprensión de los términos y códigos en su dominio de aplicación. Normalmente no tienen una estructura determinada, pero su capacidad para representar sinónimos y homónimos aporta una importante ayuda para la resolución de conflictos relacionados con este tipo de situaciones.

- Las taxonomías o listas con estructuras relacionales jerárquicas de términos que representan conceptos.

Constituyen el nivel siguiente de complejidad. Facilitan el tratamiento de conflictos semánticos que no pueden solucionar los glosarios por estar relacionados con las relaciones jerárquicas entre los términos y la submisión entre ellos.

- Los tesauros o listas con conexiones horizontales polijerárquicas en lenguaje de indexación dinámico. Según la *International Organization for Standardization (ISO)*, un tesoro es “*un vocabulario de lenguaje controlado indexado, formalmente organizado para hacer explícita a priori relaciones entre conceptos*”.

Las relaciones semánticas entre conceptos se hacen explícitas mediante la introducción de relaciones asociativas. Como consecuencia, las similitudes semánticas de los valores y esquemas de nombres, utilizados para identificar posibles sinónimos, son mucho más eficaces.

- Las ontologías o especificaciones explícitas formales de un cuerpo de conceptos compartidos

Según Ding (2002), las ontologías promueven la estandarización y reutilización de la representación de la información mediante la identificación común y compartiendo el conocimiento.

Se componen de conceptos propiamente dichos, relaciones entre conceptos o entre propiedades de conceptos, instancias y axiomas. Introducen la capacidad de definir relaciones entre los términos en lugar de restringirlas a un conjunto fijo como ocurre en los tesauros. Uno de los propósitos generales es la integración de un modelado de información y terminología que presenta la ventaja de no tener que definir estos modelos de forma exhaustiva en la parte técnica.

Las ontologías facilitan la comunicación entre máquinas y personas y pueden llegar a permitir el tratamiento inteligente de la información por máquinas.

- Las redes semánticas o estructuras complejas.

Representan conocimiento a través de grafos cuyos nodos simbolizan conceptos, propiedades e instancias y cuyas aristas implican relaciones entre los mismos. Las redes semánticas permiten el razonamiento automatizado.

Debe prestarse especial atención al tema lingüístico a la hora de considerar el uso de vocabularios. En el caso de la Unión Europea, la existencia de legislación común (reglamentos y directivas) hace presumir una terminología equivalente como parte del proceso legislativo. Sin embargo, eso exige que ese vocabulario se utilice también a nivel nacional y eso, a su vez, exige un esfuerzo de armonización.

2.4.3 Vocabularios controlados en la Unión Europea

En la sección 2.2.2 se ha hablado de las aportaciones realizadas por la Unión Europea en materia de interoperabilidad. En esta sección se van a tratar los desarrollos realizados en lo que a vocabularios controlados se refiere. En primer lugar, destacan

los artefactos desarrollados para su reutilización en los modelos de datos, como son los *Core Vocabularies*. Por otra parte, se va a mencionar la mayor aportación del proyecto TESS en el campo semántico: el Glosario de Términos.

Core Vocabularies

Entre los recursos que desde 2012 ha proporcionado ADMS (2013), están la plataforma Joinup (EU, 2017) implementada bajo el programa ISA, que incluye los *Core Vocabularies* y los *Core Concepts*.

Según se indica en el “*e-Government Core Vocabularies Handbook*” (e-Government, 2015), los *Core Vocabularies* pueden ser utilizados para reducir los conflictos semánticos de dos maneras:

- Diseñando nuevos modelos de datos que extiendan los *Core Vocabularies* haciendo uso de los definidos anteriormente (*building blocks*) tanto a nivel conceptual como sintáctico, garantizando un mínimo de interoperabilidad entre distintos sectores.
- Estableciendo correspondencias entre los modelos de datos y los *Core Vocabularies*, de forma que estos últimos sirvan de base para comparar los modelos de datos existentes.

Los “*Core Concepts*” son modelos de datos simplificados, reusables y extensibles que capturan las características fundamentales de una entidad de la manera independiente del contexto. Pueden ser representados como *Core Vocabularies* usando diferentes formalismos, tales como esquemas RDFS, XMLS o JSON.

Se han desarrollado cuatro vocabularios hasta el momento (EU, 2017):

- **Core Public Services**

Es un modelo de datos simplificado, reusable y extensible que captura las características fundamentales de un servicio ofrecido por una administración pública. Incluye características del servicio público como título, descripción, entradas, salidas, proveedores, localizaciones, etc.

- **Core Location**

Es un modelo de datos simplificado, reusable y extensible que captura las características fundamentales de una localización, representadas como dirección, nombre geográfico o geometría. Está alineado con las especificaciones INSPIRE (2017), publicadas por la W3C (2015).

- **Core Business**

Es un modelo de datos simplificado, reusable y extensible que captura las características fundamentales de una entidad legal, como el nombre, la actividad, dirección, identificador legal, tipo de compañía y sus actividades. Ha sido revisado y renombrado como “*Registered Organización Vocabulary*” (W3C, 2013). El 8 de enero de 2013, la representación RDF de este vocabulario se publicó formalmente en los estándares W3C como “*Public Working Draft*”.

- **Core Person**

Es un modelo de datos simplificado, reusable y extensible que captura las características fundamentales de una persona, como nombre, género, fecha de nacimiento, etc. (tabla 7). Su desarrollo ha sido dirigido por la Unidad de Cooperación Europea (Eurojust, 2002), quien promueve la interoperabilidad en el dominio judicial a través del proyecto EPOC (2004). Está también publicado por la W3C (2012).

SEMIC.EU Core Person
Full Name [0..1]
Given Name [0..1]
Family Name [0..1]
Gender [0..1]
Date of Birth [0..1]
Place of Birth [0..1]
Country of Birth [0..1]
Citizenship [0..*]

Tabla 7. Elementos de *Core Person* (EU, 2017)

Por encima es de estos debería completarse el modelo para un dominio concreto. En nuestro caso, Seguridad Social.

Glosario de Términos (proyecto TESS)

Una vez más, se va a citar el proyecto EESSI y su predecesor, el proyecto TESS.

En la sección 2.3.3.3 se ha descrito la creación de un repositorio de información y de un modelo conceptual de datos para los intercambios previstos en el proyecto TESS, datos que deberían servir para una estandarización de los mensajes en materia de

Seguridad Social. Los términos que integraban ese modelo de datos eran descritos en un glosario de términos del que nos ocupamos a continuación.

El Glosario de Términos definió, en primer lugar, cada uno de los datos del repositorio de información con una definición genérica, seguida de una definición específica para cada Estado miembro con la intención de permitir a la institución un mejor procesamiento del dato recibido del exterior.

Los objetivos marcados en ese momento para el GOT eran:

- Proporcionar a los funcionarios una ayuda para localizar los términos equivalentes utilizados por otro estado.
- Proporcionar descripciones genéricas y específicas de los términos intercambiados entre los Estados miembros. En ese sentido el GOT debía permitir una comunicación más eficiente entre las instituciones europeas de Seguridad Social y la necesaria optimización de las operaciones tanto en un entorno papel como electrónico.

Veamos, como ejemplo, la definición del término “nombre”. Se han incluido tanto las definiciones genéricas como las específicas por países (tabla 8).

ID	Country	Name	Data element	Description	Values	Remark	Domain
T0001	GENERIC	Surname	IC001	State the official family name of a client.			ALL
T0001	BEL	Surname	IC001	Form E6XX not used Official surname given to a person			ALL
T0001	D	Surname	IC001	HC: Official <u>family</u> surname as it appears in valid personal identity papers". When using electronic data exchange, form E110 is no longer needed. E123 is exclusively a form relating to statutory accident insurance. CB: Corresponds to the GENERIC Glossary of Terms UN: Official family name (surname) of a person.		For UN : E301-1.1-1 E303-1.1-1	ALL
T0001	IRL	Surname	IC001	State the clients family name. Note: National applications format consists of 20 characters (free text)	Free Text	E105,E206,E211, E215,E302,E402, E403,E404,E406, E407,E413 & E602 not currently used by Ireland	ALL
T0001	LUX	Surname	IC001	OAP (CPACI) : See GOT Framework if not married, same as IC002 otherwise, married name		OAP (CPACI) : wife can keep her maiden name, therefore if name and maiden name are the same, married status is not evident	
T0001	LUX	Surname	IC001	OAP (CPEP): See GOT Framework if not married, same as IC002 otherwise married name		OAP (CPEP): wife can keep her maiden name, therefore if name and maiden name are the same, marital status is not evident	
T0001	LUX	Surname	IC001	OAP (E500): See GOT Framework for a married woman: married name otherwise, maiden name (same as name at birth)			
T0001	LUX	Surname	IC001	UN: See GOT Framework maiden name in the former surnames box (E 301) if married			

ID	Country	Name	Data element	Description	Values	Remark	Domain
T0001	LUX	Surname	IC001	CB: See GOT Framework		CB: If wife, this box is completed with the maiden name for the 401/411	CB
T0001	NED	Surname	IC001	See GOT Framework		For the surname the standard NEN 1888 (the Dutch national standard concerning personal data) applies. The complete name of a person consists of several data-elements that are structured in a standard way. The data regarding the name of a person with the exception of the Christian names, consisting of a significant part and possibly a prefix.	ALL
T0001	POR	Surname	IC001	Full name as it appears on the identity card or passport.			ALL
T0001	UK	Surname	IC001	<i>That part of a person's name which is used to describe the family, clan, tribal group, marital or professional association.</i>		Further work is being done on the use and format of the surname to enable DSS to conform to the legal aspects i.e. the Data Protection Act.	
T0002	GENERIC	Surname at birth	IC002	State the official family name of a client, given at his birth. Note: The maiden name of a woman is equivalent to the surname at Birth			ALL
T0002	BEL	Surname at birth	IC002	Form E6XX not used The first official surname give to a person is the surname at birth. This birth name can be replaced by another name following an adoption or a request for a change of name.			ALL
T0002	D	Surname at birth	IC002	HC: Corresponds to GENERIC Glossary of Terms Comment: In form E122-3.2-2 there is the concept "Previous surnames". This does not correspond to data element IC002!. The concept "Maiden surname" in forms E101 and E102 is replaced in brief by "Previous surnames". OAP: See GOT Framework CB: Corresponds to the GENERIC Glossary of Terms UN: Official surname of a person at birth. In the UN forms : Former name	UN: see also IC004	For OAP: "Maiden surname" = "Surname at birth". For HC: "Previous name" (IC004) = "Previous surname held" is not the same as "Surname at birth", therefore the concept of the "Maiden surname" should not be replaced by "Previous surnames" For UN: Forms E303-1, 1-3, E303-1.1-3	All

ID	Country	Name	Data element	Description	Values	Remark	Domain
T0002	IRL	Surname at birth	IC002	State the clients surname at birth. Note: National applications format consists of 20 characters (free text) E213-2.2.2 does not refer to Maiden Name	Free Text	E213,E215 & E413 not currently used by Ireland	ALL
T0002	NED	Surname at birth	IC002	{Though maiden name is a common often used term, it is hard to describe it in a clear way. This is caused by the fact that in our opinion a so called maiden name of a person don't necessarily has to be the same as the surname at birth of that person (due to f.e. adoption or change of surname).The national standard concerning personal data, as well as the GBA (Municipal basis administration) which is the formally authorised source of personal data in the Netherlands, therefore do not distinguish maiden name as a data-element}		The complete name of a person consists of several data-elements that are structured in a standard way. Since a surname is part of the complete name of a person, the standard structure also applies to surname at birth. For addressing a married woman our Social Security domain uses the composite data-elements surname and surname husband, but only if the woman wishes to be addressed that way.	ALL
T0002	POR	Surname at birth	IC002			Not used	
T0002	UK	Surname at birth	IC002	<i>The surname given to a person at birth as used for the exchange of information with member states of the European Union.</i>		This item is also currently under discussion. The validation procedures associated with this item in the UK is very stringent and the same procedures outside the UK are not necessarily applied or indeed if the same definition is understood. We are therefore looking at the use of another item (possibly base surname) to interpret the information received from the EEC members thereby complying with our current legislation.	

Tabla 8. Definición del término nombre (TF, 2007)

En los estudios previos para el desarrollo del proyecto EESSI, se dijo que este trabajo podría ser reutilizado especialmente la el desarrollo de los SED, pero finalmente se decidió abordar su diseño desde cero.

2.5 Metadatos

Tal como se ha dicho anteriormente, para lograr el entendimiento de la información intercambiada por vía telemática, el sólo uso de vocabularios controlados no es suficiente, siendo necesarios modelos de conocimiento más ricos semánticamente, como taxonomías y ontologías, que definen los conceptos de un determinado dominio, así como sus relaciones. El primer lugar de complejidad lo constituirían los conjuntos de metadatos.

2.5.1 Definición

Conviene empezar por abordar el concepto de metadato. Según DAMA (2009), metadatos son “datos sobre datos”, si bien no siempre se utiliza de acuerdo con el estricto sentido etimológico, dependiendo su interpretación del entorno en el que se use. Así podemos encontrar metadatos sobre datos de transacciones, eventos, objetos, relaciones, etc. En definitiva, los metadatos son los bloques de información básica para el desarrollo de aplicaciones y sistemas de gobierno electrónico. Pueden ser definidos una sola vez, pero compartidos y reutilizados en múltiples ocasiones. Esta reutilización conlleva los siguientes beneficios:

- Incrementa la interoperabilidad entre distintos sistemas que comparten datos y servicios.
- Reduce el costo en el desarrollo de sistemas interoperables, ya que evita el tratamiento de las ambigüedades dentro de las aplicaciones.
- Facilita el tratamiento de los datos en sistemas de información en general, y en particular en los que integran y comparten datos, por ejemplo, los sistemas de apoyo a la toma de decisiones.

Otra definición más amplia es la de NISO (2004): información estructurada que describe, explica, localiza o bien facilita su recuperación, utiliza o administra un recurso de información. También en este caso se matiza el diferente uso del término en diferentes comunidades. Así, mientras en ocasiones se usa para referirse a información interpretable por ordenadores, en diferentes contextos se utiliza sólo para describir recursos electrónicos. En el entorno de las bibliotecas, los metadatos son normalmente utilizados para cualquier esquema formal o descripción de recursos, aplicándose a cualquier tipo de objeto, digital o no digital. MARC21 (2010), la tradicional catalogación de las bibliotecas es una forma de metadatos. El conjunto de reglas con las que se utiliza, son estándares de metadatos. Otros esquemas de metadatos han sido desarrollados para describir varios tipos de objetos textuales y no textuales, incluidos libros publicados, documentos electrónicos, herramientas de ayuda para búsquedas, objetos de arte, material educacional y de formación y conjuntos científicos.

Parecido enfoque es el que recoge Sicilia (2006), asociando la descripción propia de la naturaleza de los metadatos al propósito para el que hayan sido concebidos. Así, pueden estar enfocados a facilitar una recuperación, establecer una propiedad intelectual, dirigir el comportamiento de un sistema online, facilitar la combinación de diversos recursos, etc.

El fin de proporcionar la información recogida para identificar un recurso digital es recogida también por Zeiger et al. (2006) y ANZLIC (2001).

En Ercegovac (1999) se establece que los metadatos describen atributos de un recurso dado, sea un objeto bibliográfico, un registro de un archivo inventario, un objeto geoespacial, un recurso en un museo o una implementación de software.

A su vez, los metadatos pueden agruparse en lo que se conoce como en la bibliografía como “*metadata Schemes*” (NISO, 2004). Se trata de conjuntos de elementos concebidos para un propósito concreto, como describir un tipo concreto de información de un recurso. La definición o significado de los elementos en sí mismo es conocido como la semántica del esquema. Los valores dados a los elementos de los metadatos son su contenido. Los esquemas de metadatos generalmente especifican los nombres de los elementos y su semántica. De forma opcional, pueden especificar reglas de contenido sobre cómo su contenido puede ser establecido (como establecer el título principal) y los valores posibles para el contenido. Así, por ejemplo, los términos pueden proceder de un vocabulario controlado concreto.

Puede haber también reglas de sintaxis sobre cómo los elementos y su contenido deben ser codificados. Un esquema de metadatos sin reglas de sintaxis establecidas es llamado de sintaxis independiente. Los metadatos pueden codificarse en cualquier sintaxis definible. Entre los que se pueden utilizar están SGML, XML o JSON.

2.5.2 Historia de los metadatos

La primera referencia al concepto de metadatos puede encontrarse en la década de 1960 (Manso-Callejo et al., 2010). El término fue acuñado por Myers y aparece en el campo de la gestión de bibliotecas.

El valor de los repositorios de metadatos empezó a reconocerse en la década de los 90. A mediados de esta década fueron convirtiéndose en más relevantes especialmente en las empresas que trataban de entender cuáles eran sus recursos de información. También se asocia su extensión a la aparición de la Web (Anastasiou, 2010) y en paralelo, con la necesidad de escalar y filtrar la información. Tim Berners Lee definió los metadatos en 1997 como “*información entendible por la máquina para la Web*”.

El primer workshop tuvo lugar en Dublin (Ohio) y se creó el conjunto de elementos “*Dublin Core Metadata*” (Dublin Core, 2009).

Se extendieron principalmente por el compromiso de la fecha tope del año 2000, lo que hizo emerger iniciativas muy concretas que exigieron estandarización dentro de las aplicaciones de las empresas.

Ejemplos de esa estandarización fueron el “*CASE Definition Interchange Facility*” desarrollado por *Electronics Industries Alliance* en 1995, así como la ya mencionada iniciativa *Dublin Core* de la que hablaremos extensamente más adelante. A su vez, el *Object Management Group* (OMG) desarrolló el “*Common Warehouse Metadata Model* (CWM)” en 1998. Por otra parte, Microsoft apoyó la alternativa denominada “*Metadata Coalitions*” en 1995. En el año 2000, los dos estándares se fundieron en CWM.

Los primeros años del siglo XXI se caracterizaron por la actualización de los repositorios de metadatos existentes por cuestiones relacionadas con la Web. Los metadatos empezaron a

convertirse en un producto más en la oferta de los vendedores. Sin embargo, relativamente pocas organizaciones realmente adquirieron o desarrollaron un repositorio de metadatos. Se dio esa circunstancia por diversas razones:

- La escasez de personas con auténtica formación y habilidades.
- La dificultad del esfuerzo.
- El escaso éxito de los esfuerzos iniciales de algunas empresas.
- El relativo estancamiento del mercado de herramientas después de su lanzamiento inicial a finales de los 90.
- La poca comprensión de los beneficios para el negocio.
- El excesivo énfasis puesto en las aplicaciones heredadas y en los metadatos técnicos.

Más recientemente, las empresas han empezado a darles mayor importancia a los metadatos. Cada vez está más extendida la necesidad de incorporar metadatos más allá de las tradicionales fuentes estructuradas. Algunos de los factores que están llevando a este interés renovado son:

- La reciente entrada en el mercado de grandes vendedores.
- El desafío que algunas empresas están afrontando al tratar de abordar requerimientos tanto normativos como internos mediante herramientas poco sofisticadas.
- La aparición de iniciativas que abarcan toda la empresa, como gobernanza de la información arquitectura empresarias, reutilización del software, etc.
- La mejora de los estándares de metadatos.
- El reconocimiento al más alto nivel por algunas empresas sofisticadas y grandes organizaciones, de que la información es un activo para algunas de ellas, el más crítico, que puede ser gestionado de forma eficaz.

La historia de las herramientas de gestión de metadatos se caracteriza por la falta de enfoque metodológico que sobre la gestión de la información prevalece en las organizaciones. La falta de estándares, y la propia naturaleza de la mayoría de soluciones para la gestión de metadatos, es la causa de que muchas organizaciones hayan prescindido de ellos. La creciente importancia de la información es lo que está llevando a los productos para la gestión de metadatos a una estandarización y a seguir un enfoque más metodológico.

2.5.3 Tipos de metadatos

DAMA (2009) clasifica los metadatos en cuatro grandes tipos: negocio, proceso técnico, proceso operacional y administración de datos.

Los metadatos de negocio incluyen nombres y definiciones de las áreas, conceptuales y temáticas, entidades, atributos, rangos de descripciones, cálculos, reglas de negocio, etc. Están relacionados con la perspectiva de negocio del usuario.

Los metadatos técnicos y operacionales proporcionan a los desarrolladores y a los usuarios técnicos información sobre sus sistemas. En este tipo se incluyen las tablas de base de datos físicas, columnas de nombres, propiedades y otras propiedades. Son utilizados por los sistemas de gestión de base de datos y otros softwares similares.

Los metadatos operacionales están orientados a cubrir las necesidades de los usuarios de tecnologías de la información, como información sobre el movimiento de los datos, su fuente o los sistemas de destino, programas batch que los utilizan, anomalías de planificación, información de backup, etc.

Los metadatos sobre procedimientos son datos que definen y describen las características de otros elementos del sistema, como reglas de negocio, programas, herramientas, etc.

Los metadatos de administración de datos son datos sobre la administración, los procesos de administración y la asignación de responsabilidades. Aseguran que los metadatos sean precisos y de alta calidad en toda la empresa. Establecen y controlan la compartición de datos.

Por su parte, NISO (2004) identifica tres tipos de metadatos: descriptivos, estructurales y administrativos.

Los metadatos descriptivos tienen un enfoque de identificación y localización de la información. Pueden incluir elementos como título, resumen, autor, palabras clave, etc. Conceptualmente son similares a lo que DAMA clasifica como metadatos de negocio.

Los metadatos estructurales indican la forma en que objetos compuestos están agrupados. Por ejemplo, cómo se ordenan las páginas para formar un capítulo. En cierto modo, entrarían también en la categoría de metadatos de negocio.

Los metadatos administrativos responden a un concepto similar a la interpretación de DAMA, poniéndose como ejemplo información referida a la forma en que el recurso se ha creado, cuándo y por quién, tipo de fichero, etc.

Además de estas dos clasificaciones, podría citarse un último tipo como son los metadatos para datos no estructurados, si bien estos están fuera del contexto de esta tesis.

2.5.4 Modelo de madurez de metadatos

Como punto de referencia, en ISA (2011) se propone un esquema con cinco niveles de madurez:

- **Nivel 1 – Ignorancia de metadatos:** Los metadatos no están documentados posiblemente porque los administradores no son conscientes de su importancia. Esta situación da lugar a problemas de interoperabilidad semántica graves. Al utilizar modelos de datos propietarios, su integración con otros sistemas de *e-Government* se hace costosa y difícil.
- **Nivel 2 – Metadatos cerrados:** Los metadatos pueden estar parcialmente documentados, pero o bien no los están de forma centralizada y estructurada, o bien no están disponibles y accesibles bajo una licencia abierta para su compartición y reutilización por parte de los desarrolladores.

Es el caso de las instituciones de Seguridad Social, que se ven en la necesidad de organizar los metadatos dispersos en repositorios estructurados, catálogos o bibliotecas y proporcionar acceso a dichos recursos. Se da frecuentemente en el momento de poner en funcionamiento los acuerdos que implican intercambios de datos.

- **Nivel 3 – Metadatos abiertos para los seres humanos:** Los metadatos son documentados y están disponibles como "metadatos abiertos" para su reutilización, pero no están sistemáticamente publicados en un formato reutilizable. Por ejemplo, sólo está disponible en formato "pdf" o documentos "doc".

Es el caso de las instituciones que empiezan a estar familiarizadas con el mundo del gobierno electrónico y que comienzan a ofrecer servicios de consulta, búsqueda y exportación de sus metadatos en un formato legible por la máquina y, preferentemente, sin derechos de propiedad. Es el caso de CSV o XML.

- **Nivel 4 – Metadatos abiertos y reutilizables:** Los metadatos son administrados centralmente y publicados como "metadatos abiertos", en un formato legible por máquina o bien mediante un API para ordenador que permita el acceso, consulta y la reutilización de los repositorios de metadatos disponibles, catálogos, bibliotecas, etc. A través de estos sistemas los usuarios pueden encontrar, explorar, comparar, descargar y usar los metadatos que mejor se adapten a sus necesidades y proyectos.

Para llegar a este nivel sería necesario considerar la aplicación de políticas relacionadas con los metadatos, incluyendo el uso de RDF para documentarlos, el diseño persistente, uso y mantenimiento de las URI (*Universal Resource Identifier*), con enlaces a vocabularios externos o esquemas de datos, armonizar sus recursos a los recursos de terceros, etc.

- **Nivel 5 – Metadatos abiertos y relacionados:** Los activos semánticos se documentan con sus datos relacionados y son procesados por avanzados sistemas de gestión de metadatos. En este nivel se genera un gráfico de conceptos relacionados entre sí como definiciones de metadatos que surgen del uso de definiciones de otros vocabularios. Cada concepto representado en este gráfico corresponde a una URI única.

En este nivel, las instituciones de Seguridad Social deberían considerar la utilización de ontologías, haciendo uso de OWL para la representación de modelos de conocimiento. No se ha identificado ningún caso de institución o acuerdo internacional que pueda considerarse en este nivel.

Si se llegase a un acuerdo en el establecimiento de los metadatos entre las instituciones de la Seguridad Social, sería posible la creación de servicios públicos interoperables. Para conseguirlo se deben identificar y documentar los metadatos que las distintas instituciones utilizan para ponerlas en común. De esta forma se puede crear un sistema de gestión de metadatos (*Metadata Management System*) común a todas las instituciones.

2.6 Estándares

En esta sección se va a hacer referencia a los estándares relacionados con las tecnologías semánticas mencionadas.

2.6.1 AGLS

Dentro del ámbito de la Administración Pública, AGLS (2010) es el estándar australiano, basado en *Dublin Core* y utilizado para describir recursos de Gobierno electrónico. Los recursos pueden ser documentos, imágenes, sonido, vídeo, objetos físicos, personas y servicios. Toma trece de sus elementos de forma obligatoria, dejando los restantes de forma recomendada u opcional. Incluye una serie de extensiones hasta un total de 60 elementos.

A pesar de que el estándar australiano AS 5044 sólo exige el uso de tres elementos para crear un registro de metadatos conforme con AGLS, el gobierno australiano exige el uso de los elementos adicionales. Estos elementos son exigidos para sustentar el marco de desarrollo de portales y cualquier iniciativa futura que requiera metadatos bien constituidos para la descripción de la información y servicios del gobierno australiano. Se recomienda a las agencias el uso de cuantos elementos adicionales sean necesarios para mejorar los recursos de descripción y facilitar su localización.

2.6.2 DUBLIN CORE

El conjunto de elementos de metadatos *Dublin Core* (Dublin Core, 2009) es un vocabulario de quince propiedades o definiciones semánticas para la descripción de recursos. Se describe en la norma ISO 15836:2017 (ISO, 2017) como parte de un conjunto de vocabularios de metadatos y de especificaciones técnicas más amplio mantenido por la agencia del *Dublin Core* (DCMI).

Además de estos elementos básicos (ninguno obligatorio y todos repetibles) existen otros mecanismos que sirven para adaptar el *Dublin Core* a las necesidades concretas de información y que hacen que este modelo de metadatos sea aplicable a cualquier proyecto de sistema o servicio de información digital. Estos mecanismos son fundamentalmente son:

- Los perfiles de aplicación, desarrollados para el uso del *Dublin Core* asociado a distintas disciplinas. Un ejemplo es el perfil de aplicación para bibliotecas: DC-Lib.
- Los términos de metadatos (*metadata terms*), donde se incluyen tanto los nuevos elementos que se van incorporando al vocabulario DC, como las matizaciones de elementos ya existentes, esquemas de codificación (antes conocidos como "cualificadores") y términos de vocabularios controlados. Todos estos términos de metadatos sirven para adecuar y precisar el valor y la utilidad de la meta información expresada a través de DC.

Los metadatos de *Dublin Core* están siendo utilizados como la base para los sistemas descriptivos para varios grupos de interés, como, por ejemplo:

- Instituciones de gobierno.
- Sector científico de la investigación.
- Autores de páginas web.
- Negocios que requieren lugares más investigables.
- Corporaciones con sistemas de gerencia extensos en conocimiento.

Las implementaciones de *Dublin Core* usan generalmente XML y se basan en el *Resource Description Framework*.

2.6.3 Esquema Nacional de Interoperabilidad de España

El Esquema Nacional de Interoperabilidad de España (ENI, 2007) está regulado por el Real Decreto 4/2010, de 8 de enero. Persigue la creación de las condiciones necesarias para garantizar el adecuado nivel de interoperabilidad técnica, semántica y organizativa de los sistemas y aplicaciones empleados por las administraciones públicas, que permita el ejercicio de derechos y el cumplimiento de deberes a través del acceso electrónico a los servicios públicos, a la vez que redunde en beneficio de la eficacia y la eficiencia. Con ese fin, introduce los elementos comunes que han de guiar la actuación de las administraciones públicas en materia de interoperabilidad. En particular, introduce los siguientes elementos principales:

- Los principios específicos de la interoperabilidad.
- Las dimensiones de la interoperabilidad organizativa, semántica y técnica.
- Las infraestructuras y los servicios comunes, elementos reconocidos de dinamización, simplificación y propagación de la interoperabilidad, a la vez que facilitadores de la relación multilateral.

- La reutilización, aplicada a las aplicaciones de las administraciones públicas, de la documentación asociada y de otros objetos de información.
- La interoperabilidad de la firma electrónica y de los certificados.
- La recuperación y conservación del documento electrónico.
- Se crean las normas técnicas de interoperabilidad y los instrumentos para la interoperabilidad, para facilitar la aplicación del Esquema.

Por otra parte, las Normas Técnicas de Interoperabilidad (NTI, 2013) desarrollan aspectos concretos de diversas cuestiones necesarios para asegurar las facetas más prácticas y operativas de la interoperabilidad entre las Administraciones públicas y con el ciudadano.

En concreto, el Anexo III de la Norma Técnica de Interoperabilidad de Reutilización de recursos de la información (BOE, 2013) está dedicado a describir los **metadatos de conjuntos de datos** del catálogo y los conjuntos de datos incluidos en él. Asimismo, se incluye el término recomendado para su presentación usando vocabularios estándar que se identifican por las abreviaturas de su espacio de nombres. Además de la denominación, descripción del metadato y el tipo de dato que se debería usar para la representación, se especifica si es obligatorio y si admite más de un metadato de ese tipo. La tabla 9 recoge los vocabularios recomendados.

VOCABULARIO	URI
XML Schema	xsd: http://www.w3.org/2001/XMLSchema#
Simple Knowledge Organization System (SKOS)	skos: http://www.w3.org/2004/02/skos/core#
Dataset Catalog (dcat)	dcat: http://www.w3.org/ns/dcat#
Dublin Core Terms	dct: http://purl.org/dc/terms/
W3C Time Ontology	time: http://www.w3.org/2006/time#
Friend Of A Friend (FOAF)	foaf: http://xmlns.com/foaf/0.1/

Tabla 9. Vocabularios recomendados por el ENI (BOE, 2013)

2.6.4 HL7 (ISO 21090)

Fundada en 1987, *Health Level Seven International* (HL7, 2017) es una organización sin ánimo de lucro dedicada a establecer un marco para el desarrollo de estándares para el intercambio, integración, compartición y recuperación de información relacionada con la salud, muy especialmente en lo que se refiere a las historias clínicas. La palabra “*Health*” (Salud) hace referencia al área de trabajo de la organización y las palabras “*Level Seven*” (Nivel Siete) hacen referencia al último nivel del modelo de comunicaciones para interconexión de sistemas abiertos (OSI, *Open Systems Interconnection*) de la Organización Internacional para la Estandarización (ISO).

HL7 es en realidad un conjunto de estándares para facilitar el intercambio electrónico de información clínica. Utiliza una notación formal de modelado de datos (UML) y un lenguaje extensible de marcado de etiquetas (XML). Cuenta con un estricto y bien definido conjunto de procedimientos de operación que garantizan el consenso, la transparencia y el equilibrio de intereses.

Su especificación más utilizada es un estándar de mensajería para el intercambio electrónico de datos en salud.

Los miembros de HL7 son conocidos colectivamente como el Grupo de Trabajo, que está organizado en comités técnicos y grupos de intereses especiales. Los comités técnicos son directamente responsables por el contenido de los estándares. Los grupos de intereses especiales sirven para desarrollar información y pruebas para la exploración de nuevas áreas de cobertura de los estándares publicados por HL7.

Las principales estrategias de HL7 son:

- Desarrollar estándares cuya sintaxis coherente y extensible permita estructurar información en salud, que a su vez permita apoyar los procesos de atención al paciente, para ser intercambiada entre aplicaciones de software, conservando al mismo tiempo la semántica de la información.
- Desarrollar una metodología formal para apoyar la creación de estándares HL7 a partir del Modelo de Referencia de Información (RIM) de HL7.
- Formar al sector de la salud, a los encargados de formular políticas y al público en general, sobre los beneficios de la normalización de información sanitaria y específicamente sobre los estándares HL7.
- Promover el uso los estándares HL7 en todo el mundo a través de la creación capítulos internacionales de HL7, que participen en la elaboración y adopción de estándares HL7, donde sea necesario.
- Estimular, fomentar y facilitar expertos en dominios para que el sector salud y las organizaciones interesadas en participar en HL7 puedan desarrollar estándares de información en salud en su área de especialización.
- Colaborar con otras organizaciones desarrolladoras de estándares y organismos nacionales e internacionales de estandarización (como es el caso de ANSI e ISO), para desarrollar infraestructuras de dominios de información en salud que contribuyan a promover el uso de estándares compatibles.
- Colaborar con los usuarios de tecnologías de información en salud para asegurarse que los estándares HL7 cumplan con los requerimientos del mundo real y que los esfuerzos de desarrollo iniciados por HL7 permitan satisfacer nuevos requerimientos.

2.6.5 IEEE/LOM: ISO/ISO 19788

IEEE/LOM (*Learning Object Metadata*), es un estándar utilizado para la descripción de objetos de aprendizaje (IEEE/LOM, 2017). Se desarrolló y formalizó a través de IEEE y el *Learning Technology Standards Committee*. Estrechamente relacionado se encuentra el ISO/IEC 19788 *Metadata Learning Resource* (Pons, 2011), publicado en enero de 2011,

Su fin inicial es especificar unos elementos de metadatos y sus atributos para la descripción de recursos de aprendizaje. Incluye reglas para controlar la identificación de los elementos de datos y sus atributos. También proporciona principios, reglas y estructuras para la especificación de esos mismos recursos. Los principios clave establecidos en el ISO/IEC 19788-1:2011 son informados en el contexto de los requerimientos de usuario con el fin de permitir adaptaciones multilingüísticas y multiculturales desde una perspectiva global.

En principio, ISO/IEC 19788-1:2011 es neutral desde un punto de las tecnologías de la información y define un enfoque común metodológico que puede utilizarse para el desarrollo de las distintas partes del estándar.

Su conjunto de elementos tiene también una equivalencia con los de *Dublin Core*.

2.6.6 ISO/IEC 11179

ISO/IEC 11179 (ISO, 2017) establece las características de los metadatos para describir datos y la gestión y administración de estos metadatos en un registro de metadatos. Se aplica a la formulación de representaciones de datos, conceptos, significados y relaciones entre ellos para ser compartidos por personas y máquinas, independientemente de la organización que genere los datos. No es aplicable a la representación física de los datos como bits o bytes a nivel de máquinas.

Según DAMA (2009), este estándar es una guía importante para la industria de desarrollo de herramientas, pero es poco probable su utilización directa por las organizaciones, dado que estas suelen desarrollar utilizando herramientas comerciales. Sin embargo, fragmentos de algunas partes del ISO/IEC 11179 pueden ser útiles a las organizaciones que quieren desarrollar su propio estándar interno, puesto que contiene detalles significativos sobre todos los temas.

2.6.7 INTEROPS

“*Interopérabilité entre Organismes de la Protection Sociale*” (InterOPS, 2011) es un estándar de interoperabilidad informática que permite el establecimiento de un espacio de confianza entre los organismos de la esfera social francesa, a través de los tres siguientes modelos de intercambio:

- **InterOPS-A (entre aplicaciones):** intercambios, en protocolo de servicios web, efectuados bien en un contexto de aplicaciones sin identificación del usuario, o bien en el contexto de un usuario de un organismo cliente que da servicio a las aplicaciones de otros organismos mediante una aplicación local.

- **InterOPS-P (entre portales)**: acceso de un usuario cliente a una aplicación o a un servicio de otro organismo, vía los respectivos portales web de los organismos.
- **InterOPS-S (esfera de confianza)**: acceso de un usuario a un entorno de confianza compuesto por organismos que juegan el rol de operador de autenticación o de operador de servicios.

El estándar Interops se basa en diversos estándares reconocidos, como SOAP, WS-Security, SAML o XML *Signature*. Facilita principalmente los intercambios entre organismos de protección social y de estos con otros organismos públicos.

2.6.8 OMG (*Object Management Group*)

OMG (2017) es un consorcio sin ánimo de lucro formado por las industrias líderes del sector, dedicado a la definición, promoción y mantenimiento de los estándares de la industria para las aplicaciones interoperables de empresas. Compañías como IBM, Oracle, Unisys, NCR y otras, apoyan este consorcio. OMG es el creador del estándar de middleware CORBA y ha definido otros estándares de datos relacionados. De ellos, más de 30 están relacionados con los metadatos (OMG2, 2017). Merecen ser destacados:

- CWM (*Common Warehouse Metamodel*), que fue desarrollado entre el año 2000 y el 2008, con el objetivo de dar soporte a los metadatos de sistemas de *Data Warehousing* y *Business Intelligence*, permitiendo la interoperabilidad de los metadatos manejados en herramientas asociadas, tales como: DBMSs, OLAP, *Data Mining*, etc.
- IMM (*Information Management Metamodel*), es el modelo continuador de CWM, que también se propone dar soporte a todo tipo de sistema de información, además de DW y BI.
- ODM (*Ontology Definition Metamodel*), es una especificación para representación formal, gestión e interoperabilidad de objetos representados en lenguajes semánticos, tales como OWL y RDF.

2.7 Análisis crítico de la situación

2.7.1 La interoperabilidad

Como se ha visto anteriormente, la interoperabilidad se considera desde diferentes dimensiones que a su vez tienen sus interrelaciones. El hecho de que haya una dimensión legal, política y organizacional se traduce en que generalmente, los informes de situación o estado de la cuestión publicados sobre interoperabilidad se refieren a los ámbitos concretos en los que dicha interoperabilidad desea ser estudiada o aplicada.

En el caso de las Administraciones Públicas, un punto de referencia es el Observatorio de los Marcos Nacionales de Interoperabilidad (NIFO), establecido por la Comisión Europea en el contexto

de la aplicación del programa ISA para medir los progresos y supervisar la situación actual de la interoperabilidad en la UE. Dicho observatorio, ha puesto de manifiesto que, en 2014, la adaptación de los marcos nacionales de interoperabilidad al Marco Europeo de Interoperabilidad alcanzó una tasa media del 72 % (COM, 2017). No obstante, el nivel de aplicación de los marcos nacionales en proyectos específicos fue del 45 % en 2015, lo que muestra que **aún subsisten dificultades en la aplicación práctica de las recomendaciones actuales**. Como puntos de mejora se sugiere:

- Mejorar la gobernanza de sus actividades de interoperabilidad nacionales.
- Usar modelos operativos comunes para desarrollar mejores servicios públicos digitales e incluir las necesidades de ciudadanos y empresas de otros Estados miembros de la UE.
- Gestionar los datos de su propiedad en **formatos semánticos** y sintácticos comunes para facilitar su publicación en portales, así como los procedimientos para agregarlos, compartirlos y reutilizarlos.

De los casos estudiados a lo largo de este capítulo se desprende que, en la actualidad, las Administraciones Públicas gestionan grandes cantidades de datos en distintos formatos, con diferentes métodos de gestión de datos, alojando múltiples copias en muchos repositorios diferentes y divulgándolos, a menudo, en portales en distintos países, pero **sin ninguna armonización en términos de contenido y presentación**. Este hecho explica por qué se reutiliza la información existente sobre ciudadanos y empresas solamente en el 48 % de los casos, y por qué a menudo la prestación de servicios públicos a los ciudadanos y empresas es un proceso oneroso y que requiere mucho tiempo. También puede plantear cuestiones de confianza en relación con la protección de datos.

Por otra parte, el informe de la CEPAL (CEPAL, 2007) apuntaba una serie de temas en los que era necesario trabajar para conseguir resultados en el campo de la interoperabilidad, entre los que conviene destacar la **necesidad de estandarización**. Se señala la existencia de algunos incipientes proyectos internacionales para abordar esta temática. Sin embargo, estos esfuerzos, que pueden ser notables individualmente, se presentan como inconexos, ignorando paradójicamente algunos de los principios básicos de interoperabilidad.

2.7.2 Modelos de interoperabilidad

El modelo de madurez más genérico encontrado es LISI (1998). Sobre este modelo, el Instituto de Ingeniería de Software de la Universidad Carnegie Mellon manifiesta dos preocupaciones principales.

En primer lugar, LISI refleja un conjunto de estándares y expectativas de interoperabilidad vigentes en el ámbito del Departamento de Defensa de los EEUU en el momento de su creación. Por ese motivo, contiene ciertas suposiciones que reflejan un sesgo tecnológico, por lo que algunos de los aspectos recogidos en su modelo deberían ser actualizados para adaptarse a nuevas tecnologías.

En segundo lugar, puede suceder que dos sistemas con un nivel elevado de interoperabilidad no sean altamente interoperables, puesto que las diferencias en las cualidades del servicio, la finalidad

de uso de los sistemas, la manera en la que se emplean los datos, y otros factores diversos pueden hacer que dos sistemas no sean capaces de interoperar, incluso cuando los fundamentos técnicos de ambos sean idénticos.

Pero, además, hay que tener en cuenta que el modelo LISI se dirige a evaluar el grado de interoperabilidad alcanzado en una organización en sus distintas actividades o dominios, y su aplicación es más compleja para un ámbito más amplio que una única organización o grupo de organizaciones coordinadas. No obstante, sería posible asociar algunos niveles a entornos de cooperación entre múltiples organizaciones y distintos dominios. Así, el nivel 3 o de dominio sería extensible a un posible entorno global de Seguridad Social en el que pudiesen interoperar todos los sistemas que soporten distintas tareas y procesos en línea, empleados por cualquier organización o usuario individual y a nivel internacional.

Por otra parte, se podría asociar el nivel 4 o empresarial a un entorno que abarque organizaciones y sistemas independientes de múltiples ámbitos o dominios (*e-learning, e-Government, e-commerce*, etc.) que pretenden intercambiar y utilizar información de utilidad de forma continua o puntual. Sin embargo, en este caso se presentan diversas incompatibilidades del modelo, como el almacenamiento común de los datos y el uso de aplicaciones comunes. En este sentido, sería preferible mantener un entorno de sistemas distribuidos, pero estableciendo un modelo de datos común para el intercambio de información en el mismo dominio y entre dominios heterogéneos.

En cuanto al LCI del Ministerio de Defensa Australiano, tal como se ha dicho, extiende LISI hacia la dimensión organizativa, añadiendo cinco niveles que completan el modelo americano.

2.7.3 La interoperabilidad en los distintos escenarios de la Seguridad Social

En los casos recogidos en los programas de protección social y servicios sociales, las experiencias de interoperabilidad se han basado en:

- La implementación de recursos de información compartidos, especialmente sobre beneficiarios y programas sociales a los cuales tiene derecho.
- El acceso a datos de instituciones externas para cargar datos en el sistema de información compartido. Por ejemplo, usando servicios web.
- El acceso a los propios sistemas de información compartido a través de mecanismos programáticos, como, por ejemplo, servicios web.

En los escenarios tipo **sistemas integrados de servicios de salud**, se ha encontrado técnicas de interoperabilidad aplicadas a:

- La implementación de recursos de información compartidos, especialmente sobre beneficiarios y beneficios de salud a que tiene derecho.

- El acceso a sistema de información compartido por mecanismos programáticos, como el caso de los servicios web. En particular desde instituciones de servicios de salud que en algunos casos pueden ser privadas.
- La integración con otros sistemas que administran datos relevantes para el seguro de salud, en particular las relaciones de trabajo y vínculos familiares de los cotizantes al seguro.

En cuanto a las características principales de este tipo de escenario de **recaudación de cotizaciones** son:

- El intercambio de datos entre la institución recaudadora y los empleadores.
- El intercambio de datos, uso de datos compartidos e integración de funciones con otras instituciones de Seguridad Social, así como con instituciones de recaudación de impuestos.
- El intercambio de datos y acceso a servicios de instituciones bancarias que ofrecen servicios de pago. Las principales técnicas de interoperabilidad a utilizar son las de sistemas de información compartidos, intercambio de datos y acceso a servicios externos.

Las iniciativas de mejora de eficiencia institucional han apuntado a mejorar la efectividad y eficiencia de las plataformas informáticas de las instituciones de Seguridad Social, a través de sistematizar el reuso de servicios e información; apuntando también a facilitar la implementación ágil de nuevos programas y servicios al público. Con respecto a las técnicas de interoperabilidad, se observa la aplicación de estándares tales como XML, SOAP, SAML, el uso de servicios web para la integración de sistemas, y la aplicación de Arquitecturas Orientadas a Servicios como marco general.

De especial relevancia para esta tesis son las conclusiones relacionadas con los **intercambios de datos** para la puesta en práctica de los acuerdos internacionales. Se destaca la necesidad de contar con:

- Mecanismos de autenticación, como firma electrónica.
- Un registro de operaciones que aporte transparencia y posibilidad de seguimiento a las transacciones realizadas.
- Estándares para formateo de datos (XML) e interconexión de sistemas (servicios web) en sus versiones que permiten incluir firmas electrónicas y encriptar los datos (WS-Security).
- Herramientas que eviten **errores en el significado o interpretación** de la información.

De lo recogido en este capítulo sobre el estado del arte, se desprende la necesidad de desarrollar herramientas para solucionar los problemas relacionados con el significado (semántica) de los datos en este escenario.

2.7.4 Situación de las distintas dimensiones de interoperabilidad

2.7.4.1 Situación de la dimensión política y legal

En el capítulo 1 se ha pasado revista a los acuerdos internacionales que son el apoyo natural de esta dimensión. Son estos acuerdos los que establecen un marco de referencia y unas condiciones previas en las que luego se desarrollan los intercambios.

Sin duda, el más avanzado y ambicioso es el de la Unión Europea, si bien no se puede decir que sea en rigor un acuerdo sino una coordinación de legislaciones nacionales al amparo de unos reglamentos comunes. En cualquier caso, desde un punto de vista político y legal, la interoperabilidad cuenta con un respaldo importante.

El siguiente en importancia sería el establecido por el Convenio Iberoamericano de Seguridad Social. En este caso también hay un convenio internacional vigente que está proporcionando el marco formal que exige esta dimensión de la interoperabilidad (CIM, 2011).

El resto de los acuerdos multilaterales (CEI, CEDEAO, CIPRES, ASEAO, etc.) están todavía lejos de proponer intercambios de información regulares que necesiten de las técnicas de interoperabilidad, si bien en un próximo futuro va a darse esta circunstancia con toda seguridad.

En cuanto a los acuerdos bilaterales, entran en la categoría de acuerdos de buena voluntad, sin respaldo legal suficiente como para ser considerados un marco político o social con las garantías ideales.

2.7.4.2 Situación organizacional

Los sistemas de Seguridad Social se basan siempre en modelos de negocio que forzosamente tienen su equivalente, cualesquiera que sean los matices: afiliación, recaudación, vida laboral o prestaciones. Si lo que se pretende es facilitar los intercambios en un ámbito mundial, la creación de un instrumento similar al directorio de instituciones mencionado en la sección 2.3.3.2 parece imprescindible.

Por otra parte, sería necesario realizar un análisis de los campos actuales del único directorio existente en este momento (Repositorio de Instituciones de la Unión europea) y determinar si son apropiados en un contexto geográfico y político mayor, así como estudiar nuevos datos que fuera necesario tener en cuenta. Por ejemplo, un directorio de este tipo debería incluir la descripción semántica de los campos, así como la descripción de las particularidades por parte de los Estados que quieran utilizarlo.

Hay que tener en cuenta que, si se pretende afrontar cualquier tipo de intercambio, las diferencias de modelo y culturales pueden ser más profundas. Así, por ejemplo, en el caso de la UE los problemas sobre el juego de caracteres empleados pueden resolverse con reglas de transliteración. Eso no es posible si entran en juego lenguas como el árabe o el chino. Incluso en un ámbito geográfico cercano como podría ser España y Marruecos, estos problemas estarían presentes. Lo mismo puede decirse en unos eventuales intercambios entre Túnez y la UE, intercambios que

deben producirse como lógica consecuencia de los programas de colaboración con ese país que se están desarrollando.

Fuera del marco de la Unión Europea, el único caso de intercambios similares es el que se realiza en Mercosur en aplicación del Convenio Multilateral Iberoamericano de Seguridad Social (OISS, 2007). En el sistema de información que se ha desarrollado a estos efectos, no hay un equivalente al directorio de instituciones. Las relaciones se establecen entre países, utilizándose en cada uno de ellos un punto de referencia similar al organismo de enlace del caso europeo. A medio plazo, esto significa una limitación importante. El desarrollo de los sistemas de Seguridad Social en todo el mundo conlleva una complejidad en lo que a organizaciones se refiere de la que no escapa ningún esquema nacional. Eso forzosamente se traduce en los problemas ya citados en el caso de la Unión Europea con todos los matices propios de cada caso. En definitiva, una herramienta similar al directorio de instituciones va a ser necesaria para dar apoyo a estos intercambios.

No obstante, para que este instrumento sea útil habría que definir:

- La institución y organismo responsable de su mantenimiento.
- La información que debería contener en un ámbito mundial.
- El compromiso de los países que lo utilizaran.

2.7.4.3 Situación técnica

Entre las experiencias citadas, INTEROPS, *Banq Carrefour* o el ENI son buenos ejemplos de utilización de las tecnologías recomendadas para garantizar la interoperabilidad técnica. Así, por ejemplo, INTEROPS respeta normas técnicas internacionales como SOAP, SAML, XML o SSLV3.

Asimismo, la CEPAL propone que la interoperabilidad se defina, para las tecnologías de estructura y de semántica de informaciones y servicios, basándose en varios estándares existentes y en uso en los países. Se pide también que se posibilite la utilización de los países de la red en forma parcial sin que haya necesidad de contar con un nivel de desarrollo homogéneo de los sistemas de gobierno electrónico. Igualmente se propone desarrollar estándares que consideren la heterogeneidad de plataformas de software y hardware que existe en la región y que se asegure que cada país mantiene autonomía sobre su información. Con el soporte de firmas y certificados digitales de alta confiabilidad, acordados entre los países, se debe garantizar que los datos autorizados para cada operación específica sean transferidos para el uso de personas o entidades explícitamente autorizadas para ello y que las transacciones en línea surtan los efectos previstos en las leyes de cada país.

EESI utiliza mensajes XML con restricciones XSD para la comunicación entre las instituciones a través de los puntos de acceso. En la capa de transporte, se utiliza el protocolo ebMS/AS4 sobre HTTPS. El mismo protocolo AS4 es utilizado tanto para los mensajes de usuario como los de servicio.

En cuanto a MERCOSUR, tal como se ha apuntado en el capítulo 1, el intercambio de datos está codificado en el lenguaje XML (W3C XML), mientras que las comunicaciones se realizan utilizando servicios web (W3C WSDL).

En definitiva, la interoperabilidad cuenta con herramientas técnicamente sólidas y suficientemente divulgadas como para que no sea en ningún caso el mayor de los obstáculos. Más importante es el acuerdo de las tecnologías concretas a utilizar y la forma de ponerlas en práctica.

2.7.4.4 Situación semántica

En la casi totalidad de los casos analizados, el remitente y el receptor utilizan diferentes estructuras cuando intercambian datos, lo que da lugar a conflictos semánticos. Según Fraunhofer (2009), la mejor forma de evitarlos es el uso obligatorio de esquemas de datos estandarizados por todos los participantes. Sin embargo, en el caso de los intercambios internacionales, incluso los que tienen lugar en el interior de la Unión Europea, la autonomía de los Estados hace imposible forzarlos a utilizar un esquema común.

El proyecto EESSI, en pleno desarrollo como se ha dicho en el capítulo 1, es un claro ejemplo de los problemas semánticos que aparecen en los intercambios de datos. El autor tuvo oportunidad de analizarlos y presentar sus conclusiones a la Comisión Técnica para el Tratamiento de la Información en su reunión número 56, celebrada en Bruselas el día 27 de febrero de 2013. Dichas conclusiones se recogen en el documento AC (2012b), anexo a esta tesis y se mencionaran de nuevo en el capítulo 3.

A raíz de esta presentación y gracias al apoyo de varias delegaciones, especialmente las de Noruega y Finlandia, se formó un grupo de trabajo (grupo Ad Hoc en la terminología comunitaria) con el nombre de *Data Model Expert Group*, del que también es miembro el autor. Su objetivo principal (CASSTM, 2013) es apoyar todos los aspectos relacionados con el modelado de datos del proyecto, entre ellos, los aspectos semánticos.

En el momento de redactar esta tesis el grupo continúa trabajando en la preparación del modelo lógico de información, coordinando la labor de los grupos Ad Hoc que se encargan de la definición de los documentos electrónicos estructurados.

Es conveniente mencionar que uno de los mayores avances en la creación de un modelo de datos para los intercambios de Seguridad Social tuvo lugar con motivo del proyecto TESS, antecedente de EESSI, concretamente en el subproyecto conocido como Build 6, sin que se haya considerado conveniente aprovechar la labor realizada en su día.

Por otra parte, al ocuparse de los problemas semánticos, el Libro de la CEPAL destaca la necesidad de contar con soluciones en español, portugués e inglés, los idiomas que predominan en la región. También aboga por que se implementen repositorios de metadatos globales, resultantes de las aplicaciones en la medida en que ellas se vayan implementando. De esta forma, cuando se vayan sumando países y soluciones interoperables, se podrán reutilizar las especificaciones semánticas anteriores ampliándolas solo cuando por particularidades específicas se requiera. Se evitará así la duplicación de esfuerzos y se ampliará el alcance de la interoperabilidad.

Herramientas para mejorar la interoperabilidad semántica

Las referencias encontradas a instrumentos de interoperabilidad semántica concretos (Criado et al., 2010) ponen de manifiesto la complejidad que presentan los proyectos en este ámbito. No sólo se trata de construir los elementos y herramientas sino, también, de unir sintaxis y semántica, de determinar sus ámbitos de uso y de dejar constancia de la necesidad de adoptar medidas que logren consensuada y coordinadamente uniformizar los significados, compatibilizar los distintos elementos, tratar y reutilizar la información y vincular los datos web para hacer posible el intercambio de datos con el mismo significado.

Fraunhofer (2009) hace una clasificación y descripción de los conflictos semánticos que a su vez se toma como base para detectar y resolver estos conflictos. Algunos de ellos, como los de representación, son fáciles de resolver, mientras que otros presentan una alta complejidad, como son todos aquellos en los que diferentes estructuras o esquemas de datos están involucrados. La herramienta que más a menudo se cita para resolver los distintos tipos de conflictos, son los vocabularios controlados, especialmente para los conflictos relacionados con el significado de los datos.

Por otra parte, el libro blanco de la CEPAL (CEPAL, 2007), al ocuparse de problemas semánticos, aboga por que se implementen repositorios de metadatos globales. Asimismo, para alcanzar la interoperabilidad regional, proponer realizar una serie de acciones entre las que conviene destacar las siguientes:

- Definir la metodología de adopción de estándares
- Utilizar un vocabulario común, preciso e inequívoco.

2.7.5 Análisis de vocabularios existentes

Dado que existe consenso en cuanto a la importancia de los vocabularios controlados para afrontar los problemas de interoperabilidad semántica, se ha considerado conveniente analizar los artefactos existentes en el ámbito de la Seguridad Social o con posibilidad de ser utilizados en este ámbito.

2.7.5.1 Build 6 del proyecto TESS

Como se ha dicho anteriormente, en su día se confeccionó un modelo de datos y un glosario de términos para los intercambios de datos entre instituciones de Seguridad Social en el seno de la Unión Europea.

En los estudios previos para el desarrollo del proyecto EESSI, se contempló la posibilidad de que este trabajo pudiera ser reutilizado en el desarrollo de los SED, pero finalmente se decidió abordar su diseño desde cero. En la reunión de la Comisión Técnica celebrada en junio de 2017, el autor de esta tesis, en su calidad de representante oficial de España, reclamó una vez más unos artefactos similares, a lo que el equipo de desarrolladores contestó que ya se estaba manejando de forma

interna y más adelante se podría a disposición de las instituciones. Se proporcionó una versión provisional en octubre de 2018 que va a sufrir modificaciones en los meses siguientes.

2.7.5.2 Core Vocabularies

En el marco de esta tesis, el 27-11-2013, la DIGIT de la Comisión Europea, a petición del autor de esta tesis, en su calidad de miembro de la delegación española en la Comisión Técnica para el procesamiento de datos, realizó una presentación de los estándares semánticos desarrollados hasta el momento, con especial atención a los *Core Vocabularies*. Según concluyeron los expertos de la Comisión Europea en su exposición, es necesario promover los estándares semánticos y la definición de esquemas de metadatos es el camino indicado para hacerlo.

Por otra parte, según EU (2015), estos vocabularios pueden ser utilizados para reducir los problemas semánticos de dos maneras:

- Diseñando nuevos modelos de datos a modo de extensión de los *Core Vocabularies*.
- Mapeando los modelos de datos existentes con los *Core Vocabularies*.

Ninguna de las dos cosas se ha realizado hasta el momento.

2.7.6 Análisis de los estándares de metadatos

ISA (2011) cita buenas prácticas en este terreno por parte de Dinamarca y Finlandia a nivel nacional, en los que la gestión de metadatos juega un papel relevante en la arquitectura de información.

Otra importante línea de trabajo está relacionada con los *Core Vocabularies* y su intención de que sean utilizados como *Core Metadata* en todos los Estados miembros. Sin embargo, tal como se ha explicado a lo largo de este capítulo, esta circunstancia ni se ha producido ni está prevista en un futuro inmediato en el ámbito de la Seguridad Social. Más concretamente, se menciona que las administraciones públicas deben ser conscientes de que los metadatos son un importante activo para el desarrollo de sistemas de *e-Government*. Insiste, asimismo, en la necesidad de identificar y reaprovechar los conjuntos de metadatos existentes.

También el movimiento "*Open Data*", de ámbito mundial, ha lanzado propuestas que intentan fomentar la utilización de metadatos para un gobierno abierto, pero con nulo seguimiento hasta el momento en el ámbito que nos ocupa (intercambios de datos de Seguridad Social).

Varios de los estándares existentes analizados se basan en *Dublin Core*, el más extendido con diferencia en todos los ámbitos. Es el caso de AGLS e IEEE/LOM.

AGLS ha sido desarrollado por la administración pública australiana y se basa en *Dublin Core* con extensiones, pero está orientado a toda clase de recursos y no a datos.

Tanto éstos como los restantes, utilizan un concepto amplio de metadatos, confirmando que el término se refiere a información sobre la información más que a información sobre datos

estrictamente. Así, por ejemplo, el ENI de España trata de identificar las propiedades de los documentos, mientras que IEEE/LOM se refiere a los objetos de aprendizaje.

Sin embargo, si observamos las carencias recogidas en la interpretación de la información (interoperabilidad semántica), veremos que para evitar cualquier ambigüedad sería necesaria una definición inequívoca del significado y propiedades de cada uno de los datos individualmente.

En ese aspecto, se ha detectado de forma habitual una confusión entre la web semántica y su orientación a búsquedas inteligentes, y las propiedades semánticas de los datos que tienen que ser tratados por sistemas heterogéneos.

También se ha detectado carencias importantes en los ámbitos que cubren estos estándares. A menudo se confunde el ámbito de la salud, en el que la principal referencia es HL7, con el contexto de la Seguridad Social, ampliamente descrito en el capítulo 1, y los diferentes escenarios en los que se producen los intercambios de información entre instituciones. En la práctica, HL7 es un estándar de mensajería para el intercambio electrónico y orientado a temas médicos, concretamente historias clínicas.

En este ámbito concreto, no se han encontrado referencias sólidas para solucionar los problemas de interoperabilidad semántica, en particular, ningún estándar de metadato que sea directamente aplicable.

De los otros estándares mencionados puede decirse:

- ISO/IEC/ 11170 está orientado al desarrollo de herramientas y se utiliza muy poco por las organizaciones.
- INTEROPS facilita los intercambios entre organismos de protección social, pero se basa principalmente en los aspectos técnicos.
- OMG están enfocados al desarrollo de herramientas para el manejo de grandes bases de datos.

2.7.7 Situación de las principales experiencias de intercambios de datos de Seguridad Social en lo que a interoperabilidad se refiere

Si tomamos como referencia el modelo LSI, no se han encontrado experiencias que puedan considerarse en un nivel 3 o superior.

En el nivel 2 (funciones comunes mínimas, datos y aplicaciones separadas) podrían considerarse EESSI y MERCOSUR. Entre las funciones comunes mínimas estarían las proporcionadas por herramientas como la interfaz web que proporciona EESSI (llamada RINA) y SIBERO. En el caso de EESSI, existe un modelo de datos y el proyecto tiene capacidad de intercambiar datos estructurados, atributo característico del nivel 2. Las estructuras de datos avanzadas sólo son manejadas a nivel local por las aplicaciones propias.

En el nivel 1 (conectado) podrían considerarse los convenios bilaterales que han incluido un intercambio electrónico de datos. La conexión se realiza normalmente vía Internet o aplicaciones

propietarias (FTP securizado o similar). Este último sistema permite intercambio de ficheros de forma simple, si bien de forma muy básica. Es decir, los procedimientos se refieren únicamente a políticas locales y los datos responden a estructuras locales y a lo sumo se acuerda un fichero de intercambio.

En el siguiente capítulo de esta tesis, se presentará una propuesta para dar solución a los problemas de interoperabilidad semántica detectados en los intercambios de datos internacionales entre instituciones de Seguridad Social.

La propuesta se basa en la definición de un conjunto de metadatos para garantizar la correcta interpretación de todos y cada uno de los datos intercambiados, así como definir una especificación que en el futuro pueda convertirse en un estándar internacional.

Capítulo 3. Propuesta

3.1 Razones para una nueva especificación: recapitulación del problema

En los capítulos anteriores se ha puesto de manifiesto la existencia de problemas semánticos en los intercambios de datos de Seguridad Social, así como la falta de soluciones sólidas para afrontar estos problemas.

En el presente capítulo se va a proponer una solución en línea con lo estudiado en la literatura especializada recogida en el capítulo 2. Se trata de una propuesta de especificación de conjunto de metadatos para el intercambio internacional de datos en el ámbito de la Seguridad Social, tal como se ha hecho en otros ámbitos. La propuesta se ha concebido de tal forma que en su día pueda llegar a convertirse en un estándar ISO.

3.1.1 Alcance del problema

El capítulo uno se ha dedicado a presentar el contexto en el que se desarrolla el problema abordado por esta tesis. En datos de Naciones Unidas (UN, 2015), se mencionaba la existencia de 243 millones de emigrantes en el mundo. En uno u otro sentido, prácticamente todos los países se ven afectados por el movimiento de personas, mucho más si tenemos en cuenta los desplazamientos temporales y su repercusión en la prestación de asistencia sanitaria para nacionales de otros países y los posteriores gastos, facturas y reembolsos, que se generan.

La necesidad de garantizar derechos en materia de Seguridad Social da lugar a los acuerdos internacionales, bilaterales o multilaterales, tal como se ha expuesto en el capítulo 1 (sección 1.1.5). La puesta en marcha de estos acuerdos requiere la implementación de una serie de intercambios de datos entre sistemas de información generalmente heterogéneos. Según AISS (2014), más de mil millones de personas se pueden ver afectadas por estos flujos migratorios, con el consiguiente impacto económico en las economías locales e internacionales.

Según las mismas fuentes citadas, la tendencia actual es a un mayor movimiento de la población, lo que se traducirá en un mayor número de acuerdos y en una necesidad mayor de intercambios de datos.

En definitiva, la envergadura e impacto social del problema tratado justifica el desarrollo de una solución que pueda ser aplicada por cualquier institución de Seguridad Social, con el fin de mejorar los resultados obtenidos hasta el momento por los intercambios de datos en esta materia.

3.1.2 Automatización de los intercambios

El número de instituciones en cada país varía en función de su legislación interna y organización. A título de ejemplo, en el ámbito de los países del Espacio Económico Europeo, se habla de más de 15000 instituciones (EESSI, 2012) sin que las competencias de estas organizaciones sean tampoco homogéneas. Si esto se extiende a los 193 países reconocidos por la ONU en 2017, el número de relaciones diferentes que cada institución tiene que gestionar es imposible de abarcar sin un nivel mínimo de informatización. Sin embargo, tal como se ha visto en la sección 1.2, la automatización

de estos intercambios es todavía muy limitada. Las experiencias existentes ponen de manifiesto la necesidad de aplicar técnicas de interoperabilidad. Los trabajos de la AISS, tanto los informes técnicos como el desarrollo de directrices, así lo ponen en evidencia.

Los intentos de aplicar estas técnicas datan de 1992 y siguen en desarrollo. A lo largo del capítulo 2 se ha estudiado la situación en las distintas dimensiones de interoperabilidad, llegándose a la conclusión de que es necesario buscar soluciones especialmente en el campo semántico.

Efectivamente, los intercambios de datos de Seguridad Social tienen algunas particularidades que hacen que una definición sin ambigüedades de los componentes centrales utilizados en las instituciones tenga un impacto clave en el éxito y calidad de la interconexión de sistemas. Estos problemas semánticos de los intercambios de datos de Seguridad Social no se han resuelto hasta el momento.

Esta circunstancia se pone de manifiesto cuando, para acordar la prestación de servicios, las instituciones competentes tienen que ponerse de acuerdo en multitud de especificaciones, con la descripción de personas, productos, procesos, formatos, etc. Asimismo, es necesario consensuar cómo formular esas descripciones y dónde almacenarlas para su uso público.

Si hacemos referencia a la experiencia del proyecto EESSI, la nota presentada en 2012 por las delegaciones de Austria, Alemania, Países Bajos y Suiza, sobre mejoras necesarias para la implementación de EESSI en el ámbito nacional (AC, 2012), muestra los problemas relacionados con los documentos electrónicos a intercambiar entre instituciones (SED), la definición de esquemas XML y las validaciones de los mensajes. Todos ellos están directamente relacionados con dificultades asociadas a la interoperabilidad semántica.

Como continuación a esta nota, la delegación española presentó un análisis de la situación elaborada por el autor, en el que se indicaba que los principales problemas encontrados eran de tres tipos (AC, 2012b):

Polisemia: un mismo término tiene varios significados.

- Sinonimia: dos términos con nombres distintos tienen el mismo significado.
- Multilingüismo: no puede garantizarse que el nombre del campo se corresponda con el significado esperado.

Según Fraunhofer (2009) hay una serie de técnicas que son eficaces para paliar estos problemas. Una de las más utilizadas es la de coincidencia de esquemas, según se ha dicho en el capítulo 2, sección 2.3.3.3, especialmente útil cuando se intercambia información entre esquemas heterogéneos. En ese mismo capítulo 2 se han recogido todos los posibles problemas y alternativas de solución, así como un análisis crítico de la situación que nos lleva proponer la solución desarrollada en la sección 3.2.

3.1.3 Concepto de “dato”

En la sección 2.5.1, se mencionaban los distintos conceptos de metadatos que se puede encontrar tanto en la literatura como en las aplicaciones prácticas. En esta sección se recogía la definición de DAMA (2009), según la cual metadatos son “*datos sobre datos*”, si bien en la misma publicación se dice que su interpretación puede depender del entorno en que se utilice.

Dado que esta tesis va utilizar el concepto original de metadatos como datos que definen datos, conviene precisar cuál es el concepto de “dato” que se está utilizando.

Según la Real Academia Española (RAE, 2018), la palabra “dato” procede del latín *datum* (lo que se da) y se le atribuyen los siguientes significados:

- Información sobre algo concreto que permite su conocimiento exacto o sirve para deducir las consecuencias derivadas de un hecho.
- Documento, testimonio, fundamento.
- Información dispuesta de manera adecuada para su tratamiento por una computadora.

Una referencia más técnica nos la proporcionan Löbel et al. (1982), definiendo “dato” como “*el conjunto de signos numéricos o alfanuméricos que proporcionan información sobre cosas o acontecimientos*”.

Más recientemente, el ENI (2007), en su Guía de aplicación, define dato como “*una representación numérica, acústica o de cualquier otro tipo concerniente a persona físicas identificadas o identificables*”.

Si nos situamos en el contexto de la información a intercambiar en el ámbito de esta tesis, los datos deberán ser parte de los documentos acordados por las instituciones de Seguridad Social. Un ejemplo podría uno de los formularios definidos al amparo de los Reglamentos Comunitarios (OJEU, 2004) como el que se puede ver en la figura 19.

El contenido de cada uno de los campos o conjunto de campos de ese formulario que hacen alusión a un concepto identificado y definido en el GOT (ver sección 2.3.3.3), es lo que se entendemos como “dato”. Normalmente, debe corresponderse con los campos de un fichero tradicional o de una base de datos.

P1000

La Comisión Administrativa
de Coordinación
de los Sistemas de Seguridad Social



Periodos de educación de los hijos

Art. 44 del Regl. 987/2009

Numero de anexos:	[integer].....
Fecha de envío:	[DD/MM/YYYY].....

Institución de origen:	
Código del país*	[list ISO3166-1-alpha-2 code].....
Código de la institución*	[25].....
Nombre de la institución*	[155].....
Calle	[155].....
Localidad	[65].....
Código postal	[25].....
Provincia	[65].....
País	[list ISO3166-1-alpha-2 code].....
Teléfono	[65].....
Fax	[65].....
Correo electrónico	[255].....
Institución destinataria:	
Código del país*	[list ISO3166-1-alpha-2 code].....
Código de la institución*	[25].....
Nombre de la institución*	[155].....
Calle	[155].....
Localidad	[65].....
Código postal	[25].....
Provincia	[65].....
País	[list ISO3166-1-alpha-2 code].....
Teléfono	[65].....
Fax	[65].....
Correo electrónico	[255].....

Figura 19. Fragmento del formulario P1000 definido al amparo de lo previsto en el Reglamento 883/2004 (OJEU, 2004).

3.2 Desarrollo de la solución

3.2.1 Solución propuesta

En la propuesta inicial de tesis doctoral se sugería la creación de una ontología como solución más ambiciosa. Sin embargo, de la literatura estudiada y reflejada en el capítulo 2 se ha deducido que la complejidad de las ontologías se suele transformar en poca utilidad práctica. Por otra parte, la relación entre los términos que se maneja en el ámbito de la Seguridad Social no es lo suficientemente compleja como para justificar el uso de una ontología.

En el capítulo dedicado al estado del arte se ha visto que el estudio más profundo de la interoperabilidad semántica ha sido realizado por la Comisión Europea, en su intento de establecer una visión general del problema que permita plantear acciones concretas para mejorar los servicios públicos. Dentro de sus propuestas, la Estrategia Europea de Interoperabilidad (EIS, 2012), recomienda la utilización de metadatos como solución de compromiso.

En la misma línea, Criado et al. (2010) señalan *“como herramientas clave en la construcción de estos sistemas y, por tanto, en la referencia y organización del conocimiento, la constituyen los metadatos, datos que describen a otros datos y dan información sobre los datos contenidos o recursos, y los lenguajes controlados, listas de términos que, entre otros aspectos, reducen las*

ambigüedades semánticas, incrementan la consistencia en la representación y facilitan las búsquedas”.

Describir un recurso con metadatos permite que pueden ser interpretados tanto por humanos como por máquinas (NISO, 2004), tal como tratan de promover las técnicas de interoperabilidad.

Por otra parte, según Franhauser (2009), la forma más adecuada de evitar conflictos semánticos es el desarrollo de un estándar que sea aceptado y utilizado por todas las partes involucradas en un intercambio de datos.

Los beneficios esperados de la utilización de estándares de metadatos son fácilmente extrapolables a una definición de todos los datos implicados en un intercambio electrónico. A saber, y tal como se menciona en la nota presentada por la delegación española, citada anteriormente en la sección 2.7 e incluida como anexo (AC, 2013):

- Asegurar un procesamiento más rápido de los datos intercambiados.
- Proporcionar a los funcionarios de las instituciones remitentes unos datos de identificación y de otros tipos que las instituciones receptoras pueden obtener de sus clientes.
- Hacer posible verificar la información introducida en lo que a su estructura se refiere.
- Hacer posible, después de recibir un mensaje electrónico estructurado, identificar a las personas o cualquier información asociada con ellas y el procesamiento de la información.

Tal como se publicaba también en este documento, haciendo una vez más referencia a la experiencia del proyecto EESSI, es esencial que un mismo dato tenga la misma definición en todos los documentos a intercambiar, flujos y sectores, independientemente de la utilización que se vaya a hacer de él o del grupo que haya trabajado en su descripción. De ahí surge la necesidad de establecer los correspondientes metadatos, publicarlos y mantenerlos de forma accesible tanto para su utilización tanto por las aplicaciones como para su consulta por los funcionarios que van a tratar los documentos desde el punto de vista administrativo.

En cuanto a la puesta en práctica de una solución de este tipo, la forma más habitual es el desarrollo de una herramienta que permita la creación y mantenimiento de metadatos desde un entorno ofimático, para la posterior generación de los correspondientes formatos (como RDF/XML, JSON-LD, Turtle RDF u otros). En definitiva, sería necesario completar la descripción realizada con una arquitectura que incluya una capa de definición de datos accesible por el usuario final y con especial atención a los problemas semánticos que pudieran producirse en los intercambios. Sobre esto se hablará extensamente en el capítulo 4.

3.2.2 Elección de un estándar

Una especificación de metadatos para un dominio concreto debería diseñarse basándose en un estándar de forma que se satisfagan los requerimientos del dominio para el que se define (Baek, 2011). En primer lugar, se deberán elegir los elementos de los estándares que satisfagan el

propósito buscado, pudiendo ser más de uno. La selección apropiada de las descripciones es la clave para mejorar la interoperabilidad. También es crucial la capacidad de referenciar los distintos vocabularios o conjuntos de metadatos entre sí.

Según Harpring (2010), cuando desde una institución se desea poner información a disposición de otras instituciones, deben ser tenidos en cuenta una serie de aspectos:

- Qué elementos es importante compartir.
- Identificar la audiencia de la información compartida.
- Utilizar estándares técnicos para el intercambio de datos entre sistemas, como *Dublin Core (DC)*, *CDWA Lite* o *Visual Resources Association Core Categories (VRA Core)*.
- Acordar directrices y reglas para el contenido de datos, como CCO y CDWA.
- Acordar vocabularios controlados para asegurar la consistencia y coordinación de valores de datos.

Los “*bluiding blocks*” de cualquier esquema de metadatos, son los elementos que serán usados para definir la información proporcionada (Coyle, 2012). Generalmente son llamados “*data elements*”, si bien en la terminología de la web semántica suele llamárseles “propiedades”. En esta tesis se ha elegido la denominación de “elementos”.

En el capítulo 2 han sido revisados distintos estándares de metadatos, llegándose a la conclusión de que el más ampliamente utilizado es DC. Según Peristeras et al. (2009) es probablemente el conjunto de metadatos más ampliamente extendido y recomendado. Reforzando este argumento, varios Estados miembros de la Unión Europea o bien han establecido DC como estándar de metadatos en la Administración Pública o bien han desarrollado sus propios conjuntos de metadatos basándose en DC.

Noh (2011) recoge cinco referencias más en las que DC ha sido la base para el desarrollo de especificaciones para diversos ámbitos.

El “*Dublin Core Metadata Element Set*” surge de las discusiones que tienen lugar en el workshop organizado en 1995 por la OCLC en el *National Center for Supercomputing Applications (NCSA)*. Dicho *workshop* tuvo lugar en Dublin, Ohio, lo que dio nombre a este conjunto de elementos. Su objetivo original fue la catalogación de recursos en la web mediante unos pocos elementos y algunas reglas simples. Los 13 elementos elegidos, pronto pasaron a ser 15 y, en su nombre original en inglés, fueron:

- *Title*
- *Creator*
- *Subject*

- *Description*
- *Publisher*
- *Contributor*
- *Date*
- *Type*
- *Format*
- *Identifier*
- *Source*
- *Language*
- *Relation*
- *Coverage*
- *Rights*

Dublin Core ha sido utilizado con todo tipo de artefactos y aplicaciones que exigen cierta complejidad. Históricamente, ha habido cierta tensión entre los partidarios de una visión minimalista, que defendían la conveniencia de mantener un mínimo de elementos y unas semántica y sintaxis sencilla; y los partidarios de una visión estructurada, que proponen unas distinciones semánticas más finas y más extensiones para comunidades concretas.

Esto ha llevado a la distinción entre versiones calificadas y no calificadas de DC, según se use un mayor nivel de refinamiento para representar los posibles valores de los elementos.

Autores como Coyle (2012) se han referido a DC como “*la madre de todos los metadatos*”, especialmente por su papel clave para definir todo tipo de recursos. Su espacio de nombre ocupa el segundo lugar entre los más citados en la web semántica. DC es utilizado en aplicaciones tan comunes como las licencias *Creative Commons* o el proyecto *MusicBrainz* (2018). Como DC no está restringido a un determinado formato de registro, es habitual encontrar alguno de sus elementos en la mayor parte de conjunto de metadatos.

En principio, todos los elementos de DC son opcionales y todos son susceptibles de repetición. Los elementos pueden ser presentados en cualquier orden. Asimismo, la descripción de DC recomienda el uso de valores controlados para los campos donde sea apropiado, como es el caso de los vocabularios controlados para el campo *subject*, si bien no se considera obligatorio. Mientras que DC deja las reglas de contenido a las implementaciones particulares, la “*Dublin Core Metadata*

Inicitive” (DCMI) fomenta la adopción de reglas para dominios concretos, como podrían ser el caso de educación y gobierno. DCMI también mantiene una especificación denominada *DCMI Metadata Terms*, que toma como base los 15 elementos de DC y los refina, mediante la definición de lo que denomina “términos”, que se refiere a nuevos elementos compatibles con los de DC y a esquemas de tipos de valores que podrían asignarse a dichos elementos (DCMI, 2018).

Gracias a su simplicidad, el conjunto de elementos de DC es utilizado por cientos de proyectos en todo el mundo, en los más diversos dominios. Tanto es así, que la DCMI se ha expandido más allá del simple mantenimiento del conjunto de elementos de DC, hasta llegar a ser una organización que se describe a sí misma como “*dedicada a promover las más amplia adopción de estándares de metadatos interoperables y desarrollar vocabularios de metadatos especializados para sistemas de descubrimiento*” (DCMI, 2018).

A pesar de que muchos de los esquemas de metadatos son de desarrollo reciente, la mayoría han estado sujetos a cambios para su implementación en situaciones del mundo real. Esas modificaciones son fundamentalmente de dos tipos: extensiones y perfiles.

Una extensión consiste en añadir un elemento al esquema desarrollado para facilitar la descripción de un recurso de información, de un tipo en particular, de un tema determinado o bien para satisfacer las necesidades de un grupo de interés. Las extensiones incrementan el número de elementos.

Un perfil es un subconjunto de un esquema que se implementa para un grupo de interés determinado. Un perfil puede limitar el número de elementos que van a ser usados, refinar las definiciones de elementos para describir tipos concretos de recursos con más exactitud y especificar valores que un elemento puede tomar.

Un ejemplo puede ser el perfil desarrollado por la “*National Biological Information Infrastructure*”. Esta organización ha desarrollado un perfil que define y extiende el conjunto de datos para describir los datos biológicos (NBII, 2018).

Otro ejemplo puede ser la extensión realizada por el “*U.S. Department of Education’s Gateway to Educational Materials*”, que ha desarrollado su propio esquema de metadatos basándose en DC con algunas particularidades (GEM, 2018). En este caso, algunos elementos no son utilizados, como es el caso de *Contributor*, otros elementos son considerados obligatorios, y se añaden nuevos elementos (*Audience, Grade, Quality y Standards*). De esta forma se extiende el conjunto de DC para uso educacional.

De forma similar, se ha desarrollado el estándar ECEM para correo electrónico (Manouselis, 2006). También se basa en *Dublin Core*, tomando once de sus elementos y añadiendo extensiones para el dominio concreto en el que va a ser utilizado.

DC es también la base del estándar ISO/IEC 19788 MLR (Pons et al, 2011.) para describir los recursos educativos. Toma todos sus elementos, si bien acota sus especificaciones. Añade la posibilidad de definir tres nuevas características para estos elementos: presencia (obligatoriedad), repetitividad y ordenación (en caso de que haya repetición). Incluye también una lista de restricciones que afectan a parte de los elementos.

3.2.3 Especificación propuesta: ESSIM

En la sección 2.5 se ha analizado el concepto de metadatos y sus distintas interpretaciones según el propósito de su definición y el contexto en el que se va a utilizar. En el caso de la especificación que se va a proponer en este apartado, el fin principal es evitar cualquier posible ambigüedad en el significado de los datos intercambiados entre instituciones de Seguridad Social. En consecuencia, el concepto utilizado va a ser el de “datos que describen un dato determinado”. Ahí reside la principal novedad de la propuesta, frente a la tendencia a utilizar los metadatos para describir documentos o artefactos de forma general, pero sin controlar el contenido de cada dato en particular. En ESSIM se utilizan los metadatos para describir cada uno de los datos incluidos en los documentos, aplicando el concepto de dato descrito en la sección 3.1.3.

La especificación de metadatos propuesta se basa en los elementos de *Dublin Core* y, tal como se ha apuntado, se centra en los sistemas de información para la Seguridad Social (Delgado, 2013b). El nombre elegido es el de “**Exchange of Social Security Information Metadata**” (ESSIM).

Esta especificación se basa en once de los quince elementos de *Dublin Core* (tabla 10). Se han refinado dichos elementos con el fin de ajustarse más al uso que se pretende hacer en ESSIM. En el caso de los elementos *date* y *relation*, el refinamiento ha originado varios elementos ESSIM basados en ellos, siguiendo un procedimiento similar al que ha aplicado la *Dublin Core Metadata Initiative* (DCMI), al definir términos de DC más específicos basados en los elementos originales de DC, como es el caso de nuevos términos derivados también de *date* y *relation* (DCMI, 2018).

En ESSIM los elementos que han derivado de “*date*” son *dateCreated* y *dateModified*, y los basados en “*relation*” son *relation*, *mandate* y *references*. En este último caso coincide que también DCMI ha creado un nuevo término DC denominado “*references*” como refinamiento de elemento *relation* de DC, por lo que a su vez se podría considerar que el propuesto por ESSIM también es un refinamiento del término *references* de DC.

En la tabla 10 se muestra el conjunto de catorce elementos o metadatos propuestos bajo la denominación ESSIM para describir los datos incluidos en los intercambios de documentos entre instituciones de Seguridad Social. La columna “ESSIM” hace referencia al nombre de cada uno de los elementos propuestos.

La columna “*Obligation*” contiene la consideración del elemento como obligatorio (*Mandatory*), opcional (*Optional*) o condicionado (*Conditional*), para el caso de que sólo deba existir cuando se den ciertas condiciones.

La columna “*Multiplicity*” indica si un elemento puede tener varios valores, utilizando la notación *min..max*, que informa del mínimo y del máximo número de valores que puede tener. Así, por ejemplo, cuando aparece 0..1, significa que puede no tener ningún valor asociado o como máximo un valor. El símbolo “*” indica que no hay límite de número de valores que pueden asociarse a ese elemento. En la columna “*Based on DC*” se indican los elementos de *Dublin Core* en los que basan los elementos propuestos para ESSIM, como refinamiento del que aparece en cada caso.

La columna “*Value range*” contiene el tipo o vocabulario estándar reconocido al que deben pertenecer los valores del elemento, que se explican en la sección 3.2.5. Excepto en el caso de los elementos que admiten valores de tipo “*Literal*”, como pueden ser fechas o textos, los valores que

se pueden asignar al resto de elementos deben encontrarse en vocabularios controlados, con objeto, como indica Craig (2001), de estandarizar el contenido de los metadatos. La Unión Europea ya ha desarrollado algunas herramientas con este propósito, como es el caso del Glosario de Términos (ver sección 2.3.3.3), del Repositorio de Instituciones (EC, 2018) y de los *Core Vocabularies* (EU, 2015). Sería deseable la extensión y aplicación de estas herramientas en un ámbito mundial, si bien el desarrollo de esa propuesta se considera fuera del alcance de esta tesis.

ESSIM (element name)	Obligation	Value range	Multiplicity (min..max)	Based on DC element
creatorInstitution	Mandatory	Institution Repository (Institution)	1..1	creator
Dataset	Optional	Core Vocabulary (Mapping)	0..1	subject
dataTerm	Mandatory	Glossary of Terms (Term)	1..1	title
dateCreated	Mandatory	Literal	1..1	date
dateModified	Optional	Literal	0..1	date
Description	Mandatory	Literal	1..*	description
Format	Mandatory	Literal	1..1	format
Identifier	Optional	Literal	0..1	identifier
Language	Conditional	ISO 639	0..1	language
Mandate	Optional	EUR-Lex (Regulation)	0..*	relation
References	Optional	Standard Documents (Document)	0..*	relation
relationWithTerms	Optional	- relationTerm: Glossary of Terms (Term) - relationDescription: Literal	0..*	relation
senderInstitution	Optional	Institution Repository (Institution)	0..1	publisher
socialsecurityCategory	Mandatory	Institution Repository (Category)	1..1	coverage

Tabla 10. Elementos propuestos para ESSIM

3.2.4 Elementos de ESSIM

Como se ha dicho en la sección 2.7, las dos grandes aportaciones en el campo de la interoperabilidad semántica en el contexto de los intercambios de datos de Seguridad Social han sido el Glosario de Términos de Build 6 (proyecto TESS) y el *Institution Repository* del proyecto EESSI. A este último se le dedicó el artículo publicado en el Profesional de la Información (Delgado et al., 2012).

Por otra parte, en 2012 la Comisión Técnica de Tecnologías de la Información y las Comunicaciones de la AISS publicó un amplio informe sobre la interoperabilidad en contexto anteriormente indicado (ISSA, 2012), elaborado por un grupo de trabajo del que formaban parte representantes de distintas instituciones de Seguridad Social (entre ellos el autor), de la AISS y de la Universidad de Alcalá. Dicho informe fue presentado en la Conferencia Internacional sobre Tecnologías de la Información y las Comunicaciones que celebró en Brasilia.

Como continuación del trabajo desarrollado, en el congreso ICEIS 2013 celebrado en de Angers (Francia) se presentó un modelo de interoperabilidad específico para los intercambios de datos de Seguridad Social (Delgado et al., 2013).

Ese mismo año, en los congresos ISD2013 (celebrado en Sevilla) y ATICA2013 (celebrado en Huancayo, Perú), el modelo de datos se completó con una propuesta de conjunto de metadatos, basados en *Dublin Core*, con el propósito de solucionar los problemas de interoperabilidad semántica que se habían detectado en los intercambios de datos de Seguridad Social (Delgado et al., 2013b; Delgado, 2013d).

En el congreso ECEG 2014 (*14th European Conference on e-Government*), celebrado en Brasov (Rumanía), se presenta una versión más evolucionada del conjunto de metadatos (Delgado et al., 2014). Ese mismo año, en ATICA 2014, celebrado en Alcalá de Henares, se presentaron los esquemas de servicios web para generar y verificar el conjunto de metadatos (Viñan et al., 2014). De estos servicios web se tratará en el capítulo 4. La versión recogida en esta tesis se presentó en ATICA 2018 (Delgado, 2018).

En paralelo y dentro de la actividad profesional del autor, las cuestiones semánticas se han debatido repetidamente en la Comisión Técnica (UE), como se puede ver en los documentos referenciados en la sección 1.2, llegándose a la creación a finales de 2013 de un grupo especialmente dedicado al modelado de datos (DMEG) en el que se tratan las cuestiones semánticas.

Como resultado de lo tratado en los distintos foros antes mencionados, se ha elaborado el conjunto de metadatos que constituyen ESSIM. A lo largo de las distintas publicaciones, se observa una evolución del conjunto de metadatos y sus propiedades, hasta llegar a la propuesta que se desarrolla en los apartados siguientes. Dicha propuesta tiene que considerarse como abierta y sujeta a posibles cambios, dado que el tema de fondo está en permanente debate.

A continuación, se incluyen las definiciones y propiedades en inglés de los catorce elementos propuestos para ESSIM. En un apartado posterior se tratarán los vocabularios utilizados en cada uno de los elementos.

Para comprender mejor el significado de cada elemento, se mostrará un ejemplo de posibles valores para cada uno de ellos. Para ello se usará el caso de la información intercambiada entre Estados miembros de la Unión Europea del capítulo 2 (sección 2.3.3.3), que se basaba en el documento que se muestra de nuevo en la figura 20. Se tratará de describir el dato contenido en el campo “Surname” del documento, mediante los catorce elementos/metadatos ESSIM propuestos. Para este ejemplo no es significativo el valor concreto que tenga ese campo en el documento intercambiado, que es un dato que está incluido en el propio documento. Lo importante es entender que junto al documento se enviaría una descripción con metadatos ESSIM de cada uno de los campos más significativos del documento, entre los que se encontraría en campo “Surname”, con el objetivo de ayudar al destinatario del documento, a comprender el significado del contenido de dicho campo.

<p>ADMINISTRATIVE COMMISSION ON SOCIAL SECURITY FOR MIGRANT WORKERS</p>	<div style="border: 1px solid black; display: inline-block; padding: 2px;">E 106</div> <div style="border: 1px solid black; display: inline-block; width: 40px; height: 20px; margin-left: 10px;"></div> (*)
<p>CERTIFICATE OF ENTITLEMENT TO SICKNESS AND MATERNITY INSURANCE BENEFITS IN KIND FOR PERSONS RESIDING IN A COUNTRY OTHER THAN THE COMPETENT COUNTRY</p> <p>Employed and self-employed persons and members of their families residing with them; members of the family of unemployed persons</p> <p style="font-size: x-small;"><i>Regulation (EEC) No 1408/71: Article 19(1)(a); Article 19(2) and Article 25(3)(i) Regulation (EEC) No 574/72: Article 17(1) and (4) and Article 27 (first sentence)</i></p> <p style="font-size: x-small;"><i>The competent institution should complete Part A of the form and send two copies to the insured person, or send them - where necessary through the liaison body - to the institution in the place of residence if the form is drawn up at that institution's request. As soon as it has received the two copies, the latter institution should complete Part B and return one copy to the competent institution.</i></p> <p>Please complete this form in block letters, writing on the dotted lines only. It consists of four pages, none of which may be left out.</p>	
A. Notification of entitlement	
1.	Institution of the place of residence ⁽²⁾
1.1	Name:
1.2	Identification number of the institution:
1.3	Address:
1.4	Reference: your E 107 form of
2.	The insured person
2.1	Surname(s) ⁽³⁾ :
2.2	Forename(s) ⁽⁴⁾ : Date of birth:
2.3	Previous name(s):
2.4	Address in the country of residence:

Figura 20. Formulario de intercambio según lo previsto en la Decisión 202 (EC, 2005)

3.2.4.1 *Creator Institution*

Element specification	
Element name	creatorInstitution
Label	Creator Institution
Definition	Social security institution primarily responsible for making the data
Obligation	Mandatory
Value range	Institution Repository (Institution)
Multiplicity	1..1
Refines Dublin Core element	Creator

Tabla 11. Elemento *creatorInstitution*

Este elemento (tabla 11) se añadió después de las deliberaciones celebradas por la Comisión Técnica para la creación un directorio de instituciones (Delgado et al., 2012) y guarda relación con la necesidad de que cuente con una interface de acceso público PAI (2015) para uso tanto de los funcionarios de las instituciones como de los ciudadanos en general. De esta forma y según se había previsto en el reglamento de aplicación 987/2009 (OJEU, 2009), los anexos del antiguo reglamento 1408/71 (DOUE, 1971) eran sustituidos por esta herramienta electrónica.

El elemento ***creatorInstitution*** debe contener una referencia a la institución de Seguridad Social que ha creado el dato (el valor contenido en el campo del documento) en el inicio de un proceso de intercambio de información. Su utilización es obligatoria. Sólo puede tener un valor. Refina el elemento *creator* de DC, asociado en este caso con la organización responsable de la creación del contenido. Si bien en el caso de DC se pretende resaltar los aspectos de propiedad intelectual, en el caso de ESSIM se considera necesario conservar en todo el proceso la fuente inicial que ha registrado el dato, motivo por el que se ha modificado el nombre original del elemento.

En ocasiones, la institución propietaria o creadora del dato no es la que finalmente lo envía en el ámbito internacional. Como se ha mencionado anteriormente (sección 2.3.3.2), en el caso de los reglamentos comunitarios se contempla la figura de organismo de enlace. Es aquel que tiene la función de asumir la responsabilidad de establecer el contacto internacional, recibiendo la información de todas las instituciones competentes nacionales y encaminándola al punto de enlace del país de destino, que a su vez se encargará de hacerla llegar a la institución competente final. Figuras similares a la del organismo de enlace aparecen en otros ámbitos geográficos distintos a la Unión Europea.

El contenido de este elemento debería coincidir con una institución incluida en un vocabulario asociado a un repositorio de instituciones como el mencionado en la sección 2.3.3.2 (EC, 2018). En dicho vocabulario, una institución tendrá a su vez una descripción a través de elementos o

metadatos propios, que indiquen su nombre, código, y otros datos de interés. En la sección 3.2.5 se tratará este vocabulario.

Figura 21. Ejemplo de institución registrada en el *Institution Repository* (EC, 2018)

En el caso de la descripción del contenido (dato) del campo *Surname* del documento que se indicó que se usaría como ejemplo, el valor del elemento *creatorInstitution* debería referenciar a la institución que creó dicho contenido, que por ejemplo podría ser el “Instituto Nacional de la Seguridad Social de Madrid (Dirección provincial)”, que está incluido en el *Institution Repository* con identificador 2800 (ver figura 21).

3.2.4.2 Data Set

Element specification	
Element name	dataSet
Label	Data Set
Definition	Core Vocabulary where a similar data concept could be found
Obligation	Optional
Value range	Core Vocabulary (mapping)
Multiplicity	0..1
Refines Dublin Core element	Subject

Tabla 12. Elemento *dataSet*

El elemento *dataSet* (tabla 12) debe referenciar si existe relación (mapping) del campo con algún término de un *Core Vocabulary*. Es un elemento opcional, si bien es recomendable su uso para conseguir la mayor estandarización posible en los intercambios de datos. Puede considerarse un refinamiento del elemento *subject* de DC. Si bien su fin es en ambos casos fomentar el uso de vocabularios controlados, en el caso de ESSIM la pretensión es también facilitar su identificación con los conjuntos de datos identificados y definidos en los *Core Vocabularies*.

Tal como se ha dicho en la sección 2.4.3, el concepto de *Core Vocabularies* trata de establecer los fundamentos en los que se tienen que basar los vocabularios para ser utilizados como metadatos. Un *Core Concept* es un modelo de datos simplificado que reúne las características o atributos mínimos de una entidad de forma genérica e independiente del dominio o país. La Unión Europea ha definido este tipo de vocabularios para conceptos relacionados con personas, negocios, localizaciones, servicios públicos, entre otros (SEMIC, 2018). Para facilitar la utilización de estos vocabularios, se ha publicado una guía para realizar el mapeo de un término utilizado en un contexto determinado, con un concepto incluido en uno de los vocabularios (EU, 2015, capítulo 4).

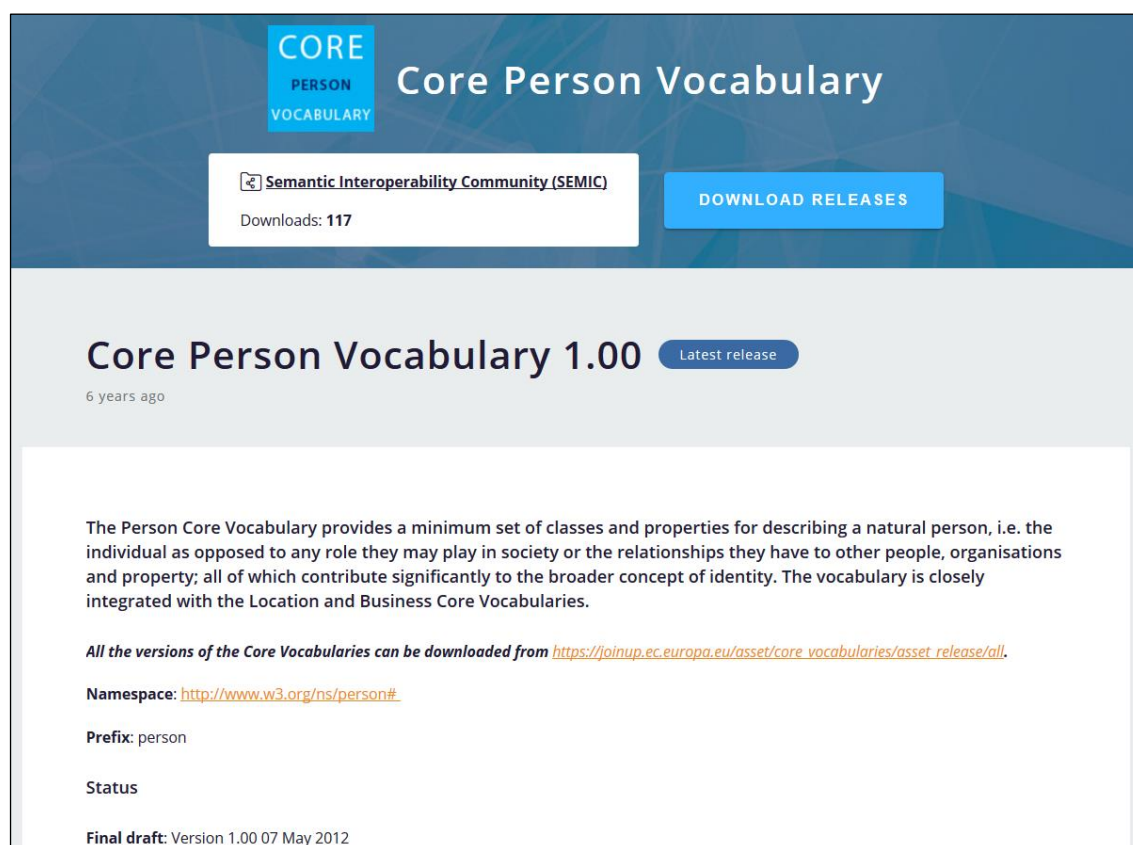


Figura 22. Sitio web del *Core Person Vocabulary* (SEMIC, 2012)

En el caso de ejemplo utilizado, para el campo *Surname* del documento intercambiado, el contenido de este elemento *dataSet* podría hacer referencia (mapear) al vocabulario denominado *Core Person Vocabulary* (figura 22), y dentro de él a la definición de la característica o propiedad *family name* de una persona, indicando un nivel de semejanza (*mapping relation*) cercano (*has closed map*). Este vocabulario, a su vez está relacionado con otros ampliamente utilizados, como

schema.org (Schema, 2018), o FOAF (Brickley y Miller, 2014), por lo que a través de estas relaciones se puede llegar a determinar que un concepto tiene otros equivalentes en dichos vocabularios. En este caso, sería equivalente al término *familyName* que está incluido tanto en schema.org como en FOAF.

3.2.4.3 Data Term

Element specification	
Element name	dataTerm
Label	Data Term
Definition	Term included in the Glossary of Terms (GOT) that defines the meaning of the data field
Obligation	Mandatory
Value range	Glossary of Terms (Term)
Multiplicity	1..1
Refines Dublin Core element	Title

Tabla 13. Elemento *dataTerm*

El elemento *dataTerm* (tabla 13) referencia un término incluido en el Glosario de Términos (GOT) descrito en la sección 2.3.3.3, dónde deberían aparecer todos los datos posibles en los intercambios de Seguridad Social. En este glosario, los términos están descritos mediante un conjunto de propiedades y atributos. En la tabla 14, se muestra el contenido para un término con código "IC001" y nombre "*Surname*". En el ejemplo utilizado, el valor del elemento *dataTerm* sobre el campo *Surname* del documento intercambiado, debería ser precisamente una referencia al término IC001 del GOT.

ID Code	Name	Description	Type	Values	Usag e
IC001	Surname	official name of a person	Text	anything, E: both names at birth	M/I

Tabla 14. Término *Surname* incluido en el GOT (TF, 2007)

Este elemento ESSIM está basado en el elemento *title* de *Dublin Core*, asumiendo la idea de hacer referencia al nombre de un recurso, en este caso a un término del GOT. Es obligatorio y sólo se puede asociar un término a un campo de un documento.

3.2.4.4 *Date Created*

Element specification	
Element name	dateCreated
Label	Date Created
Definition	A point of time associated with the starting point in the lifecycle of data
Obligation	Mandatory
Value range	Literal
Multiplicity	1..1
Refines Dublin Core element	Date

Tabla 15. Elemento *dateCreated*

El elemento ***dateCreated*** (tabla 15) es el momento en el tiempo en que se ha iniciado el ciclo de vida del dato que se está describiendo. Se basa en el elemento *date* de DC, si bien en el caso de EESSIM se refiere al momento inicial del ciclo de vida, mientras que en DC puede reflejar cualquier momento a lo largo de este ciclo. Se trata de un refinamiento de este elemento original, sobre el que el propio organismo responsable de *Dublin Core* ha propuesto otros refinamientos en forma de términos, como *dateAccepted*, *dateCopyrighted* o *dateSubmitted* (DCMI, 2018).

Aunque en principio se acepta como valor un literal que represente una fecha, en el futuro podrían referenciarse un término del Glosario de Términos (GOT) que representara este mismo concepto.

3.2.4.5 *Date Modified*

El elemento ***dateModified*** (tabla 16) debe contener la última fecha en que se modificó el dato descrito. Su interés reside en aquellos casos en que puede haber más de una versión del mismo dato y ser relevante la fecha en que ha sido actualizado por última vez.

Es un elemento opcional. Su necesidad se justifica por aquellos casos en que puede haber distintas versiones de un dato y su modificación puede afectar a los procesos objeto de los intercambios. Los campos 2.1 (*surname*) y 2.2 (*forename*) del formulario de la figura 20 pueden ser también un buen ejemplo en este caso, dado que en un gran número de países es frecuente la modificación de los apellidos, entre otras cosas, por razones de matrimonio. En el caso de intercambios de datos para el control de vivencia esta información es necesaria para evitar errores de interpretación, problema que se ha discutido en la *Death Data Community* (DDC, 2013).

Element specification	
Element name	dateModified
Label	Date Modified
Definition	Date when data was last modified or changed.
Obligation	Optional
Value range	Literal
Multiplicity	0..1
Refines Dublin Core element	Date

Tabla 16. Elemento *dateModified*

3.2.4.6 Description

El elemento *description* (tabla 17) debe contener una descripción textual del dato de una extensión aproximada de entre 25 y 30 palabras o 150 caracteres, con el fin de que sea suficientemente expresiva pero no excesivamente larga. Mantiene el mismo nombre que recibe en los elementos de DC. Se considera un elemento obligatorio.

Element specification	
Element name	description
Label	Description
Definition	An effective description is 25-30 words long or approximately 150 characters.
Obligation	Mandatory
Value range	Literal
Multiplicity	1..*
Refines Dublin Core element	Description

Tabla 17. Elemento *description*

Debería tener al menos una repetición en lengua original y otra en inglés.

El valor del elemento podría coincidir con la propiedad del mismo nombre que se utiliza en el GOT para describir el término que ha sido referenciado a través del elemento *dataTerm*, si bien puede utilizarse para mejorar aquella descripción.

En el ejemplo utilizado sobre el campo *Surname* del formulario de la figura 20, el valor del elemento *description* podría ser “*official name of a person*”. En este caso coincidiría literalmente con la descripción que existe en el GOT para el término IC001, que ha sido referenciado desde el elemento *dataTerm* usado en este ejemplo, aunque podría tener otro texto más detallado.

3.2.4.7 *Format*

Element specification	
Element name	format
Label	Format
Definition	Format of the data
Obligation	Mandatory
Value range	Literal
Multiplicity	1..1
Refines Dublin Core element	Format

Tabla 18. Elemento *format*

El elemento *format* (tabla 18) es obligatorio y se utiliza para describir el formato dentro de los posibles para un dato almacenado de forma electrónica (numérico, alfanumérico, etc.). Coincide con el elemento del mismo nombre de DC.

ID Code	Name	Description	Type	Values	Usag e
IC001	Surname	official name of a person	Text	anything, E: both names at birth	M/I
IC002	Surname at birth	official name of a person at birth, maiden name	Text	anything	O/M/I
IC003	Forenames	first name(s) associated with surname to identify a person	Text	Anything	M/I

Tabla 19. Ejemplos de términos incluidos en el GOT (TF, 2007)

Como ocurría con el elemento *description*, el valor del elemento *format* podría coincidir con la propiedad *Type* que se utiliza en el GOT para describir el formato del término que ha sido referenciado a través del elemento *dataTerm* (tabla 19), si bien puede utilizarse para detallar con más detalle el formato.

En el ejemplo utilizado sobre el campo *Surname*, el valor del elemento *format* podría ser *“Alphabetic text”*. En este caso no coincidiría literalmente con el valor tipo que existe en el GOT para el término IC001, que ha sido referenciado desde el elemento *dataTerm* usado en este ejemplo, ya que este valor en el GOT es *“Text”*, y se ha usado el elemento *format* para especificar que debe ser un texto alfabético, sólo con letras.

3.2.4.8 *Identifier*

Element specification	
Element name	identifier
Label	Identifier
Definition	Identifier of the data (field)
Obligation	Optional
Value range	Literal
Multiplicity	0..1
Refines Dublin Core element	Identifier

Tabla 20. Elemento *identifier*

El elemento *identifier* (tabla 20) contiene la información necesaria para conocer, sin posibilidad de error, a qué campo de un documento se refiere una descripción con metadatos o elementos ESSIM. Actualmente, existen tecnologías que permiten insertar en un documento metadatos asociados a los campos del documento, por ejemplo, en el caso de documentos o formularios HTML, utilizando microdatos (W3C, 2018a), RDFa (W3C, 2015a) o código JSON-LD en etiquetas <SCRIPT> (W3C, 2018b). Sin embargo, en otros casos no es posible insertar en el propio documento los metadatos y éstos se intercambian fuera del mismo, por ejemplo, en un archivo adjunto. En estos casos es necesario identificar a qué campo del documento se refiere la descripción realizada con los elementos ESSIM.

En el ejemplo utilizado como referencia, el campo que se describe es *Surname*, y en el documento tiene asociado un código numérico, que puede ayudar a identificarlo sin ninguna duda (figura 20), por lo que el valor del elemento *identifier* en este caso correspondería a ese código, que en el formulario es *“2.1”*.

3.2.4.9 Language

Element specification	
Element name	language
Label	Language
Definition	Language of the data when not numeric
Obligation	Conditional
Value range	ISO 639
Multiplicity	0..1
Refines Dublin Core element	Language

Tabla 21. Elemento *language*

El elemento **language** (tabla 21) debe contener el idioma en que está expresado el contenido del dato. Coincide conceptualmente con el elemento del mismo nombre de DC original, y del que también existe un refinamiento con el mismo nombre realizado por DCMI en su especificación de términos, en el que le asocia como rango de valores un sistema lingüístico, como pueden ser los idiomas que establece el estándar ISO 639 (DCMI, 2018).

Es un elemento considerado condicional, dado que sólo será necesario en el caso en que lo requiera el contenido del dato. Por ejemplo, no será necesario para un contenido numérico, pero sí en el caso de que se trate de un literal en determinado idioma (por ejemplo, el nombre de institución). Por tanto, se considera obligatorio para los datos no numéricos.

En cuando al contenido, debe ajustarse a uno de los códigos establecidos por el estándar ISO 639 para representar un lenguaje, bien con dos caracteres (ISO 639-1) o bien con tres caracteres (ISO 639-2 e ISO 639-3).

Considerando el ejemplo de la figura 20, si se supone que el contenido del campo *Surname* está en idioma inglés, entonces el valor del elemento *language* sería “en”, tal y como establecen ISO 639-1, o bien “eng”, si se utiliza ISO 639-2 o ISO 639-3.

3.2.4.10 Mandate

El elemento **mandate** (tabla 22) debe contener la referencia legal que da respaldo a los intercambios en los que se ve implicado el dato descrito. Es un refinamiento del elemento *relation* de DC.

Element specification	
Element name	mandate
Label	Mandate
Definition	Legal reference of the exchanges where data is involved
Obligation	Optional
Value range	EUR-Lex (Regulation)
Multiplicity	0..*
Refines Dublin Core element	Relation

Tabla 22. Elemento *mandate*

La necesidad de este elemento se deriva de la base legal que debe amparar a cualquier intercambio de datos entre instituciones de Seguridad Social y el distinto significado o tratamiento que puede tener un dato en función del acuerdo que marque su contexto. Así, por ejemplo, si se trata de intercambios de datos relacionados con el control de vivencia, puede ocurrir que se utilicen diferentes datos con fines de identificación. Sobre este tema se debatió en la reunión celebrada en 2013 por la *Death Data Community* (DDC, 2013), mencionada en la sección 1.1.

En el ámbito de la Unión Europea, su contenido debería estar relacionado con un vocabulario que representase las regulaciones europeas, que están registradas en el repositorio EUR-Lex (EU, 2018). Para el resto de los países, debería basarse en un registro internacional de acuerdos que actualmente no existe.

En el ejemplo utilizado en este capítulo, el elemento *mandate* que forma parte de la descripción semántica del campo *Surname* del formulario de la figura 20, podría hacer referencia a una regulación específica relacionada con el asunto al que se refiere el documento, como por ejemplo "*Regulation (EC) No 883/2004 of the European Parliament and of the Council of 29 April 2004 on the coordination of social security systems*".

3.2.4.11 References

El elemento *references* (tabla 23) debe contener la identificación de los documentos estandarizados en los que hay un campo con el mismo significado que el campo o dato que se está describiendo. Puede considerarse un refinamiento del elemento *relation* de DC, pero también del término *references* de DC.

Su valor debe referenciar a documentos oficiales existentes, tales como los SED (*Standard Structured Document*) de la Unión Europea, los formularios de Mercosur o los formularios acordados en los acuerdos bilaterales.

Element specification	
Element name	references
Label	References
Definition	Official documents where data is present
Obligation	Optional
Value range	Official Documents (Document)
Multiplicity	0..*
Refines Dublin Core element	Relation

Tabla 23. Elemento *references*

Se propone crear un vocabulario para describir documentos oficiales estandarizados, y poder referenciar tipos de documentos como instancias del término *Document* del vocabulario.

En el ejemplo utilizado en este apartado, se podrían referenciar documentos oficiales en los que aparezca también el campo *Surname*, como por ejemplo el documento E101 (*Certificate concerning the legislation applicable*).

3.2.4.12 Relation with Terms

El elemento *relationWithTerms* (tabla 24) indica si existe relación del campo o dato con otros términos del GOT diferentes al que ya se cita en el elemento *dataTerm*. La utilización de este elemento es útil para poner de relieve dependencias y establecer controles. Sería necesario un estudio sistemático partiendo de un glosario de términos.

Una relación básicamente consiste en una estructura con sub-elementos: *relationTerm*, cuyo valor es una referencia a un término del GOT; y *relationDescription*, que contiene una descripción del motivo de esa relación.

Element specification	
Element name	relationWithTerms
Label	Relation with Terms
Definition	Description of relationship between the data and other terms, different of the dataTerm
Obligation	Optional
Value range	- relationTerm: Glossary of Terms (Term) - relationDescription: Literal
Multiplicity	0..*
Refines Dublin Core element	Relation

Tabla 24. Elemento *relationWithTerms*

En el ejemplo de referencia de la figura 20, se podría indicar que el campo *Surname* está relacionado con el término del GOT llamado *Previous names*, y la descripción de la relación podría ser: “*The Surname value has officially substituted the previous names of the person*”.

3.2.4.13 Sender Institution

Element specification	
Element name	senderInstitution
Definition	Social security institution responsible for sending the data
Obligation	Optional
Value range	Institution Repository (Institution)
Multiplicity	0..1
Refines Dublin Core element	Publisher

Tabla 25. Elemento *senderInstitution*

El elemento ***senderInstitution*** (tabla 25) debe contener una referencia a la institución que envía el dato a otro país y debe coincidir con una institución existente en el *Institution Repository*. Refina el elemento *Publisher* de DC. En el caso de DC se hace referencia al responsable de que el recurso se encuentre disponible en la red. En el caso de ESSIM, la referencia es similar, pero se modifica su nombre para adaptarlo al contexto y reflejar la función de este elemento.

El argumento para la inclusión de este elemento es el mismo empleado en el elemento *creatorInstitution*. En este caso correspondería al código de la institución a la que se le han asignado las competencias de organismo de enlace.

Un ejemplo de esta situación suele darse en las instituciones competentes en materia de asistencia sanitaria que deben facturar por servicios prestados a ciudadanos de otros países. Las competencias de organismo de enlace normalmente las tiene asignadas una sola institución por país. Otras instituciones facturan a organismos extranjeros a través del organismo de enlace. Sin embargo, es este último quien reúne todas las facturas de un periodo determinado y las envía al organismo de enlace del país que corresponda. En tal caso, *creatorInstitution* sería la institución que factura y *senderInstitution* el organismo de enlace.

En el ejemplo utilizado, el valor del elemento *senderInstitution* podría ser “Instituto Nacional de la Seguridad Social (Dirección General)”, en su papel de organismo de enlace, mientras que el elemento *creatorInstitution* podría corresponder a una Comunidad Autónoma, si es esta quien ha creado el documento electrónico.

3.2.4.14 Social Security Category

Element specification	
Element name	socialSecurityCategory
Label	Social Security Category
Definition	Branch of social security involved in the exchange of data
Obligation	Mandatory
Value space or data type	Institution Repository (Category)
Multiplicity	1..1
Refines Dublin Core element	Coverage

Tabla 26. Elemento *socialSecurityCategory*

El elemento ***socialSecurityCategory*** (tabla 26) debe contener información sobre el sector de Seguridad Social con el que está relacionado el dato referenciado. Es un refinamiento del elemento *coverage* de DC, si bien en el caso de ESSIM, la referencia se acerca más al sentido de AENOR de

“jurisdicción” (AENOR, 2011). La información relevante en el caso de Seguridad Social son los distintos sectores del campo de aplicación (jubilación, salud, protección familiar, etc.). En ciertas ocasiones, el alcance puede comprender todos los sectores.

Su valor deber referenciar a una de las categorías establecidas en el *Institution Repository* (IR). En dicho repositorio, cada institución registrada tiene asociado una propiedad denominada “*Social Security Category*” que incluye todas las categorías en las que está implicada dicha institución (figura 23). La categoría indicada al elemento *socialSecurityCategory* debería coincidir con alguna de las categorías que tenga asociada en el IR la institución que ha creado el dato, es decir la que se referencia en el elemento *creatorInstitution*.

En el ejemplo de referencia, el valor del elemento *socialSecurityCategory* para el campo *Surname* tendría que ser algunas de las categorías reconocidas para la *creatorInstitution* en el IR, que era “Instituto Nacional de la Seguridad Social de Madrid (Dirección provincial)”. Por tanto, en este caso una categoría adecuada para ese elemento podría ser “*Maternity, paternity*”, que como puede observarse en la figura 23, es una de las asociadas a dicha institución, y es apropiada para el tipo de documento en el que se encuentra incluido el dato que se está describiendo, que es un certificado relacionado con las prestaciones por maternidad.

The screenshot shows the EESSI Directory interface for the 'INSTITUTO NACIONAL DE LA SEGURIDAD SOCIAL DE MADRID DIRECCION PROVINCIAL'. The page includes a navigation breadcrumb: 'European Commission > Employment, Social Affairs and Inclusion > EESSI Directory > Browse > Spain > Current Institution'. The main content area features a title bar with 'Export' and 'Back' buttons, followed by the institution's name and its full name in English: 'National Institute of Social Security, Provincial Head Office of Madrid'. Below this, there are two tabs: 'Institution Details' and 'Function Group'. The 'Function Group' tab is active, showing a list of 'Categories of Social Security' associated with the institution. These categories are: Maternity, paternity; Invalidity – state invalidity pension scheme; Old Age pensions – state pension scheme; Survivors' pension; Accident at work; Occupational diseases; and Death grants.

Figura 23. Algunas categorías asociadas a una institución en el IR (EC, 2018)

3.2.5 Vocabularios controlados utilizados por ESSIM

El tema central de la propuesta es una especificación de conjunto de metadatos (elementos) que llegue a convertirse en un estándar para los intercambios de datos entre instituciones de Seguridad Social. Para que la estandarización sea eficaz y consiga su fin, debe apoyarse en una serie de vocabularios controlados, de tal forma que su contenido identifique de forma inequívoca y sin posibilidad de confusión los datos que definen. Un ejemplo sería el vocabulario aportado por Build 6 (proyecto TESS), tal como se ha descrito en la sección 2.3.3.3.

En otras palabras y tal como lo expresa Harpring (2010), es necesario acordar los vocabularios controlados para asegurar la consistencia y coordinación de valores de datos. El mismo DC recomienda el uso de valores controlados para los campos donde sea apropiado, y ha definido un conjunto de términos que pueden utilizarse para ello (DCMI, 2018). En la misma línea, Craig (2001) afirma que mientras los metadatos proporcionan una descripción de los datos, los vocabularios de datos estandarizan el contenido de estos datos.

A continuación, se describen los vocabularios propuestos por ESSIM e incluidos en la tabla 10, si bien en el capítulo 5 se hablará de su evolución necesaria.

3.2.5.1 Literales

Algunos de los elementos propuestos en ESSIM tienen valores literales, entendiendo como concepto “literal” el establecido por el W3C es algunas de sus recomendaciones, como es el caso de RDF (*Resource Description Framework*), donde se indica que se refiere a valores como cadenas de caracteres, números o fechas (W3C, 2014).

Los valores de los elementos *dateCreated* y *dateModified* serían fechas, mientras que los de los elementos *description*, *format*, *identifier* y *relationDescription* serían cadenas de caracteres.

3.2.5.2 Glosario de Términos

En la sección 2.3.3.3 se ha descrito el “Glosario de Términos” creado dentro de Build 6 (proyecto TESS) con vistas a que las instituciones mejoraran el procesamiento de los datos recibidos del exterior. Como se dijo entonces, este glosario contenía las definiciones de los conceptos relacionados con los intercambios de datos de Seguridad Social. Consistía en una primera parte genérica, seguida de una definición específica para cada Estado miembro.

El elemento *dataTerm* se basa en el vocabulario que representa el GOT, y su valor debe coincidir con alguno de los términos incluidos en el GOT.

El GOT se limita al ámbito de la Unión Europea y más concretamente a los países que la integraban en 1995.

Lo que se propone es la creación de un vocabulario de libre acceso que represente dicho glosario de términos, en el que cada término quede descrito, al menos, por las siguientes propiedades (tabla 27), en principio todas con valores literales, pudiendo en el futuro ampliarse el glosario con nuevas estructuras de datos y tipos específicos.

GOT Term properties	
Code	Unique identification of the term in the glossary
Name	Readable label of the term
Description	Brief description of the term
Type	Qualifier of the term (text, data, simple choice, etc.).
Values	The values specified for the terms, if any
Usages	Social security sectors in which it is used

Tabla 27. Propiedades de los términos del GOT

Ejemplos de posibles valores para estas propiedades pueden encontrarse en la tabla 19.

Como propuesta de futuro, se propone también integrar en el GOT relaciones entre términos (tabla 28), lo cual permitiría dejar de utilizar el elemento ESSIM *relation*, pues a través del término indicado en el elemento *dataTerm* asociado a un dato, se podría acceder en el GOT a todos los otros términos relacionados con él. En este caso se podría definir en el vocabulario otro tipo concepto además del ya existente “término”, sería el concepto de “relación entre términos”, cuyas propiedades podrían ser las siguientes:

GOT Relation properties	
relationTermA	One term in the relation
relationTermB	Another term in the relation
relationDescription	Relationship description between terms A and B.

Tabla 28. Relaciones entre términos en el GOT

3.2.5.3 *Institution Repository*

En la sección 2.3.2.4 se ha descrito el repositorio de instituciones desarrollado por el proyecto EESSI, antes denominado Directorio Maestro. Tal como se ha dicho, este repositorio permite que una institución sea capaz de contactar con otras instituciones gracias a las facilidades del directorio de servicios y a los mecanismos establecidos tanto a nivel nacional como internacional por el proyecto EESSI.

En la especificación ESSIM, el repositorio de instituciones es utilizado es los elementos *creatorInstitution* y *senderInstitution* para referenciar instituciones incluidas en el repositorio. A través de esta referencia, en el vocabulario se pueden encontrar todos los detalles de la institución correspondiente, como su identificador, nombre en diferentes idiomas, dirección, etc.

También se utiliza en el elemento *socialSecurityCategory*, aunque en este caso para referirse a alguna de las categorías definidas en dicho repositorio, y que pueden ser asignadas a las instituciones a través de su propiedad "*Categories of Social Security*" en el repositorio.

Como en el caso del GOT, su ámbito se limita en estos momentos al Espacio Económico Europeo.

Lo que se propone es la creación de un vocabulario de libre acceso que represente el repositorio de instituciones, en el que cada institución quede descrita, al menos, por las propiedades indicadas en la tabla 29). Sólo se muestran algunas relevantes para los metadatos propuestos el capítulo anterior.

IR Institution properties	
Identification	Identification Number of the Institution
Acronym	Acronym of the Institution
FullName	Full name of the Institution
fullName2	Full name of the institution (lang 2)
fullName3	Full name of the institution (lang 3)
FullNameEnglish	Full name of the institution (ENGLISH)
FullNameLatin	Full name of the institution in Latin (if applicable)
SocialSecurityCategories	Coverage of social security matters
....	

Tabla 29. Propiedades de las instituciones

Hay una propiedad (*socialSecurityCategories*) cuyo valor haría referencia a otro concepto que estaría también definido en el vocabulario (*Category*). Las propiedades del concepto *Category*, serían al menos las indicadas en la tabla 30.

IR Category properties	
categoryId	Identifier of the category
categoryName	Name of the category
categoryDescription	Description of the category
...	

Tabla 30. Propiedades del concepto *Category*

3.2.5.4 Core Vocabularies

En la sección 2.4.3 se han descrito estos recursos definidos por la Comisión Europea dentro del programa ISA, ampliamente estudiados en (EU, 2015). En el caso de la especificación de metadatos que nos ocupa, se trata de utilizar modelos de datos estandarizados siempre que estos existan. Por ese motivo se sugiere que el elemento *dataSet* haga referencia a este modelo siempre que el campo descrito este recogido en alguno de los desarrollados hasta el momento. Para ello se debe aplicar un mecanismo de mapeo (*mapping*) que está descrito en (EU, 2015).

Este mecanismo permite establecer en grado de alineamiento o "*mapping relation*" de un término utilizado, con alguno de los presentes en los Core Vocabularies, que puede ser uno de los siguientes: "*Has exact match*", "*Has close match*", "*Has related match*", "*Has broad match*", "*Has narrow match*".

Como en el caso del GOT y el Repositorio de Instituciones, su ámbito se limita en estos momentos al Espacio Económico Europeo.

3.2.5.5 ISO 639

Uno de las causas habituales de los problemas semánticos es el multilingüismo. Por ese motivo, se ha incluido el elemento *language* que permita establecer en que lengua está escrito el dato que se describe. Para evitar cualquier confusión, es necesario que su contenido responda a un estándar reconocido, por lo que se ha elegido ISO 639. Este estándar proporciona un código de dos o tres caracteres y es utilizado ampliamente en el ámbito de las administraciones públicas, como es el caso del Gobierno de los Estados Unidos. Aunque en general se puede utilizar la versión ISO 639-1 (ISO, 2002), que usa dos caracteres para identificar una lengua, también se puede utilizar ISO 639-2 (ISO, 1998) o ISO 639-3 (ISO, 2007), que usa tres caracteres, en aquellos casos en los que el idioma no esté incluido en la primera versión.

Dublin Core Metadata Initiative incluye en su especificación de términos recomendados para utilizar con DC, las versiones ISO 639-2 e ISO 639-3 como posibles lenguajes (DCMI, 2018). Por lo que, se puede reutilizar dicho vocabulario disponible públicamente, y referenciarlo desde el elemento *language* de ESSIM. En la figura 24, se muestra un extracto de dicho vocabulario, en el que aparece el término ISO 639-2.

Term Name:	ISO639-2
URI:	http://purl.org/dc/terms/ISO639-2
Label:	ISO 639-2
Definition:	The three-letter alphabetic codes listed in ISO639-2 for the representation of names of languages.
See:	http://lcweb.loc.gov/standards/iso639-2/langhome.html
Type of Term:	Datatype
Version:	http://dublincore.org/usage/terms/history/#ISO639-2-003

Figura 24. Extracto del vocabulario *DCMI Metadata Terms* (DCMI, 2018)

Desde el vocabulario de DCMI se hace referencia al sitio web de la Biblioteca del Congreso de Estados Unidos donde se mantienen datos abiertos con archivos en diferentes formatos sobre algunas versiones de este estándar, como por ejemplo ISO639-1 (figura 25).

The screenshot shows the Library of Congress website interface. At the top, there is a navigation bar with the Library of Congress logo and buttons for 'ASK A LIBRARIAN', 'DIGITAL COLLECTIONS', and 'LIBRARY CATALOGS'. A search bar is also present with the text 'Search Loc.gov' and a 'GO' button. Below the navigation bar, the page title is 'The Library of Congress > Linked Data Service'. The main content area is titled 'ISO 639-1: Codes for the Representation of Names of Languages - Part 1: Two-letter codes for languages'. There is a search box with the placeholder text 'Enter Keyword or Phrase' and a dropdown menu showing a list of language codes: 'ISO639-1 Languages', 'ISO639-2 Languages', 'ISO639-5 Languages', 'Extended Date/Time Format Identifiers'. To the right of the search box, there is a note: '**Please Note: LC Classification entries are not included in general search results. You must explicitly select LC Classification in order to search the scheme. This is temporary while the impact of adding LCC to the current system is better understood.' Below the search box are 'Search' and 'Reset' buttons. At the bottom, there are tabs for 'Details' and 'Visualization'. The 'Details' tab is active, showing the title 'ISO 639-1: Codes for the Representation of Names of Languages - Part 1: Two-letter codes for languages' and a description: 'ISO 639-1 is the first part of the ISO 639 international-standard language-code family. ISO 639-1 provides two-character lowercase alphabetic strings that serve as identifiers of languages. The list contains approximately 180 discrete codes. All ISO 639-1 languages also have ISO 639-2 three-character code representations. These codes are linked to codes for the same languages in ISO 639-2 and the MARC Language Codes.' Below the description is a URI link: 'URI(s) > <http://id.loc.gov/vocabulary/iso639-1>'.

Figura 25. Web con datos abiertos sobre ISO639-1 (LC, 2018)

3.2.5.6 EUR-Lex

EUR-Lex es un servicio open data que proporciona acceso gratuito a las leyes de la Unión Europea y otros documentos públicos (EU, 2018). Esta base de datos se actualiza diariamente y contiene más de tres millones de entradas con textos que pueden datarse hasta en 1951. Cada uno de estos documentos está registrado con metadatos analíticos, como referencia de su publicación, fecha o palabras claves (figura 26).

Figura 26. Extracto de información sobre un documento en EUR-Lex (EU, 2018)

El elemento de ESSIM *mandate* se ha concebido para hacer referencia a este tipo de información, dado que la legislación aplicable es relevante para el tratamiento a realizar con los datos intercambiados. Por lo que se propone la creación de un vocabulario que defina el concepto “Regulation”, y recoja las principales propiedades que sirvan para describir cualquier regulación o legislación, como las siguientes (tabla 31):

EUR-Lex Regulation properties	
Id	Code to identify the regulation
Title	Readable title of the regulation
fullText	Document with the full text of the regulation
FullTextEnglish	Document with the full text of the regulation in English
...	

Tabla 31. Propiedades propuestas para EUR-Lex

3.2.5.7 Official Documents

Es importante saber en qué tipo de documentos oficiales (formularios) utilizados en el ámbito de la Seguridad Social aparece un determinado dato. Con ese fin se ha propuesto el elemento *references* en ESSIM. Su contenido debe referirse a un vocabulario donde figure la descripción de los documentos o formularios existentes y susceptibles de ser intercambiados.

Official Form properties	
Contributor	An entity responsible for making contributions to the form
Coverage	The spatial or temporal topic of the form, the spatial applicability of the form, or the jurisdiction under which the form is relevant
Creator	An entity primarily responsible for making the form
Date	A point or period of time associated with an event in the lifecycle of the form
Description	An account of the form
Format	The file format, physical medium, or dimensions of the form
Identifier	An unambiguous reference to the form within a given context
Language	A language of the form
Publisher	An entity responsible for making the form available
Relation	A related form
Rights	Information about rights held in and over the form
Source	A related resource from which the described form is derived
Subject	The topic of the form
Title	A name given to the form
Type	The nature of the form

Tabla 32. Vocabulario sobre tipos de documentos oficiales

Para ello se propone la creación de un vocabulario sobre tipos de documentos oficiales, en el que la descripción de un formulario se haría utilizando los quince elementos de *Dublin Core* con su definición original (tabla 32).

3.2.6 Publicación de los vocabularios creados

Se ha publicado en GitHub la definición de los vocabularios creados. Para ello se ha creado un proyecto denominado “socialsec” y un repositorio “voc”, ubicado en: <https://github.com/socialsec/voc/> (figura 27).

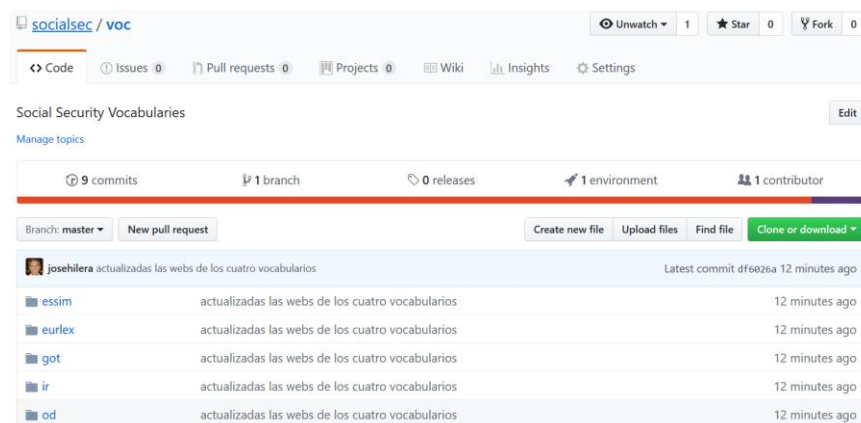


Figura 27. Publicación de los vocabularios creados en GitHub

Los vocabularios creados están publicados en:

- Vocabulario ESSIM: <https://socialsec.github.io/voc/essim/>
- Vocabulario GOT: <https://socialsec.github.io/voc/got/>
- Vocabulario IR: <https://socialsec.github.io/voc/ir/>
- Vocabulario EUR-LEX: <https://socialsec.github.io/voc/eurlex/>
- Vocabulario OD: <https://socialsec.github.io/voc/od/>

Por simplicidad, se ha utilizado una plantilla similar a la usada por el programa ISA (*Interoperability Solutions for public Administrations*) de la Unión Europea en vocabularios como *Person Core*: <https://www.w3.org/ns/person> (figura 28).

ESSIM Vocabulary

This document is also available as an [RDF schema](#)

Abstract

The ESSIM Vocabulary provides a minimum set of classes and properties for describing data (files) in Social Security forms.

Status

The ESSIM Vocabulary was developed by UAH and ISSA and all documentation is available on <https://socialsec.github.io/voc/>.

Copyright © 2018 [UAH](#), [ISSA](#)

The vocabulary is published under [CC BY Licence](#).

Namespaces used

prefix	namespace
cvmap	http://data.europa.eu/core-vocabularies/
dc	http://purl.org/dc/elements/1.1/
dcterms	http://purl.org/dc/terms/
iso639-1	http://id.loc.gov/vocabulary/iso639-1
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs	http://www.w3.org/2000/01/rdf-schema#
xsd	http://www.w3.org/2001/XMLSchema#
essim	https://socialsec.github.io/voc/essim/#
eurlex	https://socialsec.github.io/voc/eurlex/#
got	https://socialsec.github.io/voc/got/#
ir	https://socialsec.github.io/voc/ir/#
od	https://socialsec.github.io/voc/od/#

Namespaces used in this schema

Summary of terms

The ESSIM Vocabulary defines the following terms.

Class	essim:Field
label	Data field
comment	Data field in a Social Security form.

Figura 28. Publicación de los vocabularios creados en GitHub

Capítulo 4. Demostración

4.1 Introducción

En el capítulo 3 se ha propuesto una especificación de conjunto de metadatos para solucionar los problemas semánticos en los intercambios de datos de Seguridad Social.

Según se decía en el mencionado capítulo, para la puesta en práctica de la solución, la forma más habitual es el desarrollo de una herramienta que permita la creación y mantenimiento de metadatos desde un entorno ofimático, para la posterior generación de los correspondientes esquemas XML. En definitiva, partiendo de un modelo de datos en UML, sería necesario completar la descripción con una capa de definición de datos accesible por el usuario final y con especial atención a los problemas semánticos que pudieran producirse en los intercambios.

Por otra parte, según (Harping, 2010), para mejorar los escenarios de los intercambios, los sistemas TIC deberían incluir mecanismos para:

- Automatizar validaciones, incluyendo mecanismos para revisar los datos recibidos.
- Proporcionar información sobre el estatus (trazas, etc.,) de las operaciones
- Notificar los cambios sobre la situación personal y familiar de los beneficiarios.
- Controlar no solo el formato de los datos sino también su significado, por ejemplo, subrayando los conceptos, con el fin de mejorar la correspondencia de datos y las validaciones.

Teniendo en cuenta estas premisas, en el presente capítulo se van a recoger las pruebas realizadas para demostrar la viabilidad de la especificación propuesta en esta tesis.

La primera prueba realizada se centró en la creación y validación de los metadatos. Con ese fin, se construyó una aplicación que permitía la generación manual del conjunto inicial de elementos de ESSIM presentados en las publicaciones (Delgado et al., 2013c; Delgado et al., 2013b). A través de un formulario HTML, la aplicación muestra los elementos agrupados por el vocabulario controlado de referencia, permite introducir sus valores y verificar su corrección accediendo a una función de validación. A continuación, se crea un fichero RDF susceptible de ser tratado por otra aplicación o bien transmitido a otra institución junto con los datos correspondientes.

Para hacer posible esta prueba, previamente fue necesario tratar los vocabularios existentes para contar con una versión compatible con el procedimiento previsto.

La segunda prueba realizada fue mucho más amplia. Consistió en crear una arquitectura simplificada para un caso de estudio, basada en servicios web. Mediante cinco servicios se construyó el proceso completo desde la solicitud de una pensión hasta su resolución, incluyendo la creación y validación de los metadatos. Si bien el proceso real es mucho más complejo, la simplificación realizada permitió reflejar la parte semántica de la interoperabilidad que se pretende garantizar.

Para poder validar el vocabulario ESSIM se montó un SPARQL Endpoint de pruebas con la colaboración la Universidad Técnica Particular de Loja (Ecuador). Este centro de datos se montó con

la ayuda de Virtuoso, un servidor que soporta datos RDF y permite la gestión de datos relacionales, datos XML, implementación de servicios web, etc.

Para realizar las pruebas se introdujeron datos ficticios de tres instituciones de España, tres de Bélgica y una de Italia. Por ese motivo se agregaron los valores necesarios para identificar el país y la ciudad a la que pertenecen, así como su dirección de página web.

Esta arquitectura fue presentada en el VI Congreso Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones Avanzadas (ATICA 2014), celebrado en Alcalá de Henares del 29 al 31 de octubre de 2014 (Viñán et. al, 2014).

En las dos pruebas descritas pueden encontrarse diferencias entre la especificación propuesta en esta tesis y la que se maneja en el software creado. Se deben a la evolución de la propuesta a la vista de las investigaciones realizadas hasta llegar a su versión final, si bien las diferencias no afectan a lo que podría considerarse como objetivo principal del testeo.

Asimismo, ambas pruebas tienen en común la utilización de la tecnología RDF para manejar los metadatos. Dicha tecnología permite no sólo asignar valores concretos a las propiedades sino también referenciarlas desde otros recursos de información que pueden ayudar a entender los valores adscritos. Estos recursos son definidos formalmente de acuerdo con el estándar RDF y, en consecuencia, cumplen plenamente las reglas establecidas por esta tecnología. Su estructura y valores son determinados por vocabularios y ontologías.

Dada su relevancia en las comprobaciones realizadas, se va a dedicar la sección 4.2 de este capítulo a introducir las tecnologías de la web semántica, en particular a la utilización de RDF.

4.2 Tecnologías utilizadas

4.2.1 La Web semántica

Las tecnologías de Web semántica, como RDF, RDF Schema, OWL o SPARQL; permiten la creación de almacenes de datos en la web, construir vocabularios y ontologías, así como escribir reglas para manejar los datos (Shadbolt et al., 2006).

El objetivo principal es mejorar Internet ampliando la interoperabilidad entre los sistemas informáticos y reducir la necesaria mediación de operadores humanos. Con ese fin principal, la Web semántica propone superar las limitaciones de la Web actual mediante la introducción de descripciones explícitas del significado, la estructura interna y la estructura global de los contenidos y servicios disponibles en la Web. A diferencia de la semántica implícita, el crecimiento caótico de recursos y la ausencia de una organización clara de la Web actual, la Web semántica aboga por clasificar, dotar de estructura y anotar los recursos con semántica explícita procesable por máquinas.

Con ese propósito, se articula en torno a una arquitectura formada por una serie de capas que permite resolver los problemas a los que se enfrenta. En la figura 29 podemos apreciar las distintas capas de la Web semántica.

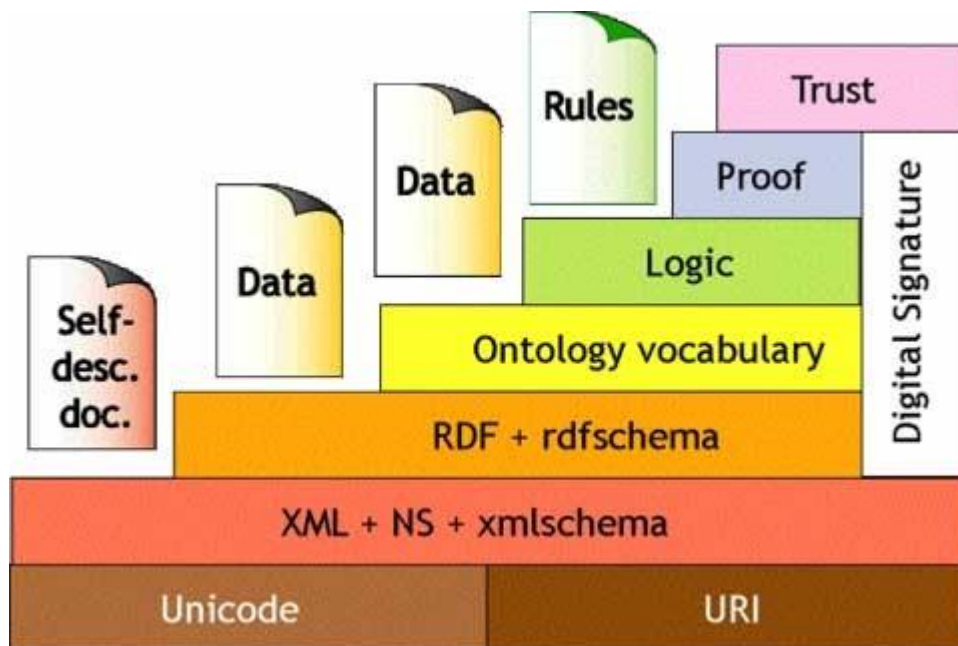


Figura 29. Arquitectura de la Web Semántica (Shadbolt et al., 2006)

En la primera capa se sitúan los estándares necesarios para la transmisión de información. Esto implica la existencia de un lenguaje de descripción de datos (Unicode) y un esquema de nombrado (*Uniform Resource Identifier*, URI). Unicode es un estándar cuyo objetivo es proporcionar el medio por el cual un texto en cualquier forma e idioma pueda ser codificado para el uso informático. Las URI podemos considerarlas como una extensión de las URL que permiten nombrar no solo recursos disponibles en la Web, sino que es un esquema genérico de nombrado, independientemente de la naturaleza del objeto que se nombra. Esta tecnología es fundamental en la Web semántica y es la base por la cual se representan los objetos, propiedades y conceptos.

En la segunda capa se encuentran estándares como XML (*Extensible Markup Language*) que ofrece un formato común para el intercambio de documentos. También en esta capa, *Namespaces* (NS) proporciona un método para cualificar elementos y atributos de nombres usados en documentos XML asociándolos con espacios de nombre identificados por referencias de URI. Asimismo, *XML Schema* es un lenguaje que permite describir la estructura y restringir el contenido de documentos XML.

La tercera capa corresponde a la representación del conocimiento. Conformar la base de la arquitectura proporcionando un lenguaje de representación del conocimiento. Mediante este lenguaje se expresan los contenidos de la Web semántica. El estándar utilizado es RDF (*Resource Description Framework*), al que se dedica la sección siguiente. Se trata de un modelo de datos para los recursos y las relaciones que se puedan establecer entre ellos. Aporta una semántica básica para este modelo de datos que puede representarse mediante XML.

El resto de las capas se consideran fuera del alcance de las pruebas realizadas.

A continuación, se va a profundizar en las dos tecnologías más utilizadas: RDF y servicios web.

4.2.2 Resource Description Framework (RDF)

“*RDF Schema*” es un vocabulario para describir las propiedades y las clases de los recursos RDF, con una semántica para establecer jerarquías de generalización entre dichas propiedades y clases. Esta información puede ser serializada en diferentes formatos, como Turte, JSON-LD, RDFa, N-Triples o RDF/XML. Una descripción de recurso en RDF es un formato de dato directo, etiquetado gráficamente, para representar la información sobre un recurso.

RDF (2004) nació en el seno de la W3C como consecuencia de la necesidad de tener un lenguaje para describir los recursos existentes en la Web, o como sugiere su nombre, debido a la necesidad de tener un lenguaje para expresar metainformación.

Este modelo se basa en la idea de convertir las declaraciones de los recursos en expresiones con la forma sujeto-predicado-objeto (conocidas en términos RDF como triples). El sujeto es el recurso, es decir, aquello que se está describiendo. El predicado es la propiedad o relación que se desea establecer acerca del recurso. Por último, el objeto es el valor de la propiedad o el otro recurso con el que se establece la relación.

Debido a la similitud entre los conceptos “información” y “metainformación”, se entiende que el uso de RDF se puede extender no sólo a describir recursos, como indican sus siglas, sino que se puede aplicar a la descripción de cualquier objeto que pueda ser identificado. Cuando un objeto sea identificado, éste podrá ser determinado en términos de propiedad o valor, debido a lo cual RDF puede utilizarse para describir objetos que se encuentren en la Web, fuera de ella, e incluso elementos abstractos sin representación física.

Así RDF surgió como un lenguaje genérico para la descripción de elementos (cualquier entidad que pueda ser identificada) a través de metainformación; la cual podrá incluir las propiedades de dicho elemento, así como las relaciones que se puedan dar con el resto de ellos, elaborando representaciones más complejas, pero siempre procesables por un sistema informático diseñado para tal efecto. A esa descripción de los diferentes elementos, junto con sus relaciones establecidas (contexto), se denominará conocimiento. En definitiva, RDF es un lenguaje para la descripción de conocimiento.

En líneas generales, RDF permitirá establecer una serie de afirmaciones que constarán de sujeto, verbo y predicado. Como se ha comentado anteriormente, estas afirmaciones podrán estar relacionadas, por lo que el predicado de una de ellas podrá ser sujeto de otra, formando con ello redes semánticas o de conocimiento.

La especificación de desarrollo RDF está definida en los siguientes documentos:

- *RDF/XML Syntax Specification*. Encargado de la descripción de la sintaxis utilizada para expresar el modelo RDF así como sus esquemas.
- *RDF Vocabulary Description Language: RDF Shema*. En este documento se describen los esquemas RDF (RDFS).

- *RDF Primer*. Se trata de un documento introductorio a RDF, que pretende dar una visión general para introducir al lector en el estándar. Este documento no forma parte de lo que sería la especificación formal.
- *Resource Description Framework (RDF). Concepts and Abstract Syntax*. En este documento se procede a realizar una explicación del modelo abstracto en el que está basado RDF.
- *RDF Semantics*. Se especifica en este documento una semántica precisa para el vocabulario definido por RDF, así como las reglas de inferencia tanto para RDF como para los esquemas RDFS.
- *RDF Test Cases*. Se trata de un conjunto de test que recogen las reglas gramaticales de RDF.

Un esquema RDF (RDFS) es un lenguaje de descripción de vocabularios. Utilizar esquemas RDF permite crear vocabularios para su utilización en gráficos RDF. El uso de gráficos y vocabularios RDF creados con un esquema RDF, tiene importantes beneficios (Schutte, M. 2009), como el de permitir relacionar los datos con datos en RDF.

Como ejemplo práctico relacionado con la propuesta de esta tesis, conviene señalar que DCMI decidió en 2008 modelar *Dublin Core* como un conjunto de elementos RDF. El conjunto comprende los elementos originales, sus extensiones, así como elementos añadidos por ser apropiados al entorno de los *linked data* (Coyle, 2012).

4.2.3 Servicios web

Internet introduce un nuevo entorno donde el software se puede ofrecer y acceder como servicio. En los últimos años la mayoría de los procesos de negocio han cambiado en flexibilidad, interconectividad y autonomía debido a las condiciones del mercado, a los nuevos modelos organizacionales y a los escenarios de uso de los sistemas de información. En este contexto, Internet y la Web están cambiando la forma en la que se ofrecen los negocios y los servicios a la sociedad global, y en la que estos negocios interoperan. Esta tendencia nos lleva a sistemas de información conectados e integrados a través de la infraestructura que proporciona Internet.

Para el desarrollo de aplicaciones basadas en la Web, podemos encontrar una gran cantidad de tecnologías, muchas de ellas incompatibles entre sí, por esta razón surgieron los servicios web, como una alternativa para facilitar la intercomunicación entre diferentes arquitecturas software.

Los servicios web proporcionan la plataforma ideal para conseguir la completa integración de los procesos de negocio de una determinada organización con el resto de implicados en dichos procesos (clientes, proveedores, etc.).

Un servicio web es un sistema de software diseñado para hacer posible la interoperabilidad en la interacción máquina-máquina a través de una red.

La definición de servicio web la podemos encontrar en la W3C: un sistema software identificado por una URI cuyos interfaces públicos y enlaces se definen y describen utilizando XML. Otros sistemas

de software pueden interactuar con el servicio Web de la forma prescrita por su definición, usando mensajes basados en XML, a través de protocolos y estándares de Internet (W3C, 2009).

Estos servicios se apoyan en un conjunto de estándares (WSDL, UDDI y SOAP) basados en XML que permiten a los desarrolladores implementar aplicaciones distribuidas, utilizando herramientas muy distintas para crear aplicaciones que utilizan una combinación de módulos de software que son llamados desde diversos sistemas distribuidos en regiones geográficas distintas.

Los servicios Web son aplicaciones auto-contenidas y modulares que pueden ser:

- Descritas mediante un lenguaje de descripción de servicio, como el lenguaje WSDL (*Web Service Description Language*) (WSDL, 2007).
- Publicadas y encontradas, al incluir las descripciones y políticas de uso en algún registro conocido, utilizando el registro UDDI (*Universal Description, Discovery and Integration*) (UDDI, 2004).
- Asociadas, al utilizar la información contenida en la descripción del servicio para crear una instancia de servicio disponible.
- Invocadas sobre la red, al utilizar la información contenida en los detalles de enlace de la descripción del servicio; en un documento WSDL. Durante la invocación, utilizaremos el protocolo SOAP de comunicación (SOAP, 2007).
- Compuestas con otros servicios para integrar servicios y aplicaciones nuevas, en lo que constituirá la base de SOA (*Service Oriented Architecture*), que será explicada en detalle en un apartado posterior.

Los estándares explicados anteriormente (WSDL, UDDI y SOAP) se consideran la primera generación dentro del desarrollo de los servicios web, posteriormente surgió una segunda generación denominada "WS-*" compuesta por un conjunto de nuevas especificaciones que tratan de resolver diversos problemas surgidos tras la aplicación de los servicios web de primera generación.

Una versión más reciente la constituyen los servicios web RESTful. Se basan en los siguientes principios:

Utiliza un protocolo cliente/servidor sin estado: Cada mensaje HTTP contiene toda la información necesaria para comprender la petición, por lo tanto, ni el cliente ni el servidor necesitan recordar el estado de las comunicaciones entre mensajes.

Poseen un conjunto de operaciones bien definidas que se aplican a todos los recursos de información, entre las más importantes están: GET (obtener recursos), PUT (introducir un nuevo recurso), POST (modificar un recurso) y DELETE (borrar un recurso).

Tiene una sintaxis universal para identificar los recursos. En un sistema REST, cada recurso es direccionable únicamente a través de su URI.

Los distintos recursos se pueden interrelacionar y referenciar entre si mediante sus identificadores globales. La representación de los recursos en un sistema REST son típicamente XML o JSON. Como resultado de esto, es posible navegar de un recurso REST a muchos otros, simplemente siguiendo enlaces sin requerir el uso de registros u otra infraestructura adicional.

Un servicio web que cumple las restricciones de la arquitectura REST se denomina RESTful API (Mehta, 2014) y se define incluyendo los siguientes aspectos:

- Se basa en una URL, como por ejemplo <http://example.org/essim/webservices>.
- El tipo de datos debe ser adecuado para Internet, tal como JSON (*JavaScript Object Notation*) o XML.
- Debe utilizar métodos estándar HTTP (., GET, PUT, POST, or DELETE)

Los servicios web REST se pueden modelar utilizando algunas de las especificaciones que actualmente tratan de imponerse como estándares, como son RAML y OpenAPI.

4.3 Representación de los vocabularios con RDF

Antes de abordar las pruebas realizadas, era necesaria la formalización de los vocabularios propuestos en el capítulo 3 mediante su representación utilizando tecnologías de la web semántica.

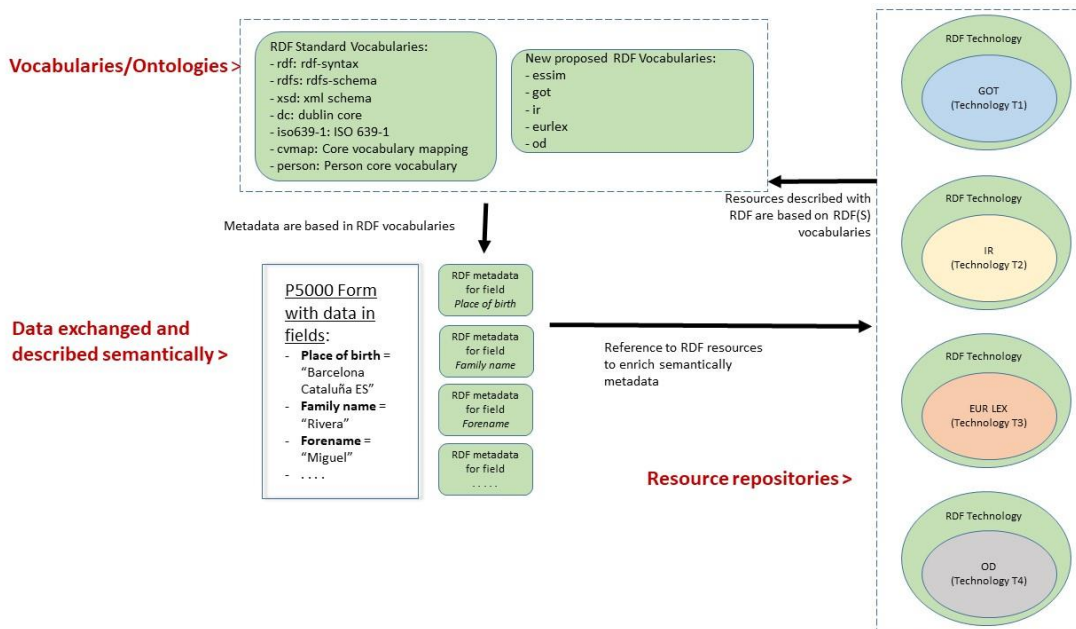


Figura 30. Esquema general de descripción de datos utilizando metadatos ESSIM

La idea general que subyace en la propuesta ESSIM que se presenta en esta tesis es que la tecnología RDF permita la interoperabilidad entre los diferentes recursos existentes basados en los conceptos representados en dichos vocabularios (Delgado, 2018b). En la figura 30 se muestra esta idea. Se trata de describir datos o campos incluidos en formularios o documentos oficiales intercambiados entre organismos de seguridad social, utilizando metadatos. Los valores que pueden asignarse a algunos de dichos metadatos hacen referencia a recursos a los que se debe acceder utilizando diferentes tecnologías. Por ejemplo, el glosario de términos está disponible en formato PDF (tecnología T1, en la figura 30), mientras que la información de las instituciones está disponible en una base de datos en un servidor de la Comisión Europea, a la que hay que acceder desde una interfaz de usuario determinada (tecnología 2 en la figura 30).

Mediante la tecnología RDF se consigue describir el contenido de estos repositorios y exponerlos a sus usuarios de una manera uniforme, en un formato único. Para ello es necesario crear previamente el vocabulario ESSIM, así como el resto en el que éste se basa, que parecen en la figura 4.2 y fueron descritos en el capítulo 3.

En la figura 31 se muestra un diagrama conceptual en UML que representa estos vocabularios. El diagrama se ha realizado utilizando el perfil de UML conocido como “*Ontology Definition Metamodel*” (ODM), creado por OMG para representar modelos RDF (OMG, 2014) utilizando estereotipos. Para simplificar el diagrama se han usado los prefijos indicados en el apartado 3.3 para cada vocabulario, que son los siguientes:

- `essim`: <https://socialsec.github.io/voc/essim/#>
- `got`: <https://socialsec.github.io/voc/got/#>
- `ir`: <https://socialsec.github.io/voc/ir/#>
- `eurlex`: <https://socialsec.github.io/voc/eurlex/#>
- `od`: <https://socialsec.github.io/voc/od/#>

En el diagrama aparecen sombreados dos elementos utilizados de otros vocabularios ya existentes, cuyos prefijos son los siguientes:

- `rdf`: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
- `rdfs`: <http://www.w3.org/2000/01/rdf-schema#>
- `xsd`: <http://www.w3.org/2001/XMLSchema#>
- `dc`: <http://purl.org/dc/elements/1.1/>
- `iso639-1`: <http://id.loc.gov/vocabulary/iso639-1>

- cvmap: <http://data.europa.eu/core-vocabularies/>

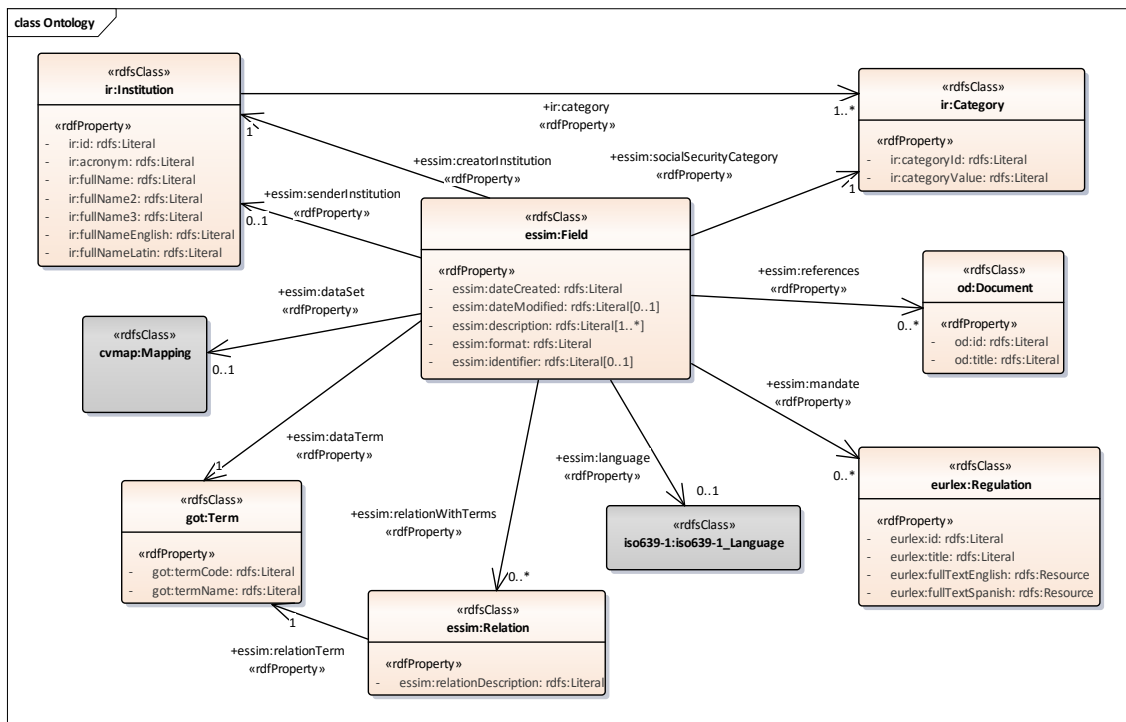


Figura 31. Diagrama conceptual de los vocabularios

Las descripciones de los vocabularios propuestos se han publicado en las URL indicadas por los prefijos, donde también están disponibles en formato RDF. En las tablas 33 a 37 se muestra el contenido de los archivos dónde están almacenados, utilizando la Serialización RDF-Turtle.

```

@prefix cvmap: <http://data.europa.eu/core-vocabularies/> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix iso639-1: <http://id.loc.gov/vocabulary/iso639-1/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix essim: <https://socialsec.github.io/voc/essim/#> .
@prefix eurlex: <https://socialsec.github.io/voc/eurlex/#> .
@prefix got: <https://socialsec.github.io/voc/got/#> .
@prefix ir: <https://socialsec.github.io/voc/ir/#> .
@prefix od: <https://socialsec.github.io/voc/od/#> .

*****
*****ESSM VOCABULARY*****
*****

# Classes
essim:Field
  a rdfs:Class ;
  rdfs:label "Data field"@en ;
  rdfs:comment "Data field in a Social Security form"@en .
essim:Relation
  a rdfs:Class ;
  rdfs:label "Relation"@en ;

```

```

    rdfs:comment "Relation between a data field and a term in the Glossary of
Terms"@en .

# Properties (domain: Field)
essim:creatorInstitution
  a rdf:Property ;
  rdfs:label "Creator Institution"@en ;
  rdfs:domain essim:Field ;
  rdfs:range ir:Institution ;
  rdfs:subPropertyOf dc:creator .

essim:dataSet
  a rdf:Property ;
  rdfs:label "Data Set"@en ;
  rdfs:domain essim:Field ;
  rdfs:range cvmap:Mapping ;
  rdfs:subPropertyOf dc:subject .

essim:dataTerm
  a rdf:Property ;
  rdfs:label "Data Term"@en ;
  rdfs:domain essim:Field ;
  rdfs:range got:Term ;
  rdfs:subPropertyOf dc:date .

essim:dateCreated
  a rdf:Property ;
  rdfs:label "Date Created"@en ;
  rdfs:domain essim:Field ;
  rdfs:range rdfs:Literal ;
  rdfs:subPropertyOf dc:date .

essim:dateModified
  a rdf:Property ;
  rdfs:label "Date Modified"@en ;
  rdfs:domain essim:Field ;
  rdfs:range rdfs:Literal ;
  rdfs:subPropertyOf dc:date .

essim:description
  a rdf:Property ;
  rdfs:label "Description"@en ;
  rdfs:domain essim:Field ;
  rdfs:range rdfs:Literal ;
  rdfs:subPropertyOf dc:description .

essim:format
  a rdf:Property ;
  rdfs:label "Format"@en ;
  rdfs:domain essim:Field ;
  rdfs:range rdfs:Literal ;
  rdfs:subPropertyOf dc:format .

essim:identifier
  a rdf:Property ;
  rdfs:label "Identifier"@en ;
  rdfs:domain essim:Field ;
  rdfs:range rdfs:Literal ;
  rdfs:subPropertyOf dc:identifier .

essim:language
  a rdf:Property ;
  rdfs:label "Language"@en ;
  rdfs:domain essim:Field ;
  rdfs:range iso639-1:iso639-1_Language ;
  rdfs:subPropertyOf dc:language .

essim:mandate

```

```

a rdf:Property ;
rdfs:label "Mandate"@en ;
rdfs:domain essim:Field ;
rdfs:range eurlex:Regulation .

essim:references
a rdf:Property ;
rdfs:label "References"@en ;
rdfs:domain essim:Field ;
rdfs:range od:Document .

essim:relationWithTerms
a rdf:Property ;
rdfs:label "Relation with Terms"@en ;
rdfs:domain essim:Field ;
rdfs:range essim:Relation ;
rdfs:subPropertyOf dc:relation .

essim:senderInstitution
a rdf:Property ;
rdfs:label "Sender Institution"@en ;
rdfs:domain essim:Field ;
rdfs:range ir:Institution ;
rdfs:subPropertyOf dc:publisher .

essim:socialSecurityCategory
a rdf:Property ;
rdfs:label "Social Security Category"@en ;
rdfs:domain essim:Field ;
rdfs:range ir:Category ;
rdfs:subPropertyOf dc:coverage .

# Properties (domain: Relation)
essim:relationTerm
a rdf:Property ;
rdfs:label "Relation Term"@en ;
rdfs:domain essim:Relation ;
rdfs:range got:Term .

essim:relationDescription
a rdf:Property ;
rdfs:label "Relation Description"@en ;
rdfs:domain essim:Relation ;
rdfs:range rdfs:Literal .

```

Tabla 33. Vocabulario ESSIM en formato RDF-Turtle

```

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix got: <https://socialsec.github.io/voc/got/#> .

*****
*****GLOSSARY OF TERMS VOCABULARY*****
*****

# Classes
got:Term
a rdfs:Class ;
rdfs:label "Term"@en ;
rdfs:comment "Term included in the Glossary of Terms"@en .

# Properties

```



```

got:termCode
  a rdf:Property ;
  rdfs:domain got:Term ;
  rdfs:range rdfs:Literal ;
  rdfs:comment "Match with column4 of Term description in GOT"@en .

got:termName
  a rdf:Property ;
  rdfs:domain got:Term ;
  rdfs:range rdfs:Literal ;
  rdfs:comment "Match with column3 of Term description in GOT"@en .

# Resources (Individuals, Instances)
got:IC001
  a got:Term ;
  got:termCode "IC001" ;
  got:termName "Surname"@en .

got:IC004
  a got:Term ;
  got:termCode "IC004" ;
  got:termName "Previous names"@en .

got:IC007
  a got:Term ;
  got:termCode "IC007" ;
  got:termName "Place of birth"@en .

got:IC009
  a got:Term ;
  got:termCode "IC009" ;
  got:termName "Nationality"@en .

```

Tabla 34. Vocabulario GOT en formato RDF-Turtle y ejemplos de instancias

```

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix ir: <https://socialsec.github.io/voc/ir/#> .

#*****
#*****INSTITUTION REPOSITORY VOCABULARY*****
#*****

# Classes
ir:Institution
  a rdfs:Class ;
  rdfs:label "Institution"@en ;
  rdfs:comment "Institution included in the Institution Repository."@en .

ir:Category
  a rdfs:Class ;
  rdfs:label "Category"@en ;
  rdfs:comment "Category of Social Security (CSS)."@en .

# Properties
ir:id
  a rdf:Property ;
  rdfs:domain ir:Institution ;
  rdfs:range rdfs:Literal .

ir:acronym
  a rdf:Property ;
  rdfs:domain ir:Institution ;

```

```

rdfs:range rdfs:Literal .

ir:fullName
  a rdf:Property ;
  rdfs:domain ir:Institution ;
  rdfs:range rdfs:Literal .

ir:fullName2
  a rdf:Property ;
  rdfs:domain ir:Institution ;
  rdfs:range rdfs:Literal .

ir:fullName3
  a rdf:Property ;
  rdfs:domain ir:Institution ;
  rdfs:range rdfs:Literal .

ir:fullNameEnglish
  a rdf:Property ;
  rdfs:domain ir:Institution ;
  rdfs:range rdfs:Literal .

ir:fullNameLatin
  a rdf:Property ;
  rdfs:domain ir:Institution ;
  rdfs:range rdfs:Literal .

ir:category
  a rdf:Property ;
  rdfs:domain ir:Institution ;
  rdfs:range ir:Category .

ir:categoryId
  a rdf:Property ;
  rdfs:domain ir:Category ;
  rdfs:range rdfs:Literal .

ir:categoryValue
  a rdf:Property ;
  rdfs:domain ir:Category ;
  rdfs:range rdfs:Literal .

# Resources
ir:2800
  a ir:Institution ;
  ir:id "2800" ;
  ir:acronym "INSS MADRID" ;
  ir:fullName "INSTITUTO NACIONAL DE LA SEGURIDAD SOCIAL DE MADRID DIRECCION
PROVINCIAL" ;
  ir:fullNameEnglish "National Institute of Social Security. Provincial Head Office
of Madrid" ;
  ir:category ir:CSS21001, ir:CSS21003, ir:CSS21004, ir:CSS21008, ir:CSS21009,
ir:CSS21011, ir:CSS21014 .

ir:3656
  a ir:Institution ;
  ir:id "3656" ;
  ir:acronym "ISM VILLAGARCIA" ;
  ir:fullName "INSTITUTO SOCIAL DE LA MARINA DIRECCION PROVINCIAL DE VILLAGARCIA" ;
  ir:fullNameEnglish "Social Institute for Mariners. Provincial Head Office of
Villagarcía" ;
  ir:category ir:CSS21001, ir:CSS21003, ir:CSS21004, ir:CSS21008, ir:CSS21009,
ir:CSS21011, ir:CSS21012 .

ir:6000
  a ir:Institution ;
  ir:id "6000" ;
  ir:acronym "INSS ORG. ENLACE" ;

```

```

ir:fullName "INSTITUTO NACIONAL DE LA SEGURIDAD SOCIAL DIRECCION GENERAL"@es ;
ir:fullNameEnglish "National Institute of Social Security. General
Directorate"@en ;
ir:category ir:CSS21001, ir:CSS21002, ir:CSS21003, ir:CSS21004, ir:CSS21006,
ir:CSS21008, ir:CSS21009, ir:CSS21011, ir:CSS21014 .

ir:CSS21001
a ir:Category ;
ir:categoryId "21001" ;
ir:categoryName "Sickness benefits"@en .

ir:CSS21002
a ir:Category ;
ir:categoryId "21002" ;
ir:categoryName "Long-term care benefits"@en .

ir:CSS21003
a ir:Category ;
ir:categoryId "21003" ;
ir:categoryName "Maternity, paternity benefits"@en .

ir:CSS21004
a ir:Category ;
ir:categoryId "21004" ;
ir:categoryName "Invalidity - state invalidity pension scheme"@en .

ir:CSS21006
a ir:Category ;
ir:categoryId "21006" ;
ir:categoryName "Old Age pensions - state pension scheme"@en .

```

Tabla 35. Vocabulario IR en formato RDF-Turtle y ejemplos de instancias

```

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix eurlex: <https://socialsec.github.io/voc/eurlex/#> .
#*****
#*****EUR-LEX VOCABULARY*****
#*****

# Classes
eurlex:Regulation
a rdfs:Class ;
rdfs:label "Regulation"@en ;
rdfs:comment "Regulation about Social Security"@en .

# Properties
eurlex:id
a rdf:Property ;
rdfs:domain eurlex:Regulation ;
rdfs:range rdfs:Literal .

dc:title
a rdf:Property ;
rdfs:domain ir:Regulation ;
rdfs:range rdfs:Literal .

eurlex:fullTextEnglish
a rdf:Property ;
rdfs:domain ir:Regulation ;
rdfs:range rdfs:Resource .

```

```

eurlex:fullTexSpanish
  a rdf:Property ;
  rdfs:domain ir:Regulation ;
  rdfs:range rdfs:Resource .

# Resources
eurlex:32004R0883
  a eurlex:Regulation ;
  eurlex:id "32004R0883" ;
  dc:title "Regulation (EC) No 883/2004 of the European Parliament and of the
Council of 29 April 2004 on the coordination of social security systems (Text with
relevance for the EEA and for Switzerland)" ;
  eurlex:fulltextenglish <http://eur-lex.europa.eu/legal-
content/EN/TXT/PDF/?uri=CELEX:32004R0883&qid=1437065889270&from=EN> ;
  eurlex:fulltextspanish <http://eur-lex.europa.eu/legal-
content/ES/TXT/PDF/?uri=CELEX:32004R0883&qid=1437065889270&from=EN> .

eurlex:32009R0987
  a eurlex:Regulation ;
  eurlex:id "32009R0987" ;
  dc:title "Regulation (EC) No 987/2009 of the European Parliament and of the
Council of 16 September 2009 laying down the procedure for implementing Regulation
(EC) No 883/2004 on the coordination of social security systems (Text with
relevance for the EEA and for Switzerland)" ;
  eurlex:fulltextenglish <http://eur-lex.europa.eu/legal-
content/EN/TXT/PDF/?uri=CELEX:32009R0987&qid=1437066151142&from=EN> ;
  eurlex:fulltextspanish <http://eur-lex.europa.eu/legal-
content/ES/TXT/PDF/?uri=CELEX:32009R0987&qid=1437066151142&from=EN> .

```

Tabla 36. Vocabulario EUR-LEX en formato RDF-Turtle y ejemplos de instancias

```

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix od: <https://socialsec.github.io/voc/od/#> .

*****
*****OFFICIAL DOCUMENTS VOCABULARY*****
*****

# Classes
od:Document
  a rdfs:Class ;
  rdfs:label "Document"@en ;
  rdfs:comment "Official document for social security issues."@en .

od:Form
  a rdfs:Class ;
  rdfs:subClassOf od:Document ;
  rdfs:label "Form"@en ;
  rdfs:comment "A form document."@en .

od:P-Form
  a rdfs:Class ;
  rdfs:subClassOf od:Form ;
  rdfs:label "P-Form"@en ;
  rdfs:comment "A form related to pensions."@en .

# Properties
od:id
  a rdf:Property ;
  rdfs:domain od:Document ;
  rdfs:range rdfs:Literal .

od:title

```

```

a rdf:Property ;
rdfs:domain od:Document ;
rdfs:range rdfs:Literal .

od:template
a rdf:Property ;
rdfs:domain od:Document ;
rdfs:range rdfs:Resource .

# Resources
od:E101
a od:Form ;
od:id "E101" ;
od:title "E101 Certificate concerning the legislation applicable"@es ;
od:template <http://od.org/templates/E101.docx> .

od:E106
a od:Form ;
od:id "E106" ;
od:title "E106 certificate of entitlement to sickness and maternity insurance
benefits in kind for persons residing in a country other than the competent
country"@es ;
od:template <http://od.org/templates/E106.docx> .

od:P5000
a od:P-Form ;
od:id "P5000" ;
od:title "P5000 Insurance/residence periods"@en ;
od:template <http://od.org/templates/P5000.docx> .

od:P9000
a od:P-Form ;
od:id "P9000" ;
od:title "P9000 Reply to request for additional information"@es ;
od:template <http://od.org/templates/P9000.docx> .

```

Tabla 37. Vocabulario OD en formato RDF-Turtle y ejemplos de instancias

Además, estos vocabularios, así como instancias de los mismos, se han instalado en un SPARQL Endpoint disponible en la web del Grupo de Investigación TIFYC de la Universidad de Alcalá, ubicado en: <http://tifyc-pmi.cc.uah.es/fuseki/>. Se han publicado en forma de *dataset* con el nombre de *socialsecurity*. En la figura 32 se muestra la creación del *dataset* mediante la carga de los 5 archivos RDF que contienen los vocabularios con Serialización Turtle (ttl).

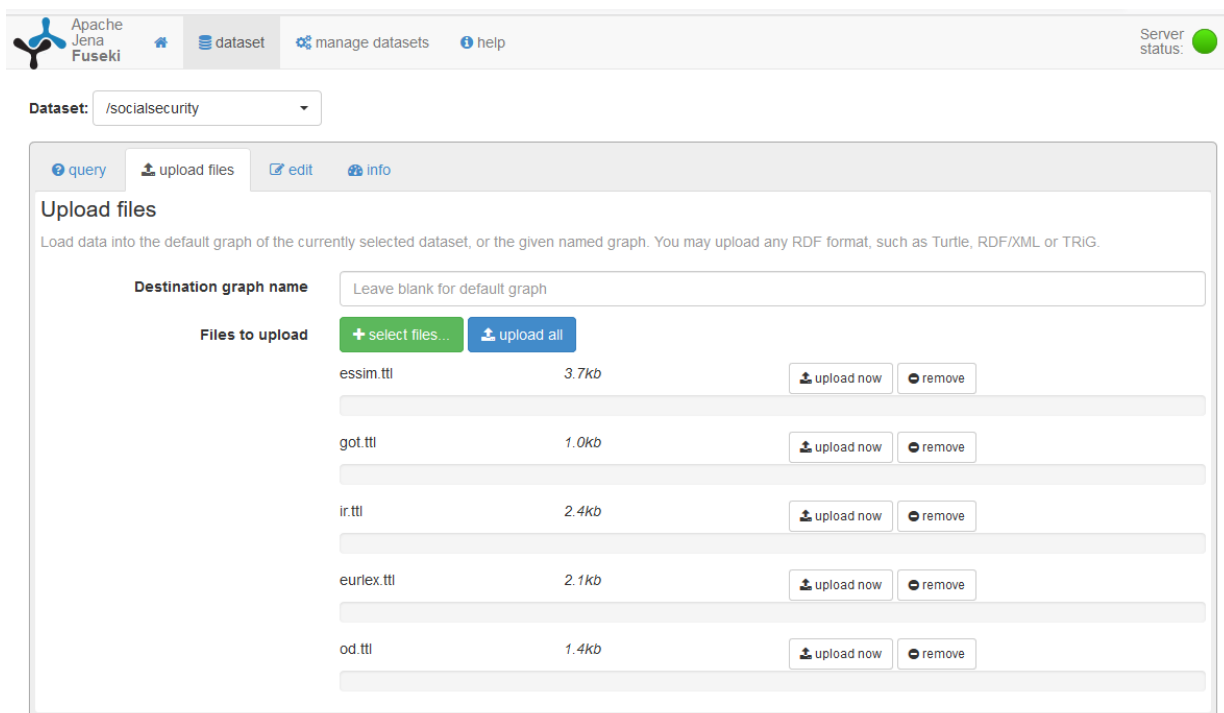


Figura 32. Creación de un *dataset* en el SPARQL Endpoint

De esta forma se puede utilizar el editor disponible para ejecutar consultas SPARQL sobre los vocabularios, además de integrarse con otros SPARQL Endpoints instalados en el contexto de los datos enlazados de la Web, e integrarse con otros data sets.

Por ejemplo, una sencilla consulta sobre el nombre en inglés de la institución cuyo acrónimo es “INSS Madrid” sería la siguiente:

```
PREFIX ir: <https://socialsec.github.io/voc/ir/#>
SELECT ?name
WHERE {
  ?institution ir:acronym "INSS MADRID" .
  ?institution ir:fullNameEnglish ?name
}
```

En la figura 33 se muestra el resultado de esta consulta.

The screenshot shows a web interface for running SPARQL queries. At the top, there are navigation links: 'query', 'upload files', 'edit', and 'info'. Below this is the title 'SPARQL query' and a brief instruction: 'To try out some SPARQL queries against the selected dataset, enter your query here.' There are two buttons for 'EXAMPLE QUERIES': 'Selection of triples' and 'Selection of classes'. A 'PREFIXES' section contains buttons for 'rdf', 'rdfs', 'owl', 'xsd', 'ir', and a plus sign. The 'SPARQL ENDPOINT' is set to 'http://tifyc-pmi.cc.uah.es/fuseki/socialsecurity/query'. The 'CONTENT TYPE (SELECT)' is 'JSON' and the 'CONTENT TYPE (GRAPH)' is 'Turtle'. The query editor contains the following code:

```

1
2 PREFIX ir: <https://socialsec.github.io/voc/ir/#>
3
4 SELECT ?name
5 WHERE {
6   ?institution ir:acronym "INSS MADRID" .
7   ?institution ir:fullNameEnglish ?name
8 }
9 |

```

Below the query editor, the 'QUERY RESULTS' section shows 'Table' selected over 'Raw Response'. It indicates 'Showing 1 to 1 of 1 entries'. A search box and a 'Show 50 entries' dropdown are present. The results table has one entry:

name
"National Institute of Social Security. Provincial Head Office of Madrid"

Figura 33. Ejemplo de consulta SPARQL

4.4 Prueba 1: desarrollo de una aplicación de creación de metadatos

Para describir un campo de cualquier formulario de un documento oficial utilizando los metadatos ESSIM es suficiente con disponer de un procesador de texto y conocer la notación RDF, así como el contenido de los vocabularios relacionados con ESSIM. Sin embargo, para permitir que cualquier persona pudiera realizar dicha descripción sin conocer RDF ni el detalle de los vocabularios, se ha creado un prototipo en forma de aplicación web desarrollada en Java que permite introducir los valores de los metadatos y generar un archivo RDF verificado con la estructura establecida por la propuesta ESSIM.

Para procesar los vocabularios RDF se ha utilizado el *framework* Jena (Apache, 2018). Se trata de una librería Java que permite leer grafos RDF contenidos en archivos y almacenarlos en memoria en forma de un modelo con una estructura de objetos, lo que facilita su procesamiento y permite hacer consultas sobre ellos con el lenguaje SPARQL, y también puede leer archivos con reglas y realizar inferencias sobre el modelo de conocimiento en memoria para ampliarlo generando nuevo conocimiento.

En la figura 34 se muestra el aspecto de la aplicación, desarrollada como una página JSP.

The screenshot shows a web form for creating metadata. The form is organized into two main columns. The left column contains the following fields: 'Creator Institution (IR)', 'Data Set (Core Vocabulary)' (with a dropdown menu showing 'Country of Birth'), 'Data Term (GOT)', 'Date Created', 'Date Modified', 'Description', and 'Format'. The right column contains: 'Identifier', 'Language (ISO 639-1)', 'Mandate (Euro-Lex)', 'References', 'Relation with Terms (GOT)' (with 'Term' and 'Relation description' sub-fields), 'Sender Institution (IR)', 'Social Security Category (IR)', and 'Path save files in your PC'. A 'Submit' button is located at the bottom right.

Figura 34. Aplicación web de creación de metadatos

En aquellos campos (metadatos) asociados a valores de un vocabulario controlado, lo que se ofrece al usuario es una lista desplegable de valores posibles, en un formato legible, sin que el usuario tenga que conocer su representación interna. Por ejemplo, en el caso del primero de los metadatos: *creatorInstitution*, si está buscando una institución en cuyo nombre se incluye el término “Madrid”, aparece una lista de organismos del repositorio de instituciones con dicho término, y el usuario podría elegir uno de ellos, como el “*National Institute of Social Security. Provincial Head Office of Madrid*” (figura 35). Sin embargo, internamente en formato RDF, el valor que se asignaría a la propiedad *creatorInstitution* sería la IRI del recurso correspondiente a esa institución en el vocabulario IR, que sería <https://socialsec.github.io/voc/ir/#2800>.



Figura 35. Selección de una institución para el metadato *creatorInstitution*

En las siguientes figuras se irán mostrando el resto de elementos ESSIM que habría que cumplimentar para describir el campo *Surname* del documento oficial utilizado en el capítulo 3 como ejemplo (figura 20). En dicho capítulo se proponía un valor concreto para cada elemento o metadato ESSIM usado para describirlo, serán esos mismos valores los que se utilizarán en este apartado, como caso de prueba de funcionamiento de la aplicación.

El siguiente elemento ESSIM que hay que cumplimentar es *dataSet*. En el ejemplo del capítulo 3 se proponía utilizar el término *familyName* del *Core Vocabulary Person*. En la figura 36 puede verse que el usuario puede elegir dicho valor de entre la lista que se despliega correspondiente a todos los términos del vocabulario *Person*. Por simplicidad, el usuario sólo elige el término, pero internamente la aplicación le asocia la propiedad *mappingRelation* con valor por defecto *hasCloseMatch*. Si el usuario tiene conocimientos de RDF, posteriormente cuando se genere el archivo RDF correspondiente, podrá editarlo y modificar este valor. En la primera versión de prototipo sólo se ha considerado el vocabulario *Person*.

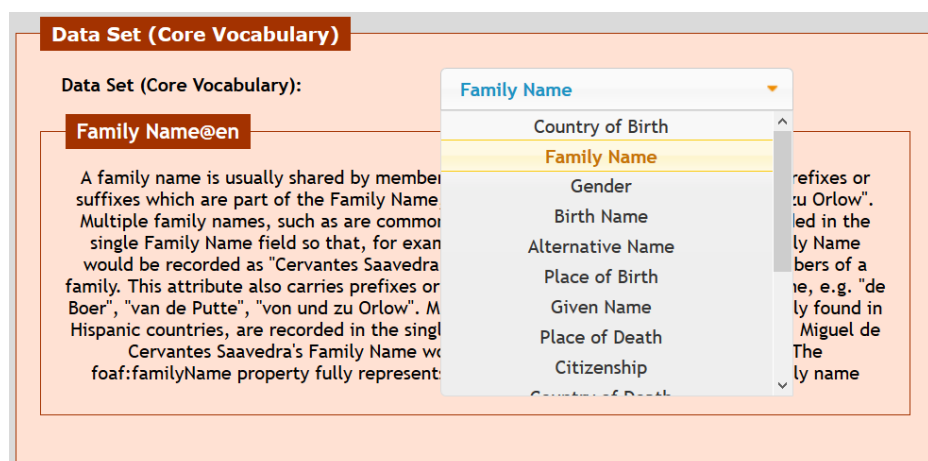


Figura 36. Selección de un valor para el metadato *dataSet*

El siguiente elemento es *dataTerm*. En este caso se ofrece una lista con los términos del GOT que contienen la palabra *“name”* y el usuario elegiría el primero de la lista (*Surname*), como se proponía en el capítulo 3 para este metadato (figura 37).

Data Term (GOT)

Data Term (GOT)*:
name

[IC001] : [Surname] : [GENERIC] : State the official family name of a client.

[IC001] : [Surname] : [BEL] : Form E6XX not used Official surname given to a person

[IC001] : [Surname] : [D] : HC: Official family surname as it appears in valid personal identity papers". When using electronic data exchange, form E110 is no longer needed. E123 is exclusively a form relating to statutory accident insurance. CB: Corresponds to the GENERIC Glossary of Terms UN: Official family name (surname) of a person.

[IC001] : [Surname] : [IRL] : State the clients family name. Note: National applications format consists of 20 characters (free text)

[IC001] : [Surname] : [LUX] : OAP (CPACI) : See GOT Framework if not married, same as IC002 otherwise, married name

[IC001] : [Surname] : [LUX] : OAP (CPEP) : See GOT Framework if not married, same as IC002 otherwise married name

[IC001] : [Surname] : [LUX] : OAP (E500) : See GOT Framework for a married woman: married name otherwise, maiden name (same as name at birth)

[IC001] : [Surname] : [LUX] : UN: See GOT Framework maiden name in the former surnames box (E 301) if married

[IC001] : [Surname] : [LUX] : CB: See GOT Framework

[IC001] : [Surname] : [NED] : See GOT Framework

[IC001] : [Surname] : [POR] : Full name as it appears on the identity card or passport.

Figura 37. Selección de un valor para el metadato *dataTerm*

Los dos siguientes elementos son *dateCreated* y *dateModified*. En ambos casos se ofrece al usuario un calendario para elegir las fechas (figura 38).

Date Created

Date Created*:
2014-05-30

Date Modified

Date Modified:

May 2014

Su	Mo	Tu	We	Th	Fr	Sa
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17

Figura 38. Selección de un valor para el metadato *dateCreated*

Los elementos *description*, *format* e *identifier* se pueden rellenar con texto libre. En la figura 39 se muestra con los valores recomendados para este ejemplo en el capítulo 3.

Description

Description*:
Official name of a person

Format

Format*:
Alphabetic text

Identifier

Identifier:
surname

Figura 39. Introducción de valores para los metadatos

En el caso del elemento *language*, aparece la lista de nombres de los lenguajes incluidos en el estándar ISO 639-1 (figura 40).

The image shows a web form with a section titled "Language (ISO 639-1)". Inside this section, there is a label "Language (ISO 639-1):" followed by a dropdown menu. The dropdown menu is open, showing "English : [English]" as the selected option. Below the dropdown, there is a list of other language options: "Esperanto : [Esperanto]" and "Estonian : [eesti, eesti keel]".

Figura 40. Selección de un valor para el metadato *language*

Los metadatos *mandate* y *references* se deben cumplimentar con valores de los vocabularios EurLex y *Official Documents*. En ambos casos sólo se ofrece al usuario introducir un valor (figura 41). Si se necesitaran más se podrían introducir editando el archivo RDF final.

The image shows two form sections. The top section is titled "Mandate (Euro-Lex)" and contains a label "Mandate (Euro-Lex):" followed by a text input field containing the text "Regulation (EC) No 883/2004 of the European Parliament and of the Council". The bottom section is titled "References" and contains a label "References:" followed by a text input field containing the text "E101: Certificate concerning the legislation applicable".

Figura 41. Valores para los metadatos *mandate* y *references*

El elemento *relationWithTerms* ofrece al usuario la posibilidad de elegir un término del GOT, y de introducir un texto en el campo "*Relation Description*" (Figura 42).

The image shows a form section titled "Relation with Terms (GOT)". It contains a label "Term:" followed by a text input field containing the text "Previous". Below this, there is a label "Relatio" followed by a text input field containing the text "[IC004] : [Previous names] : [GENERIC] : State the name(s) that a client has been known by previously." and "[IC004] : [Previous names] : [UK] : The surname(s) that a person has been known by previously. The Surname value has officially substituted the previous names of the pers".

Figura 42. Valores para el metadato *relationWithTerms*

El metadato *senderInstitution* se cumplimenta eligiendo un valor del *Institution Repository* (figura 43). En este caso la Dirección General del INSS.

Sender Institution (IR)

Sender Institution (IR)*:

General Directorate
6017 : State Public Employment Service (General Directorate)
6000 : National Institute of Social Security. General Directorate

Figura 43. Selección de un valor para el metadato *senderInstitution*

Por último, el metadato opcional *socialSecurityCategory* permite asignar un ámbito de actuación de la Seguridad Social, de los recogidos en el repositorio de Instituciones (figura 44).

Social Security Category (IR)

Social Security Category (IR)*:

Maternity, paternity

Figura 44. Valor para el metadato *socialSecurityCategory*

Una vez cumplimentados los campos, se puede pulsar el botón “Submit” para obtener el archivo RDF correspondiente a los valores introducidos. En la tabla 38 se muestra un ejemplo de código que corresponde al ejemplo utilizado en este caso, es decir la descripción del campo *Surname* incluido en el documento E106 utilizado en el capítulo 3, con los valores para los metadatos ESSIM que se han ido mostrando en las figuras anteriores.

```
<?xml version="1.0" encoding="utf-8" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:essim="https://socialsec.github.io/voc/essim/#"
  xmlns:cvmmap="http://data.europa.eu/cv/">

  <rdf:Description>
    <rdf:type rdf:resource="https://socialsec.github.io/voc/essim/#Field"/>
    <essim:creatorInstitution
rdf:resource="https://socialsec.github.io/voc/ir/#2800"/>
    <essim:dataSet>
      <rdf:Description>
        <rdf:type rdf:resource="http://data.europa.eu/cv/Mapping"/>
        <cvmmap:coreVocURI
rdf:resource="https://joinup.ec.europa.eu/sites/default/files/distribution/2013-
10/person-v1.00.rdf#familyName"/>
        <cvmmap:mappingRelation
rdf:resource="http://data.europa.eu/cv/hasCloseMatch"/>
      </rdf:Description>
    </essim:dataSet>
    <essim:dataTerm rdf:resource="https://socialsec.github.io/voc/got/#IC001"/>
    <essim:dateCreated rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2014-
05-30</essim:dateCreated>
    <essim:dateModified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2015-
02-24</essim:dateModified>
    <essim:description xml:lang="en">Official name of a person</essim:description>
    <essim:format xml:lang="en">Alphabetic text</essim:format>
    <essim:identifier>surname</essim:identifier>
    <essim:language rdf:resource="http://id.loc.gov/vocabulary/iso639-1/en"/>
    <essim:mandate
```

```

rdf:resource="https://socialsec.github.io/voc/eurlex/#32004R0883"/>
  <essim:references rdf:resource="https://socialsec.github.io/voc/od/#E101"/>
  <essim:relationWithTerms>
    <essim:Relation>
      <essim:relationTerm
rdf:resource="https://socialsec.github.io/voc/got/#IC004"/>
        <essim:relationDescription xml:lang="en">The Surname value has officially
substituted the previous names of the person</essim:relationDescription>
      </essim:Relation>
    </essim:relationWithTerms>
    <essim:senderInstitution
rdf:resource="https://socialsec.github.io/voc/ir/#6000"/>
    <essim:socialSecurityCategory
rdf:resource="https://socialsec.github.io/voc/ir/#21003"/>
    </rdf:Description>
</rdf:RDF>

```

Tabla 38. Descripción del campo *Surname* utilizando los elementos de ESSIM (formato RDF/XML)

Un grafo RDF como el de la tabla 38 también puede codificarse o serializarse utilizando otras posibles notaciones, como es el caso de Turtle (tabla 39) o JSON-LD (tabla 40).

```

@prefix cvmap: <http://data.europa.eu/cv/> .
@prefix essim: <https://socialsec.github.io/voc/essim/#> .
@prefix eurlex: <https://socialsec.github.io/voc/eurlex/#> .
@prefix got: <https://socialsec.github.io/voc/got/#> .
@prefix ir: <https://socialsec.github.io/voc/ir/#> .
@prefix iso639-1: <http://id.loc.gov/vocabulary/iso639-1/> .
@prefix od: <https://socialsec.github.io/voc/od/#> .
@prefix person: <https://joinup.ec.europa.eu/sites/default/files/distribution/2013-10/person-v1.00.rdf#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

[]
  a essim:Field ;
  essim:creatorInstitution ir:2800;
  essim:dataSet [a cvmap:Mapping;
                cvmap:coreVocURI person:familyName;
                cvmap:mappingRelation cvmap:hasCloseMatch];
  essim:dataTerm got:IC001;
  essim:dateCreated "2014-05-30"^^xsd:date;
  essim:dateModified "2015-02-24"^^xsd:date;
  essim:description "Official name of a person"@en;
  essim:format "Alphabetic text"@en;
  essim:identifier "surname"^^xsd:string;
  essim:language iso639-1:en;
  essim:mandate eurlex:32004R0883;
  essim:references od:E101;
  essim:relationWithTerms [a essim:Relation;
                            essim:relationTerm got:IC004;
                            essim:relationDescription "The Surname value has
officially substituted the previous names of the person"@en];
  essim:senderInstitution ir:6000;
  essim:socialSecurityCategory ir:CSS21003.

```

Tabla 39. Descripción del campo *Surname* con formato Turtle

```

{
  "@context": {
    "cvmap": "http://data.europa.eu/cv/",
    "essim": "https://socialsec.github.io/voc/essim/#",
    "eurlex": "https://socialsec.github.io/voc/eurlex/#",
    "got": "https://socialsec.github.io/voc/got/#",
    "ir": "https://socialsec.github.io/voc/ir/#",
    "iso639-1": "http://id.loc.gov/vocabulary/iso639-1/",
    "od": "https://socialsec.github.io/voc/od/#",
    "person": "https://joinup.ec.europa.eu/sites/default/files/distribution/2013-10/person-v1.00.rdf#",
    "xsd": "http://www.w3.org/2001/XMLSchema#"
  },
  "@type": "essim:Field",
  "essim:creatorInstitution": {"@id": "ir:2800"},
  "essim:dataSet": {
    "@type": "cvmap:Mapping",
    "cvmap:coreVocURI": {"@id": "person:familyName"},
    "cvmap:mappingRelation": {"@id": "cvmap:hasCloseMatch"}
  },
  "essim:dataTerm": {"@id": "got:IC001"},
  "essim:dateCreated": {
    "@value": "2014-05-30",
    "@type": "http://www.w3.org/2001/XMLSchema#date"
  },
  "essim:dateModified": {
    "@value": "2015-02-24",
    "@type": "http://www.w3.org/2001/XMLSchema#date"
  },
  "essim:description": {
    "@value": "Oficial name of a person",
    "@language": "en"
  },
  "essim:format": {
    "@value": "Alphabetic text",
    "@language": "en"
  },
  "essim:identifier": "surname",
  "essim:language": {"@id": "iso639-1:en"},
  "essim:mandate": {"@id": "eurlex:32004R0883"},
  "essim:references": {"@id": "od:E101"},
  "essim:relationWithTerms": {
    "@type": "essim:Relation",
    "essim:relationTerm": {"@id": "got:IC004"},
    "essim:relationDescription": {
      "@value": "The Surname value has officially substituted the previous names of the person",
      "@language": "en"
    }
  },
  "essim:senderInstitution": {"@id": "ir:6000"},
  "essim:socialSecurityCategory": {"@id": "ir:CSS21003"}
}

```

Tabla 40. Descripción del campo *Surname* con formato JSON-LD

Una vez obtenido el archivo con los metadatos que describen el campo *Surname*, éste debería enviarse junto al documento que contiene dicho campo (E102). Si el documento tuviera un formato HTML, entonces sería posible integrar la descripción con metadatos en el propio documento, utilizando alguna de las alternativas existentes para ello: RDFa, Microdata o JSON-LD.

En la tabla 41 se muestra un posible código HTML del documento E102 con los metadatos integrados utilizando la tecnología JSON-LD (W3C, 2018b). En este caso se incluye la descripción

entre etiquetas <SCRIPT></SCRIPT>. El código es el mismo que el de la tabla 40, pero en este caso añadiendo la línea "@id": "#surname", para indicar que los metadatos son para describir el elemento HTML del documento identificado como *Surname*. Se trata del campo que en el documento E102 aparece como "2.1 Surname(s)" (figura 20), que en HTML se podría codificar como:

```
<form>
  <label for="surname">2.1 Surname(s:</label>
  <input type="text" id="surname" value="Martínez">
</form>
```

```
<html>
...
<body>
<h1>E106 CERTIFICATE OF ENTITLEMENT TO SICKNESS AND MATERNITY INSURANCE BENEFITS IN
KIND FOR PERSONS RESIDING IN A COUNTRY OTHER THAN THE COMPETENT COUNTRY</h1>

<script type="application/ld+json">
{
  "@context": {
    "cvmap": "http://data.europa.eu/cv/",
    "essim": "https://socialsec.github.io/voc/essim/#",
    "eurlex": "https://socialsec.github.io/voc/eurlex/#",
    "got": "https://socialsec.github.io/voc/got/#",
    "ir": "https://socialsec.github.io/voc/ir/#",
    "iso639-1": "http://id.loc.gov/vocabulary/iso639-1/",
    "od": "https://socialsec.github.io/voc/od/#",
    "person": "https://joinup.ec.europa.eu/sites/default/files/distribution/2013-
10/person-v1.00.rdf#",
    "xsd": "http://www.w3.org/2001/XMLSchema#"
  },
  "@type": "essim:Field",
  "@id": "#surname",
  "essim:creatorInstitution": {"@id": "ir:2800"},
  "essim:dataSet": {
    "@type": "cvmap:Mapping",
    "cvmap:coreVocURI": {"@id": "person:familyName"},
    "cvmap:mappingRelation": {"@id": "cvmap:hasCloseMatch"}
  },
  "essim:dataTerm": {"@id": "got:IC001"},
  "essim:dateCreated": {
    "@value": "2014-05-30",
    "@type": "http://www.w3.org/2001/XMLSchema#date"
  },
  "essim:dateModified": {
    "@value": "2015-02-24",
    "@type": "http://www.w3.org/2001/XMLSchema#date"
  },
  "essim:description": {
    "@value": "Oficial name of a person",
    "@language": "en"
  },
  "essim:format": {
    "@value": "Alphabetic text",
    "@language": "en"
  },
  "essim:identifier": {"@value": "surname"},
  "essim:language": {"@id": "iso639-1:es"},
  "essim:mandate": {"@id": "eurlex:32004R0883"},
  "essim:references": {"@id": "od:E101"},
  "essim:relationWithTerms": {
    "@type": "essim:Relation",
```

```

    "essim:relationTerm": {"@id": "got:IC004"},
    "essim:relationDescription": {
      "@value": "The Surname value has officially substituted the previous
names of the person",
      "@language": "en"
    }
  },
  "essim:senderInstitution": {"@id": "ir:6000"},
  "essim:socialSecurityCategory": {"@id": "ir:CSS21003"}
}
</script>
<form>
...
  <label for="surname">2.1 Surname(s):</label>
  <input type="text" id="surname" value="Martínez">
...
</form>
</body>
</html>

```

Tabla 41. Metadatos con la descripción del campo *Surname* incluidos en el propio documento utilizando JSON-LD

Otra opción de integración sería utilizando la tecnología RDFa (W3C, 2015a). En este caso los metadatos se introducen utilizando atributos como *property*, *typeof*, *content*, *resource* y *prefix* (tabla 42).

```

<html>
<head>
</head>
<body prefix="cvmap: http://data.europa.eu/cv/
essim: https://socialsec.github.io/voc/essim/#
xsd: http://www.w3.org/2001/XMLSchema#" >

<h1>E106: CERTIFICATE OF ENTITLEMENT TO SICKNESS AND MATERNITY INSURANCE BENEFITS
IN KIND FOR PERSONS RESIDING IN A COUNTRY OTHER THAN THE COMPETENT COUNTRY</h1>

<form>

  <div about="#surname" typeof="essim:Field">
    <meta property="essim:creatorInstitution"
resource="https://socialsec.github.io/voc/ir/#2800"/>
    <div property="essim:dataSet" typeof="cvmap:Mapping">
      <meta property="cvmap:coreVocURI"
resource="https://joinup.ec.europa.eu/sites/default/files/distribution/2013-
10/person-v1.00.rdf#familyName"/>
      <meta property="cvmap:mappingRelation"
resource="http://data.europa.eu/cv/hasCloseMatch"/>
    </div>
    <meta property="essim:dataTerm"
resource="https://socialsec.github.io/voc/got/#IC001"/>
    <meta property="essim:dateCreated" content="2014-05-30" datatype="xsd:date"/>
    <meta property="essim:dateModified" content="2015-02-24" datatype="xsd:date"/>
    <meta property="essim:description" lang="en" content="Official name of a person"
/>
    <meta property="essim:format" lang="en" content="Alphabetic text" />
    <meta property="essim:identifier" content="surname" />
    <meta property="essim:language" resource="http://id.loc.gov/vocabulary/iso639-
1/es"/>
    <meta property="essim:mandate"
resource="https://socialsec.github.io/voc/eurlex/#32004R0883"/>
    <meta property="essim:references"

```



```

resource="https://socialsec.github.io/voc/od/#E101"/>
  <div property="essim:relationWithTerms" typeof="essim:Relation">
    <meta
      resource="https://socialsec.github.io/voc/got/#IC004"/>
      <meta property="essim:relationDescription" lang="en" content="The Surname
value has officially substituted the previous names of the person"/>
    </div>
    <meta
      resource="https://socialsec.github.io/voc/ir/#6000"/>
      <meta
        resource="https://socialsec.github.io/voc/ir/#CSS21003"/>

    <label for="surname">2.1 Surname(s) :</label>
    <input type="text" id="surname" value="Martínez">
  </div>
</form>

</body>
</html>

```

Tabla 42. Metadatos con la descripción del campo *Surname* utilizando la tecnología RDFa

Por último, también se pueden insertar utilizando la tecnología Microdata (W3C, 2018a), que utiliza atributos como *itemid*, *itemtype* e *itemscope* (tabla 43).

```

<html>
...
<body>

<h1>E106: CERTIFICATE OF ENTITLEMENT TO SICKNESS AND MATERNITY INSURANCE BENEFITS
IN KIND FOR PERSONS RESIDING IN A COUNTRY OTHER THAN THE COMPETENT COUNTRY</h1>

<form>

  <div
    itemtype="https://socialsec.github.io/voc/essim/#Field" itemscope
    <link
      href="https://socialsec.github.io/voc/ir/#CSS21003" />
      <link
        href="https://socialsec.github.io/voc/ir/#2800" />
        <meta itemprop="identifier" content="surname" />
        <link
          href="https://socialsec.github.io/voc/eurlex/#32004R0883" />
          <div itemprop="dataSet" itemtype="http://data.europa.eu/cv/Mapping"
itemscope>
            <link
              href="https://joinup.ec.europa.eu/sites/default/files/distribution/2013-10/person-
v1.00.rdf#familyName" />
              <link
                href="http://data.europa.eu/cv/hasCloseMatch" />
                </div>
                <link
                  href="https://socialsec.github.io/voc/got/#IC001" />
                  <meta itemprop="dateCreated" content="2014-05-30" />
                  <meta itemprop="description" content="Oficial name of a person" />
                  <meta itemprop="format" content="Alphabetic text" />
                  <link itemprop="language" href="http://id.loc.gov/vocabulary/iso639-1/es"
/>
                <div
                  itemtype="https://socialsec.github.io/voc/essim/#Relation" itemscope>
                    <link
                      href="https://socialsec.github.io/voc/got/#IC004" />

```

```

        <meta itemprop="relationDescription" content="The Surname value has
officially substituted the previous names of the person" />
    </div>
    <link                                itemprop="senderInstitution"
href="https://socialsec.github.io/voc/ir/#6000" />
    <meta itemprop="dateModified" content="2015-02-24" />
    <link                                itemprop="references"
href="https://socialsec.github.io/voc/od/#E101" />
    </div>

    <label for="surname">2.1 Surname(s):</label>
    <input type="text" id="surname" value="Martínez">

</form>

</body>
</html>

```

Tabla 43. Metadatos con la descripción del campo *Surname* utilizando la tecnología Microdata

4.5 Prueba 2: arquitectura basada en servicios web

Esta sección presenta una arquitectura simplificada para la implementación de los intercambios de datos utilizando las tecnologías de web semántica como prueba de aplicación de la idea general. La arquitectura se basa en un estudio de caso que sirvió asimismo de guía para las pruebas.

4.5.1 Estudio de caso

Se propone como caso de prueba la solicitud de una pensión por parte de un ciudadano que ha cotizado en más de un país.

El proceso (figura 45) se inicia mediante la presentación de la solicitud por el interesado. La solicitud normalmente empieza a tramitarse en la institución de residencia o institución de instrucción del afiliado. En dicha solicitud se indica que ha trabajado en distintos países, por lo que debe haber realizado cotizaciones a distintos sistemas de Seguridad Social a través de diferentes instituciones.

La institución instructora deberá comprobar que existe un acuerdo en vigor por el que las cotizaciones en otros países deban ser tenidas en cuenta para el reconocimiento de la pensión. Si la contestación es positiva, se pondrá en marcha un intercambio de datos con las instituciones implicadas a través de los organismos de enlace designados en cada uno de los países que intervienen en el proceso. Previamente al intercambio de datos, deberán generarse y validarse los correspondientes conjuntos de metadatos de acuerdo con la especificación ESSIM.

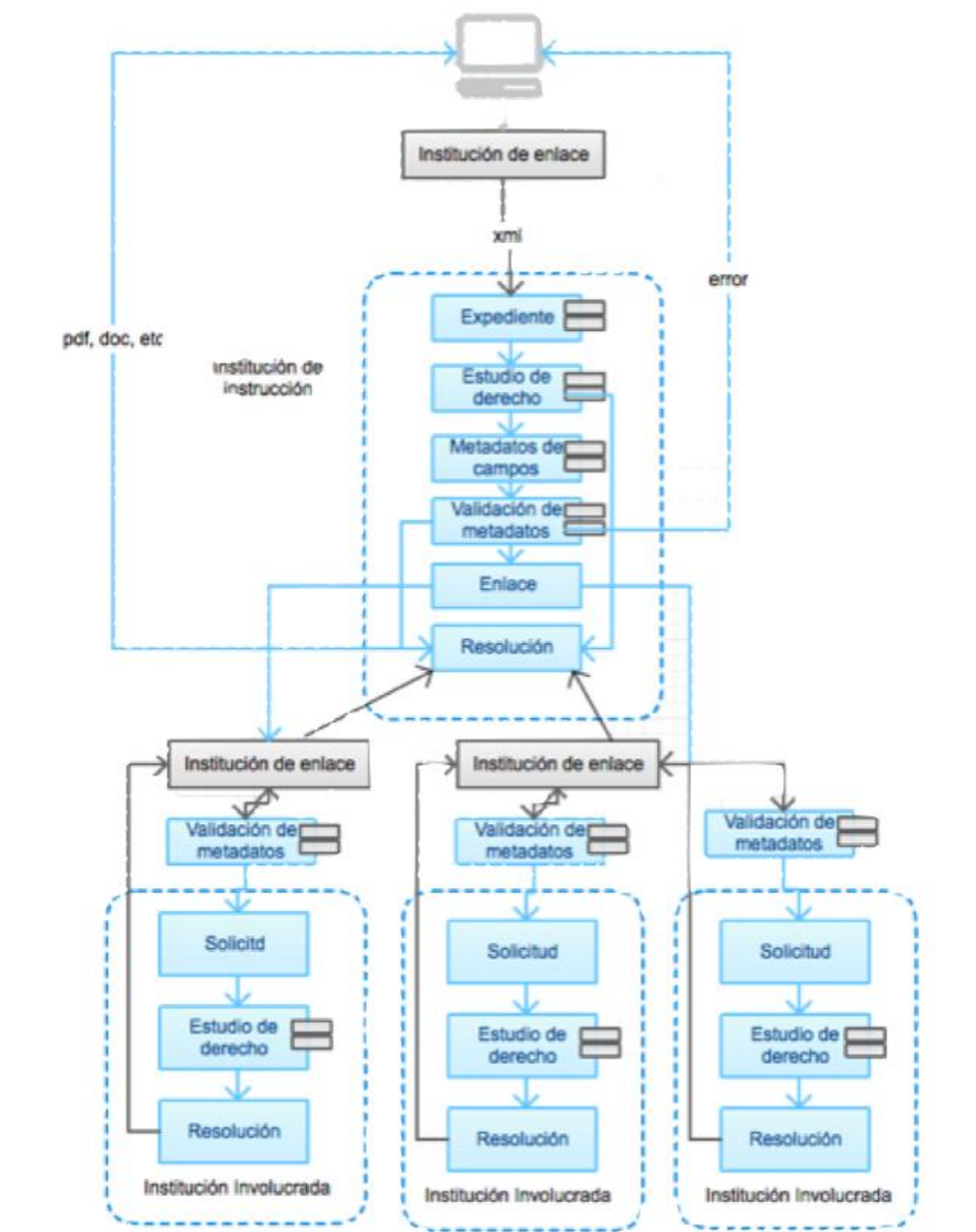


Figura 45. Proceso de concesión de una pensión

El intercambio deber realizarse en dos sentidos, de tal forma que la institución instructora reciba de las restantes instituciones involucradas la información necesaria para determinar si el solicitante tiene derecho a la pensión y en qué cuantía.

Si circunscribimos nuestro caso al ámbito de la Unión Europea, uno de los formularios que tendrá que ser objeto de intercambio será el llamado P5000 (figura 46), en el que se deben hacer constar los periodos a tener en cuenta para el reconocimiento de derecho y cálculo de la pensión.


Administrative Commission for the Coordination of Social Security Systems	<div style="border: 1px solid black; display: inline-block; padding: 2px 5px; font-weight: bold;">P5000</div> 				
Insurance/residence periods <i>Articles 1, 6, 45, 49, 51, 52, 57, 60 of Regulation (EC) No 883/2004; Articles 12(1), 47(4-5), 52(1b) of Regulation (EC) No 987/2009</i>					
<table style="width: 100%; border: 1px solid black;"> <tr> <td style="width: 30%;">Number of attachments</td> <td>[integer].....</td> </tr> <tr> <td>Date sent</td> <td>[DD/MM/YYYY].....</td> </tr> </table>		Number of attachments	[integer].....	Date sent	[DD/MM/YYYY].....
Number of attachments	[integer].....				
Date sent	[DD/MM/YYYY].....				
3. Insured person* 3.1 Person identification* 3.1.1 Personal Identification Number in the sending institution* [65]..... 3.1.2 Family name(s)* [155]..... 3.1.3 Forename(s)* [155]..... 3.1.4 Birth date* [DD/MM/YYYY]..... 3.1.5 Sex* <input type="checkbox"/> Female <input type="checkbox"/> Male <input type="checkbox"/> Unknown 3.1.6 Family name(s) at birth [155]..... 3.1.7 Forename(s) at birth [155]..... <hr style="border-top: 1px dashed black;"/> 3.1.9 If you do not have the Personal Identification Number of the person in the receiving institution, please fill in the following: <input type="checkbox"/> Identification of the person, without Personal Identification Number					

Figura 46. Formulario oficial P5000

Cada uno de los campos de este formulario deberá corresponder a un dato que deberá ser intercambiado entre las instituciones en su versión electrónica. Fijémonos por ejemplo en el campo 3.1.9.1 de dicho formulario, recogido en la tabla 44: *Place of birth*.

Place of birth		
Town of birth	Region of birth	Country of birth
<i>Barcelona</i>	<i>Cataluña</i>	<i>ES</i>

Tabla 44. Ejemplo de contenido para el campo *Place of birth*

Para describir este campo será necesario crear el conjunto de metadatos de acuerdo con la especificación ESSIM, de forma similar a cómo se hizo en el ejemplo del apartado anterior. Se muestra en la tabla 45 en formato JSON-LD.

```
{
  "@context": {
    "cvmap": "http://data.europa.eu/cv/",
    "essim": "https://socialsec.github.io/voc/essim/#",
    "eurlex": "https://socialsec.github.io/voc/eurlex/#",
    "got": "https://socialsec.github.io/voc/got/#",
    "ir": "https://socialsec.github.io/voc/ir/#",
    "iso639-1": "http://id.loc.gov/vocabulary/iso639-1/",
    "od": "https://socialsec.github.io/voc/od/#",
    "person": "https://joinup.ec.europa.eu/sites/default/files/distribution/2013-10/person-v1.00.rdf#",
    "xsd": "http://www.w3.org/2001/XMLSchema#"
  },
  "@type": "essim:Field",
  "essim:creatorInstitution": {"@id": "ir:3656"},
  "essim:dataSet": {
    "@type": "cvmap:Mapping",
    "cvmap:coreVocURI": {"@id": "person:placeOfBirth"},
    "cvmap:mappingRelation": {"@id": "cvmap:hasCloseMatch"}
  },
  "essim:dataTerm": {"@id": "got:IC007"},
  "essim:dateCreated": {
    "@value": "2014-05-30",
    "@type": "http://www.w3.org/2001/XMLSchema#date"
  },
  "essim:dateModified": {
    "@value": "2015-02-24",
    "@type": "http://www.w3.org/2001/XMLSchema#date"
  },
  "essim:description": {
    "@value": "State the place of birth",
    "@language": "en"
  },
  "essim:format": {
    "@value": "Alphanumeric",
    "@language": "en"
  },
  "essim:identifier": {"@value": "place-of-birth"},
  "essim:language": {"@id": "iso639-1:es"},
  "essim:mandate": [{"@id": "eurlex:32004R0883"}, {"@id": "eurlex:32009R0987"}],
  "essim:references": [{"@id": "od:P5000"}, {"@id": "od:P9000"}],
  "essim:relationWithTerms": {
    "@type": "essim:Relation",
    "essim:relationTerm": {"@id": "got:IC009"},
  }
}
```

```

"essim:relationDescription": {
  "@value": "Usually one nationality matches the place of birth",
  "@language": "en"
},
"essim:senderInstitution": {"@id": "ir:2800"},
"essim:socialSecurityCategory": {"@id": "ir:CSS21006"}
}

```

Tabla 45. Descripción del campo *Place of birth* con metadatos ESSIM

4.5.2 Esquema de la arquitectura

La arquitectura propuesta refleja el sistema de intercambio de datos (figura 47), mostrando como las instituciones en diferentes países intercambian información a través de sus correspondientes organismos de enlace e interactúan con una entidad responsable de gestionar los vocabularios RDF compartidos y los recursos referenciados en la especificación de metadatos. Esta entidad, llamada organismo de enlace o *“International Link Organización”*, también realiza el proceso de validación basado en los metadatos utilizando consultas SPARQL.

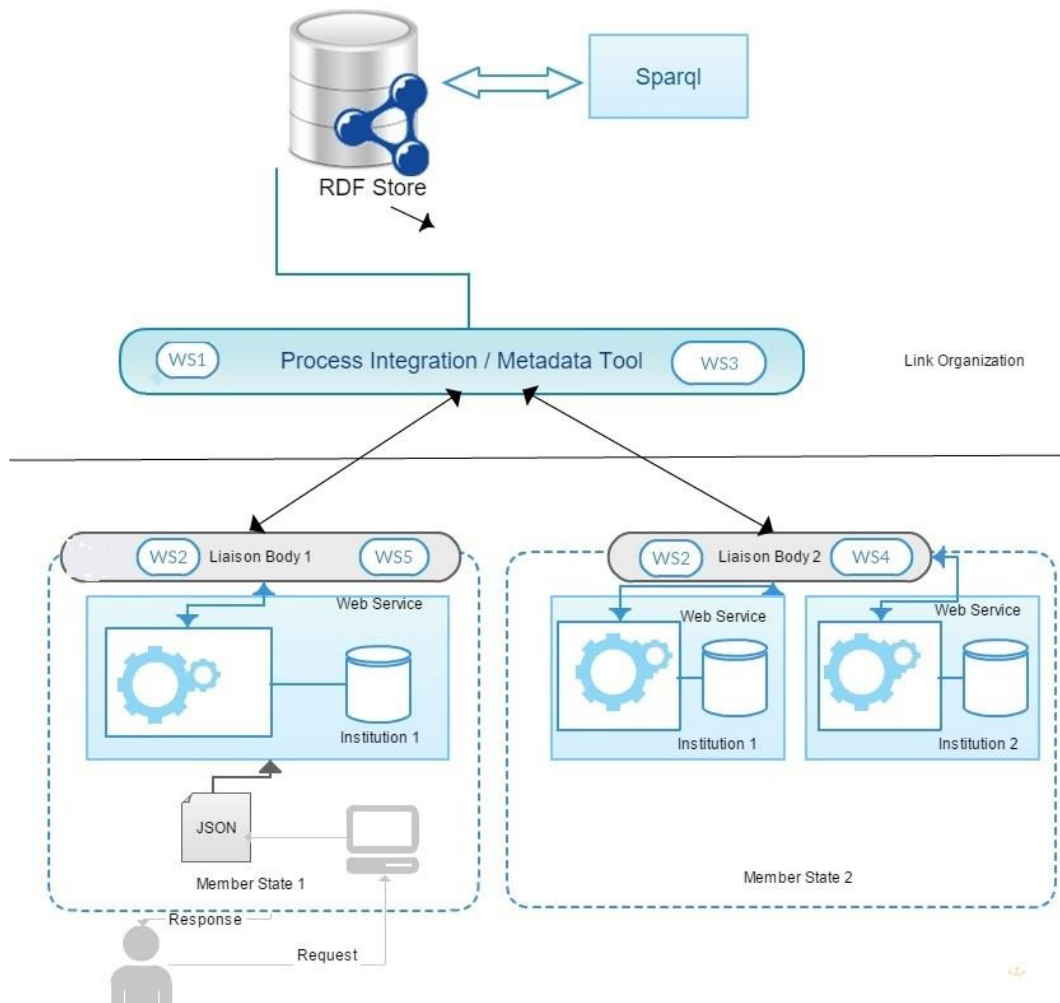


Figura 47. Arquitectura de intercambio propuesta

Ejecutar estas consultas requiere lo que se ha llamado “SPARQL end point”, servidor que interpreta el código de las *queries* SPARQL y lo lanza contra el almacén RDF o el conjunto de datos que incluyen los vocabularios y sus valores.

La figura 47 muestra un solo almacén RDF que comprende los vocabularios y valores desarrollados para esta prueba (essim, got, ir, eurlex, od), pero los vocabularios podrían distribuirse en varios servidores proporcionando servicios semánticos a varias instituciones en el mismo país.

Como se ha dicho en la sección 4.2.3, técnicamente hay dos formas de implementar servicios web. Por una parte, los llamados servicios web tradicionales, basados en SOAP y WSDL. Por otra parte, los llamados RESTful, basados en la tecnología REST (Pautasso et. al, 2008). Estos últimos fueron los elegidos para construir la prueba de la arquitectura.

4.5.3 Servicios web desarrollados

Para llevar a cabo la prueba de la arquitectura se crearon cinco servicios web con tecnología REST (tabla 46). Las principales interacciones relacionadas con la semántica entre los elementos de la arquitectura se basan en estos cinco servicios. Tres de ellos están relacionados directamente con los metadatos mientras que otros dos llevan a cabo los intercambios de datos entre dos instituciones. Todos los servicios implementan un único método tipo GET.

ID	Método	Descripción
WS1	verification (GET)	Verifica si hay acuerdo en vigor entre dos instituciones
WS2	metadata (GET)	Crea la estructura de metadatos ESSIM asociados a un campo de un documento
WS3	validation (GET)	Comprueba los valores de una estructura de metadatos ESSIM
WS4	request (POST)	Acepta la petición de un trámite administrativo y genera un ticket de referencia
WS5	reply (POST)	Responde a la institución instructora sobre un trámite administrativo solicitado

Tabla 46. Servicios web/métodos incluidos en la arquitectura

Tal como se ha explicado en la sección 4.5.1, como ejemplo para mostrar cómo funciona esta arquitectura, se tomó el caso de negocio real de un proceso de Seguridad Social para reconocer el derecho a una pensión de jubilación solicitada por un ciudadano que había trabajado en más de un país. En estas circunstancias, debe haber un intercambio de información entre las distintas

instituciones. La institución donde el ciudadano ha presentado la solicitud tiene que pedir información (periodos de trabajo del solicitante) a las otras instituciones involucradas y recopilar las contestaciones a través de los organismos de enlace en los diferentes países. En esta información intercambiada entre las instituciones debe evitarse cualquier problema semántico, razón por la que se propone acompañarla de los correspondientes metadatos. Además, la exactitud de los metadatos asociados a cada campo se garantiza mediante la utilización de los vocabularios RDF y se comprueba utilizando SPARQL.

La figura 48 muestra el *workflow* que representa el ejemplo descrito en formato BPMN. El diagrama representa los cinco servicios web involucrados en el proceso de negocio en el contexto de la arquitectura propuesta. Las siguientes secciones describen más en detalle estos cinco servicios web y su utilización en el ejemplo sobre solicitud de pensión de jubilación.

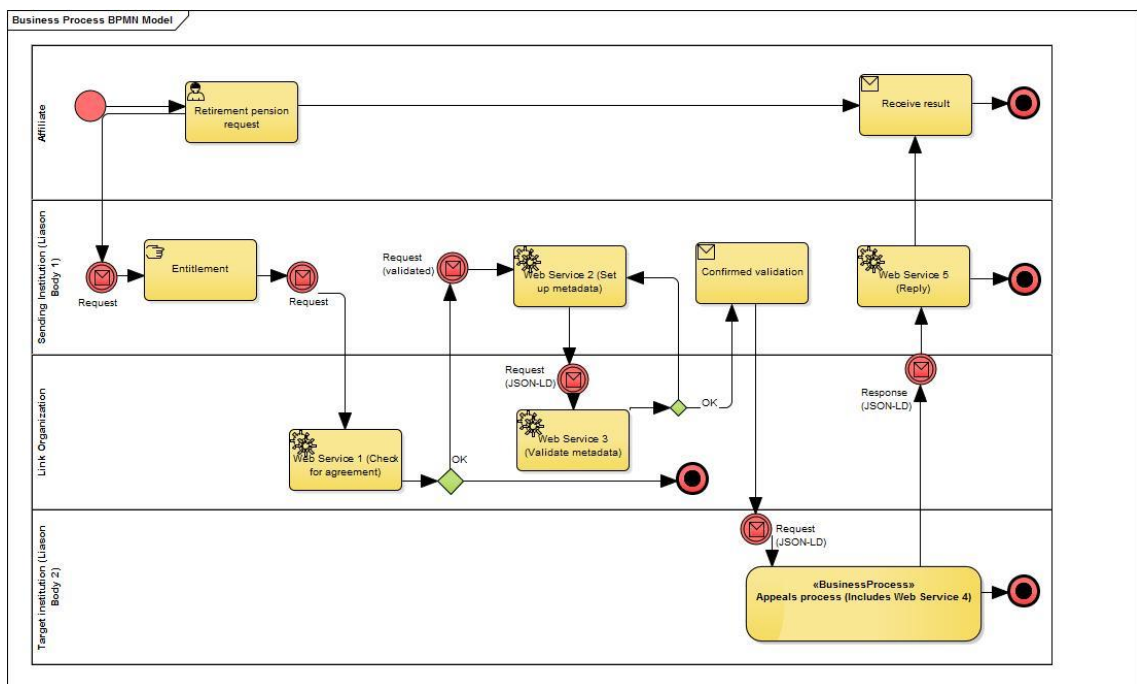


Figura 48. Ejemplo de *workflow* del proceso de negocio

La prueba realizada es un caso simplificado para mostrar la utilidad de la propuesta ESSIM en el intercambio de información en el contexto internacional de organizaciones de Seguridad Social. En el caso real, estos servicios deberían ser parte de una arquitectura más desarrollada y desdoblarse en servicios mucho más complejos.

Es importante indicar que las instituciones de seguridad social de diferentes países pueden contar con muy distintas configuraciones. La arquitectura se abstrae de los detalles sobre los procesos internos de las instituciones, así como de los detalles de interoperabilidad técnica que debieran establecerse para los intercambios.

La implementación de los servicios se realizó con tecnología Java. Antes de implementarlos se realizó su diseño o especificación utilizando el estándar de facto Open API, inicialmente conocido como Swagger (OpenAPI, 2018).

4.5.4 Descripción de los servicios web

4.5.4.1 Servicio WS1: Verificación de acuerdo en vigor

Este servicio permite comprobar si existe un acuerdo en vigor entre la institución donde se haya presentado la solicitud y los restantes países en los que el solicitante manifieste haber trabajado. La comprobación se realiza en el ámbito nacional a través del organismo de enlace (*Link Organization*) correspondiente a la institución donde se ha realizado la solicitud. En la tabla 47 se muestra la especificación de este servicio utilizando la notación OpenAPI.

```
openapi: 3.0.0
info:
  version: 1.0.0
  title: EESSIM (Link Organization)
  description: Services offered by a Link Organization
servers:
  - url: URL_BASE_LINK_ORGANIZATION
paths:
  /WS1/verification:
    get:
      summary: Verify the existence of a type of agreement between two Social Security
      Institutions
      operationId: verifyAgreement
      tags:
        - WS1
      parameters:
        - name: typeOfAgreement
          in: query
          description: IRI of a type of Social Security Agreement in a controlled
          vocabulary
          required: true
          schema:
            type: string
            example: 'https://socialsec.github.io/voc/agreements/#Pension'
        - name: institution1
          in: query
          description: IRI of an institution included in the Institutions Repository
          required: true
          schema:
            type: string
            example: 'https://socialsec.github.io/voc/ir/#2800'
        - name: institution2
          in: query
          description: IRI of an institution included in the Institutions Repository
          required: true
          schema:
            type: string
            example: 'https://socialsec.github.io/voc/ir/#1410000000'
      responses:
        '200':
          description: General response
          content:
            application/json:
              schema:
                $ref: '#/components/schemas/AgreementVerification'
components:
  schemas:
    AgreementVerification:
      required:
```

```

- eessim_error_code
- value
properties:
  eessim_error_code:
    type: string
    description: Standarized error code
    example: '2045'
  eessim_error_desc:
    type: string
    description: Error description
    example: 'Agreement not found'
  value:
    type: boolean
    description: Result
    example: False

```

Tabla 47. Descripción del servicio WS1 con notación OpenAPI

Existen editores de especificaciones de servicios con OpenAPI que permiten ver la descripción de forma más amigable, como es el caso, por ejemplo, del editor Swagger (figura 49). Además, esta herramienta ofrece la posibilidad de generar el código del esqueleto de la programación del servicio (en diferentes lenguajes y *frameworks* de programación) tanto en la parte del servidor donde se implementará, como el del cliente que lo vaya a invocar.

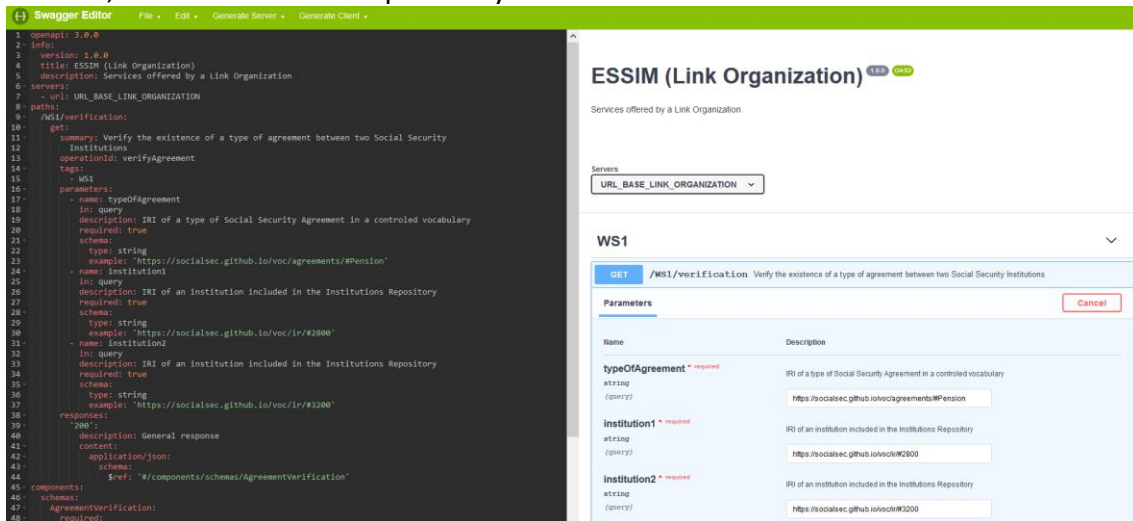


Figura 49. Especificación del servicio WS1 con Swagger Editor

En la figura 50 se muestra el detalle visual de la especificación del servicio WS1 en la herramienta indicada. Puede comprobarse que la verificación se ofrece como un recurso denominado “/WS1/verification” que se invoca mediante un método GET, y suministrando tres parámetros de tipo *query*, es decir a continuación de la URL del método, de la siguiente forma:

```

URL_BASE_LINK_ORGANIZATION/WS1/verification?typeOfAgreement={IRI}&institution1={IRI}&institution2{IRI}

```

GET /agreements/verification Verify the existence of a type of agreement between two Social Security Institutions

Parameters Cancel

Name	Description
typeOfAgreement * required string (query)	IRI of a type of Social Security Agreement in a controlled vocabulary <input type="text" value="https://socialsec.github.io/voc/agreements/#Pension"/>
institution1 * required string (query)	IRI of an institution included in the Institutions Repository <input type="text" value="https://socialsec.github.io/voc/ir/#2800"/>
institution2 * required string (query)	IRI of an institution included in the Institutions Repository <input type="text" value="https://socialsec.github.io/voc/ir/#1410000000"/>

Execute

Responses

Code	Description	Links
200	<p>General response</p> <p>application/json</p> <p>Content-Access header:</p> <p>Example Value Model</p> <pre>{ "essim_error_code": "2045", "essim_error_desc": "Agreement not found", "value": false }</pre>	No links

Figura 50. Especificación visual del servicio WS1 en Swagger Editor

Como puede verse en la figura 50 y en la documentación del servicio en la tabla 47, este método recibe tres parámetros que son IRIs de recursos disponibles en vocabularios o repositorios RDF. En la especificación de un servicio con OpenAPI se pueden incluir ejemplos de valores para los parámetros y para los resultados, en este caso con los ejemplos incluidos en la descripción, una posible invocación del servicio podría ser el siguiente:

```
URL_BASE_LINK_ORGANIZATION/WS1/verification?
  typeOfAgreement='https://socialsec.github.io/voc/agreements/#Pension'
  &institution1='https://socialsec.github.io/voc/ir/#2800'
  &institution2='https://socialsec.github.io/voc/ir/#1410000000'
```

Se han suministrado los parámetros siguientes:

- *typeOfAgreement*: Representa un tipo de acuerdo descrito en un vocabulario o repositorio RDF. En el ejemplo se hace referencia a un acuerdo denominado “Pension” sobre la tramitación de pensiones de jubilación.
- *Institution1*: Representa una institución de Seguridad Social registrada en el Repositorio de Instituciones, en el ejemplo se hace referencia al “INSTITUTO NACIONAL DE LA SEGURIDAD SOCIAL DE MADRID DIRECCION PROVINCIAL”, que se identifica como recurso 2800 en el repositorio.
- *Institution2*: Como en el caso anterior, representa otra institución. En el ejemplo, se refiere a la “ASSEMBLEE NATIONALE SERVICE DES AFFAIRES SOCIALES” de Francia, con identificador 1410000000, que tiene competencias para gestionar pensiones de jubilación en Francia.

Lo que hace este servicio es comprobar en una base de datos interna, si existe un convenio del tipo indicado entre las dos instituciones. El resultado es una estructura JSON con tres campos, cuya descripción y ejemplos de posibles valores está incluido en la especificación del servicio con OpenAPI (Tabla 47). Un ejemplo de posible resultado al invocar el servicio podría ser el siguiente:

```
{
  "eessim_error_code": "2045",
  "esessim_error_desc": "Type of agreement unknown",
  "value": false
}
```

Los dos primeros campos indican si se ha producido un error, y el tercer campo es el que señala si existe (*true*) o no existe (*false*) acuerdo entre ambas instituciones.

La implementación de este servicio y de los siguientes se realizó en un servidor web Java EE, utilizando la tecnología JAX-RS (*Java API for RESTful Web Services*) y el *framework* de persistencia JPA. Para la prueba del servicio se utilizó una base de datos con una tabla en la que se registraron diferentes combinaciones de instituciones y tipos de acuerdos entre ellas.

4.5.4.2 Servicio WS2: Creación de metadatos

Este servicio se encarga de crear los metadatos asociados a un campo incluido en un documento oficial, de acuerdo con lo previsto en la especificación ESSIM. La propuesta es que sea un servicio que se aloje en el dominio nacional, en el organismo de enlace correspondiente a la institución solicitante (*Liason Body* en la figura 47). En los ejemplos se utilizará como URL base del servicio: URL_BASE_LIASON_BODY_1, para diferenciarla del organismo de la institución de destino, que en los siguientes apartados se nombrará como URL_BASE_LIASON_BODY_2.

Este servicio se ofrece asociado a una aplicación web con interfaz de usuario. Se trata de la aplicación o formulario web de creación de metadatos de la sección 4.4 (figura 34). Aquí se presenta sólo la especificación del método que recibe los valores que ha introducido o seleccionado en cada caja de texto de la aplicación para cada uno de los metadatos, y devuelve la estructura ESSIM de dichos metadatos en formato JSON-LD. Es decir, para los valores que en el ejemplo del capítulo anterior se mostraba en las figuras 35 a 44, devolvería la estructura JSON-LD de la tabla 39.

El servicio incorpora otros microservicios o métodos auxiliares necesarios para la interacción con el usuario, como son aquellos necesarios para autocompletar las listas desplegables de aquellos metadatos asociados a vocabularios controlados. Por ejemplo, en el caso del metadato *dataSet*, si se trabaja con el *Person Core Vocabulary*, en la caja de texto en la que se le debe asignar un valor, se debe ofrecer al usuario la lista de términos que incluye el vocabulario “*Person*”, para que elija el que le interese. Para ello, en servicio W2 ofrece también un método GET que es invocado desde el código cliente JavaScript de la página JSP de la aplicación para rellenar la lista de valores.

El código HTML de la caja de texto en la que se solicita un valor para ese metadato es:

```
<label for="person">Data Set (Core Vocabulary):</label><br>
<select id="personrdf"></select>
```

Es decir, se trata de un elemento `<select>` que debería tener anidados tantos elementos `<option>` como posibles valores que se le ofrecen al usuario para que seleccione uno de ellos. Estos elementos `<option>` se generan al cargar la página web, ejecutando el código JavaScript de la tabla 48, que hace una llamada al método *personterms* que es ofrecido como parte del servicio WS2 (se trata de un extracto simplificado, que hace uso del *framework* jQuery).

```
<script type="text/javascript">
...
var selectList = [];
jQuery.ajax({
url: ' URL_BASE_LIASON_BODY_1/WS2/personterms
  type: 'GET',
  dataType : 'json',
  success: function(result){
    jQuery.each(result, function(index, personTerm){
      selectList.push(result.comment);
      $("#person").append("<option></option>")
        .text(result.literal)
        .val(result.literalId));});
...
</script>
```

Tabla 48. Código JavaScript que invoca un método GET del servicio WS2

Lo que devuelve el método GET del recurso *personterms* es una estructura de tipo JSON (no de tipo JSON-LD), compuesta por una lista de objetos que se refieren a cada uno de los términos del vocabulario *Person*, con tres campos que describen cada término: *literalId*, *literal* y *comment*. Con el código JavaScript se rellena una variable local *selectlist* con dichos objetos, y se asigna al texto de cada opción de la lista desplegable en el interior del elemento “*#person*” de la página, el valor del campo *literal* de cada término, que es el campo que contiene el nombre del término. El resto de campos se utilizan después en otros elementos asociados al metadato *dataSet* que no se muestran en el listado.

Por tanto, el servicio WS2 incluye varios métodos que son invocados por la aplicación web de ayuda al usuario para rellenar los metadatos, y finalmente un método que recibe todos los valores de las cajas de texto del formulario de la aplicación y devuelve una estructura, en este caso JSON-LD, con

el paquete de metadatos con la estructura propuesta en esta tesis, denominada ESSIM. En la tabla 49 se muestra la especificación OpenAPI de este método como parte del servicio WS2.

```
openapi: 3.0.0
info:
  version: 1.0.0
  title: EESSIM (Link Organization)
  description: Services offered by a Liason Body Organization
servers:
  - url: URL_BASE_LIASON_BODY
paths:
  /WS2/metadata:
    get:
      summary: Create the ESSIM metadata to describe one field of a form
      operationId: setUpMetadata
      tags:
        - WS2
      parameters:
        - name: ci
          in: query
          description: Creator Institution
          required: true
          schema:
            type: string
            example: '3656: INSTITUTO SOCIAL DE LA MARINA DIRECCION PROVINCIAL DE
VILLAGARCIA'
        - name: ds
          in: query
          description: Data Set
          required: false
          schema:
            type: string
            example: Place Of Birth
        - name: dt
          in: query
          description: Data Term
          required: true
          schema:
            type: string
            example: 'IC007: Place of birth'
        - name: dc
          in: query
          description: Date Created
          required: true
          schema:
            type: string
            example: '2014-05-30'
        - name: dm
          in: query
          description: Date Modified
          required: false
          schema:
            type: string
            example: '2015-02-24'
        - name: de
          in: query
          description: Description
          required: true
          schema:
            type: string
            example: State the place of birth
        - name: fo
          in: query
          description: Format
          required: true
          schema:
            type: string
            example: Alphanumeric
        - name: id
          in: query
          description: Identifier
          required: false
```

```

    schema:
      type: string
      example: place-of-birth
  - name: la
    in: query
    description: Language
    required: false
    schema:
      type: string
      example: es
  - name: ma
    in: query
    description: Mandate
    required: true
    schema:
      type: string
      example: '32004R0883, 32009R0987'
  - name: re
    in: query
    description: References
    required: false
    schema:
      type: string
      example: 'P5000, P9000'
  - name: rt
    in: query
    description: Relation Term
    required: false
    schema:
      type: string
      example: 'IC009'
  - name: rd
    in: query
    description: Relation Description
    required: false
    schema:
      type: string
      example: 'Usually a nationality matches the place of birth'
  - name: si
    in: query
    description: Sender Institution
    required: true
    schema:
      type: string
      example: '2800'
  - name: sc
    in: query
    description: Social Security Category
    required: true
    schema:
      type: string
      example: 'Old Age Pensions'
responses:
  '200':
    description: Correct response
    content:
      application/json:
        schema:
          $ref: '#/components/schemas/MetadataSchema'
  default:
    description: unexpected error
    content:
      application/json:
        schema:
          $ref: '#/components/schemas/Error'
components:
  schemas:
    Error:
      required:
        - code
        - message
      properties:
        code:
          type: integer
          format: int32
        message:
          type: string

```

```

MetadataSchema:
  required:
    - essim:creatorInstitution
    - essim:dataTerm
    - essim:dateCreated
    - essim:description
    - essim:format
    - essim:socialSecurityCategory
  properties:
    "@context":
      type: string
      description: Prefixes used
      example: '{"cvmap": "http://data.europa.eu/cv/", "essim":
"https://socialsec.github.io/voc/essim/#", "eurlex":
"https://socialsec.github.io/voc/eurlex/#", "got":
"https://socialsec.github.io/voc/got/#", "ir": "https://socialsec.github.io/voc/ir/#",
"iso639-1": "http://id.loc.gov/vocabulary/iso639-1/", "od":
"https://socialsec.github.io/voc/od/#", "person":
"https://joinup.ec.europa.eu/sites/default/files/distribution/2013-10/person-
v1.00.rdf#", "xsd": "http://www.w3.org/2001/XMLSchema#"}'
    "@type":
      type: string
      description: Type of an ESSIM resource
      example: 'essim:Field'
    essim:creatorInstitution:
      type: string
      description: Social security institution primarily responsible for making
the data
      example: '{"@id": "ir:3656"}'
    essim:dataSet:
      type: string
      description: Core Vocabulary where a similar data concept could be found
      example: '{"@type": "cvmap:Mapping", "cvmap:coreVocURI": {"@id":
"person:placeOfBirth"}, "cvmap:mappingRelation": {"@id": "cvmap:hasCloseMatch"}}'
    essim:dataTerm:
      type: string
      description: Term included in the Glossary of Terms (GOT) that defines the
meaning of the data field
      example: '{"@id": "got:IC007"}'
    essim:dateCreated:
      type: string
      description: A point of time associated with the starting point in the
lifecycle of data
      example: '{"@value": "2014-05-30", "@type":
"http://www.w3.org/2001/XMLSchema#date"}'
    essim:dateModified:
      type: string
      description: Date when data was last modified or changed
      example: '{"@value": "2015-02-24", "@type":
"http://www.w3.org/2001/XMLSchema#date"}'
    essim:description:
      type: string
      description: An effective description is 25-30 words long or approximately
150 characters
      example: '{"@value": "State the place of birth", "@language": "en"}'
    essim:format:
      type: string
      description: Format of the data
      example: '{"@value": "Alphanumeric", "@language": "en"}'
    essim:identifier:
      type: string
      description: Identifier of the data (field)
      example: 'place-of-birth'
    essim:language:
      type: string
      description: Language of the data when not numeric
      example: '{"@id": "iso639-1:es"}'
    essim:mandate:
      type: string
      description: Legal reference of the exchanges where data is involved
      example: '[{"@id": "eurlex:32004R0883"}, {"@id": "eurlex:32009R0987"}]'
    essim:references:
      type: string
      description: Official documents where data is present
      example: '[{"@id": "od:P5000"}, {"@id": "od:P9000"}]'
    essim:relationWithTerms:
      type: string

```



```

description:      Description of relationship between the data and other
terms, different of the dataTerm
example:      '{"@type":      "essim:Relation",      "essim:relationTerm":      {"@id":
"got:IC009"}, "essim:relationDescription": {"@value": "Usually one nationality matches
the place of birth", "@language": "en"}}'
essim:senderInstitution:
type: string
description: Social security institution responsible for sending the data
example: '{"@id": "ir:2800"}'
essim:socialSecurityCategory:
type: string
description: Branch of social security involved in the exchange of data
example: '{"@id": "ir:CSS21006"}'

```

Tabla 49. Descripción del servicio WS2 con notación OpenAPI

La herramienta *Swagger Editor* con la que se ha diseñado el servicio, ofrece una utilidad para probar servicios en la que el usuario puede cumplimentar un formulario con los valores para los parámetros de entrada (figura 51).

WS2 ▼

GET /WS2/metadata Create the ESSIM metadata to describe one field of a form

Cancel

Name	Description
ci * required string (query)	Creator Institution <input style="width: 80%;" type="text" value="3656: INSTITUTO SOCIAL DE LA MARINA DIRECCION PR"/>
ds string (query)	Data Set <input style="width: 80%;" type="text" value="Place Of Birth"/>
dt * required string (query)	Data Term <input style="width: 80%;" type="text" value="IC007: Place of birth"/>
dc * required string (query)	Date Created <input style="width: 80%;" type="text" value="2014-05-30"/>
dm string (query)	Date Modified <input style="width: 80%;" type="text" value="2015-02-24"/>
de * required string (query)	Description <input style="width: 80%;" type="text" value="State the place of birth"/>
fo * required string (query)	Format <input style="width: 80%;" type="text" value="Alphanumeric"/>

Figura 51. Parámetros de entrada del método GET metadata

En la figura 52 se muestra el detalle visual de la parte de la especificación del servicio WS2 en la misma herramienta, en la que aparece el formato del resultado que devuelve el método GET "metadata".

```

MetadataSchema v {
  @context      string
                 example: {"cvmap": "http://data.europa.eu/cv/", "essim":
                 "https://socialsec.github.io/voc/essim/#", "eurlex": "https://socialsec.github.io
                 /voc/eurlex/#", "got": "https://socialsec.github.io/voc/got/#", "ir":
                 "https://socialsec.github.io/voc/ir/#", "iso639-1": "http://id.loc.gov/vocabulary
                 /iso639-1/", "od": "https://socialsec.github.io/voc/od/#", "person":
                 "https://joinup.ec.europa.eu/sites/default/files/distribution/2013-10/person-
                 v1.00.rdf#", "xsd": "http://www.w3.org/2001/XMLSchema#"}
                 Prefixes used

  @type         string
                 example: essim:Field
                 Type of an ESSIM resource

  essim:creatorInstitution* string
                 example: {"@id": "ir:3656"}
                 Social security institution primarily responsible for making the data

  essim:dataSet  string
                 example: {"@type": "cvmap:Mapping", "cvmap:coreVocURI": {"@id":
                 "person:placeOfBirth"}, "cvmap:mappingRelation": {"@id": "cvmap:hasCloseMatch"}}
                 Core Vocabulary where a similar data concept could be found

  essim:dataTerm* string
                 example: {"@id": "got:IC007"}
                 Term included in the Glossary of Terms (GOT) that defines the meaning of the data
                 field

  essim:dateCreated* string
                 example: {"@value": "2014-05-30", "@type": "http://www.w3.org/2001/XMLSchema#date"}
                 A point of time associated with the starting point in the lifecycle of data

  essim:dateModified string
                 example: {"@value": "2015-02-24", "@type": "http://www.w3.org/2001/XMLSchema#date"}
                 Date when data was last modified or changed

  essim:description* string
                 example: {"@value": "State the place of birth", "@language": "en"}
                 An effective description is 25-30 words long or approximately 150 characters

  essim:format*   string
                 example: {"@value": "Alphanumeric", "@language": "en"}
                 Format of the data

  essim:identifier string
                 example: place-of-birth
                 Identifier of the data (field)

```

Figura 52. Formato del resultado del método GET metadata

En la implementación de este servicio, además de las tecnologías Java usadas en el servicio anterior, también se ha utilizado el *framework* Apache Jena, para manipular en la memoria del servidor como modelos de objetos los vocabularios RDF implicados en la creación de la estructura de metadatos ESSIM.

4.5.4.3 Servicio WS3: Validación de metadatos

Este servicio comprueba los metadatos de acuerdo con los vocabularios y recursos establecidos por ESSIM. El principal objetivo es verificar si los valores de los metadatos de una estructura ESSIM se corresponden con los vocabularios y recursos propuestos para el dominio de Seguridad Social: *Glossary of Terms, Insitutions Repository, EUR-Lex y Official Documents*. Asimismo, se comprueba si se cumplen otras restricciones o reglas incluidas en los vocabularios, como por ejemplo la multiplicidad de valores permitidos.

Este servicio debe estar alojado en la *Link Organization* de la arquitectura de intercambio propuesta (figura 47), para ofrecer su funcionalidad de forma centralizada a todas las organizaciones de Seguridad Social. La URL para las solicitudes HTTP de este servicio deber ser del tipo:

```
URL_BASE_LINK_ORGANIZATION/WS3/validation?metadata={ESSIM metadata}
```

El parámetro de entrada en este caso es una estructura JSON-LD que describe los metadatos de ESSIM. Esta misma estructura es la salida del WS2, por lo que es posible combinar ambos servicios en la misma solicitud:

```
URL_BASE_LINK_ORGANIZATION/WS3/validate?metadata=URL_BASE_LIASON_BODY/WS2/metadata?fieldMetadata={raw values for metadata}
```

En el caso en que siempre se generasen los metadatos mediante el servicio WS2 podría no ser necesaria esta validación. Por ese motivo, no es obligatorio el uso de este servicio y se podrían recibir estructuras ESSIM que hayan sido creadas con otras aplicaciones o con un editor RDF. No obstante, es necesario asegurarse de que se cumple la especificación ESSIM antes de utilizar los metadatos para el intercambio con otra institución de Seguridad Social.

En la tabla 50 se muestra la especificación OpenAPI del método de validación como parte del servicio WS3.

```
openapi: 3.0.0
info:
  version: 1.0.0
  title: ESSIM (Link Organization)
  description: Services offered by a Link Organization
servers:
  - url: URL_BASE_LINK_ORGANIZATION
paths:
  /WS3/validation:
    get:
      summary: Validate ESSIM metadata format
      operationId: validateMetadata
      tags:
        - WS3
      parameters:
        - name: metadata
          in: query
          description: Field description using EESSIM metadata in JSON-LD format
          content:
            application/json:
              schema:
                $ref: '#/components/schemas/MetadataSchema'
      responses:
        '200':
          description: General response
          content:
            application/json:
              schema:
                $ref: '#/components/schemas/MetadataValidation'
components:
  schemas:
    MetadataSchema:
      required:
        - 'essim:creatorInstitution'
        - 'essim:dataTerm'
```

```

- 'essim:dateCreated'
- 'essim:description'
- 'essim:format'
- 'essim:socialSecurityCategory'
properties:
  '@context':
    type: string
    description: Prefixes used
    example: >-
      {"cvmap": "http://data.europa.eu/cv/", "essim":
      "https://socialsec.github.io/voc/essim/#", "eurlex":
      "https://socialsec.github.io/voc/eurlex/#", "got":
      "https://socialsec.github.io/voc/got/#", "ir":
      "https://socialsec.github.io/voc/ir/#", "iso639-1":
      "http://id.loc.gov/vocabulary/iso639-1/", "od":
      "https://socialsec.github.io/voc/od/#", "person":
      "https://joinup.ec.europa.eu/sites/default/files/distribution/2013-
      10/person-v1.00.rdf#",
      "xsd": "http://www.w3.org/2001/XMLSchema#"}
  '@type':
    type: string
    description: Type of an ESSIM resource
    example: 'essim:Field'
  'essim:creatorInstitution':
    type: string
    description: >-
      Social security institution primarily responsible for making the
      data
    example: '{"@id": "ir:3656"}'
  'essim:dataSet':
    type: string
    description: Core Vocabulary where a similar data concept could be found
    example: >-
      {"@type": "cvmap:Mapping", "cvmap:coreVocURI": {"@id":
      "person:placeOfBirth"}, "cvmap:mappingRelation": {"@id":
      "cvmap:hasCloseMatch"}}
  'essim:dataTerm':
    type: string
    description: >-
      Term included in the Glossary of Terms (GOT) that defines the
      meaning of the data field
    example: '{"@id": "got:IC007"}'
  'essim:dateCreated':
    type: string
    description: >-
      A point of time associated with the starting point in the lifecycle
      of data
    example: >-
      {"@value": "2014-05-30", "@type":
      "http://www.w3.org/2001/XMLSchema#date"}
  'essim:dateModified':
    type: string
    description: Date when data was last modified or changed
    example: >-
      {"@value": "2015-02-24", "@type":
      "http://www.w3.org/2001/XMLSchema#date"}
  'essim:description':
    type: string
    description: >-
      An effective description is 25-30 words long or approximately 150
      characters
    example: '{"@value": "State the place of birth", "@language": "en"}'
  'essim:format':
    type: string
    description: Format of the data
    example: '{"@value": "Alphanumeric", "@language": "en"}'
  'essim:identifier':
    type: string
    description: Identifier of the data (field)
    example: place-of-birth
  'essim:language':
    type: string
    description: Language of the data when not numeric
    example: '{"@id": "iso639-1:es"}'
  'essim:mandate':
    type: string
    description: Legal reference of the exchanges where data is involved

```

```

    example: '[{"@id": "eurlex:32004R0883"}, {"@id": "eurlex:32009R0987"}]'
  'essim:references':
    type: string
    description: Official documents where data is present
    example: '[{"@id": "od:P5000"}, {"@id": "od:P9000"}]'
  'essim:relationWithTerms':
    type: string
    description: >-
      Description of relationship between the data and other terms,
      different of the dataTerm
    example: >-
      {"@type": "essim:Relation", "essim:relationTerm": {"@id":
        "got:IC009"}, "essim:relationDescription": {"@value": "Usually one
        nationality matches the place of birth", "@language": "en"}}
  'essim:senderInstitution':
    type: string
    description: Social security institution responsible for sending the data
    example: '{"@id": "ir:2800"}'
  'essim:socialSecurityCategory':
    type: string
    description: Branch of social security involved in the exchange of data
    example: '{"@id": "ir:CSS21006"}'
  MetadataValidation:
    required:
      - eessim_error_code
      - value
    properties:
      eessim_error_code:
        type: string
        description: Standarized error code
        example: '34'
      esessim_error_desc:
        type: string
        description: Error description
        example: dataTerm metadata does not match with a term in the Glossary of
terms (GOT)
      value:
        type: boolean
        description: 'True if the validation is correct'
        example: false

```

Tabla 50. Descripción del servicio WS3 con notación OpenAPI

En la figura 53 se muestra el detalle visual de la parte de la especificación del servicio WS3 en la herramienta *Swagger Editor*, en la que aparece el formato del resultado que devuelve el método GET “validation”.

GET /ws3/validation Validate ESSIM metadata format

Models v

MetadataSchema ←

MetadataValidation ←

```

MetadataValidation {
  eessim_error_code string
    example: 34
    Standarized error code

  esessim_error_desc string
    example: dataTerm metadata does not match with a term in the Glossary of terms (GOT)
    Error description

  value* boolean
    example: false
    True if the validation is correct
}

```

Figura 53. Formato del resultado del método GET validation

En la implementación de este servicio se han utilizado tecnologías Java para procesar modelos RDF dado que recibe un parámetro JSON-LD que representa un grafo RDF, que debe procesarse para compararlo con otros grafos con los vocabularios y recursos utilizados.

Una decisión que hay que tomar antes de la implementación es la ubicación de estos vocabularios y recursos RDF. En este sentido se consideran dos opciones:

- Instalar un RDF Store en un servidor alojado en la institución de enlace, es decir en URL_BASE_LINK_ORGANIZATION, y almacenar todos los vocabularios y recursos de forma centralizada.
- Utilizar referencias a repositorios instalados en otros servidores con *RDF Stores* o que ofrecieran las tripletas RDF en forma de archivos.

En el prototipo creado para la validación se aplicó la primera opción, instalando en el servidor un *RDF Store*. Existen en el mercado diferentes servidores de este tipo, como son Virtuoso, AllegroGraph o Jena (Fuseki), entre otros (W3C, 2018; DB, 2018). Finalmente se hicieron las pruebas con un servidor Virtuoso (OpenLink, 2018).

Virtuoso ofrece un “*Jena RDF Data Provider*”, que permite la manipulación desde código Java de las tripletas RDF almacenadas en el servidor, para lo cual hace uso del *framework* de Apache conocido como Jena (Apache, 2018), utilizado en la implementación del servicio WS2. Jena ofrece una API Java para extraer y escribir datos en grafos RDF. Los grafos se representan en memoria como un modelo (*model*), a partir de datos procedentes de archivos, bases de datos, URI o una combinación

de ellos. Un modelo también puede ser consultado y actualizado utilizando expresiones SPARQL. En la figura 54 se muestra la forma de utilizar esta tecnología. En el contexto de la figura, el servicio WS3 se asemejaría al componente *Application*, y realiza llamadas a la API Jena para manipular grafos RDF como objetos o modelos en memoria, y también para acceder al almacenamiento RDF del servidor Virtuoso.

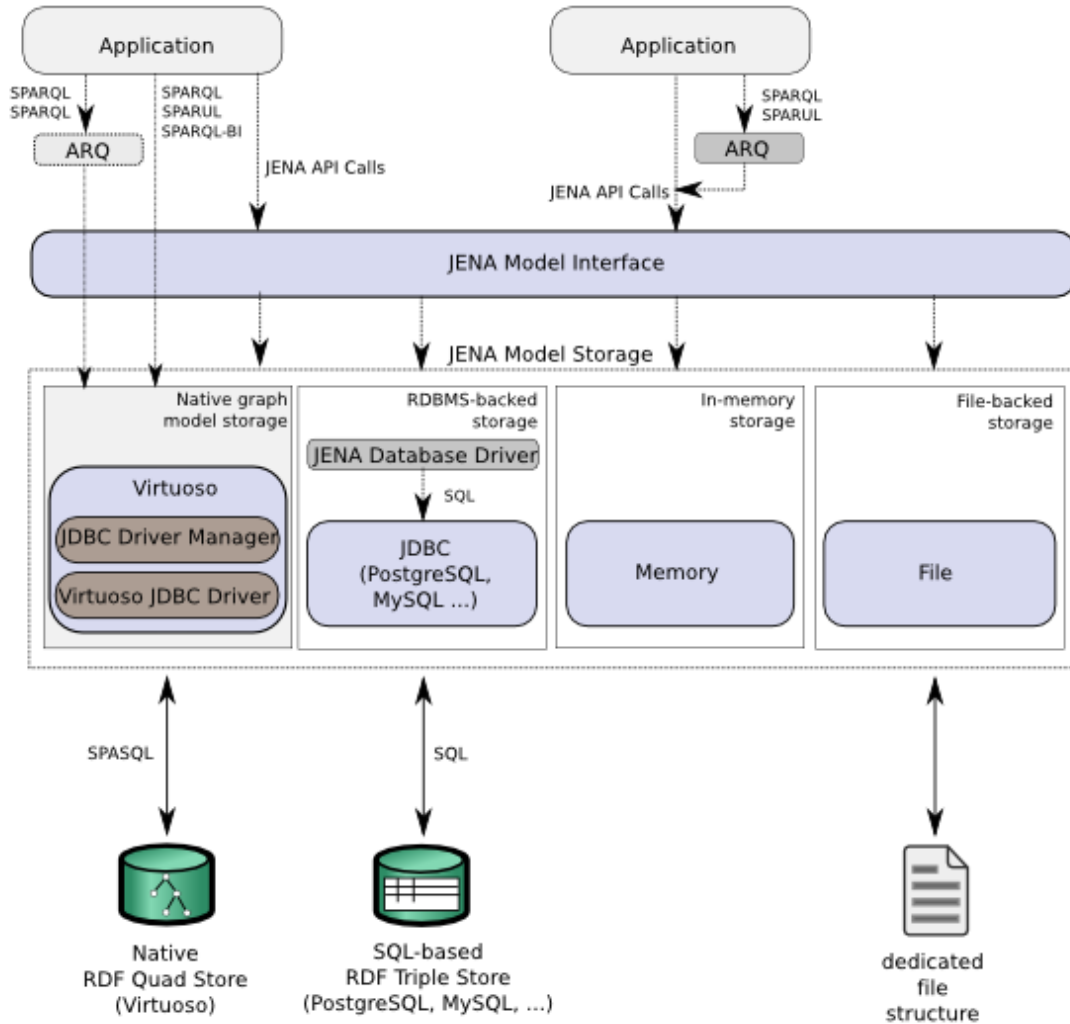


Figura 54. Virtuoso Jena RDF Data Provider (OpenLink, 2018)

En la tabla 51 se muestra un extracto del código Java del servicio WS3, en el que se recibe la estructura de metadatos ESSIM, y se procede a la validación de los metadatos. Sólo se muestra la sección en la que se invoca a una función interna encargada de la validación del metadato *dataTerm*, para confirmar que su valor se corresponde con la IRI de alguno de los términos del Glosario de Términos.

```
import virtuoso.jena.driver.VirtGraph;
import virtuoso.jena.driver.VirtuosoQueryExecution;
import virtuoso.jena.driver.VirtuosoQueryExecutionFactory;
import com.hp.hpl.jena.query.Query;
import com.hp.hpl.jena.query.QueryFactory;
```

```

import com.hp.hpl.jena.query.QuerySolution;
import com.hp.hpl.jena.query.ResultSet;
...
@GET
@Path("/validation")
@Consumes({"application/json;charset=UTF-8"})
@Produces({"application/json;charset=UTF-8"})
public Response validateMetadata(@QueryParam("metadata") String metadata) {
    String dataTerm = extractDataTermURI(metadata);
    MetadataValidation result = new MetadataValidation();
    if (!(validateDataTerm(dataTerm)) {
        result.setErrorCode ("34");
        result.setErrorDesc ("dataTerm metadata does not match with a term in GOT");
        result.setValue (false);
        return Response.ok(result).build();
    } else {
        // Validate the rest of "metadata" parameter: creatorInstitution, language, etc.
        ...
    }
}
...
public boolean validateDataTerm(String dataSet) {
    boolean exists = false;
    List<String> list = new ArrayList<String>();
    VirtGraph set = new VirtGraph(virtuoso_url, user, pwd);
    Query sparql = QueryFactory.create
        ("SELECT ?term WHERE {?term a <https://socialsec.github.io/voc/got/#Term>}");
    VirtuosoQueryExecution vqe = VirtuosoQueryExecutionFactory.create(sparql, set);
    ResultSet results = vqe.execSelect();
    while (results.hasNext()) {
        QuerySolution rs = results.nextSolution();
        String term = rs.getLiteral("term").getString();
        if (term != null) {
            list.add(term);
        }
    }
    for (int i = 0; i < list.size(); i++) {
        String t = list.get(i);
        if (t.equals(dataSet)) {
            exists = true;
            break;
        }
    }
    return exists;
}
...

```

Tabla 51. Descripción del servicio WS2 con notación OpenAPI

Como puede observarse en la tabla 51, lo primero que se hace es validar el contenido del metadato *essim:dataTerm* incluido como un campo de la estructura JSON-LD correspondiente al parámetro "metadata". Para ello se llama a una función local del servicio, denominada *extractDataTermURI*, para obtener el contenido de dicho campo, resolviendo el posible uso de prefijos. A continuación, se invoca una función *validateDataTerm*, a la que se le pasa como parámetro el valor extraído, para que lo compare con todos los términos que forman parte del Glosario de Términos.

Como ya se ha indicado, se ha decidido almacenar en el servidor Virtuoso todos los vocabularios y los recursos que derivan de éstos, como es el caso del GOT. Por ello lo que hace la función *validateDataTerm* es pedir al servidor Virtuoso, a través del constructor de la clase *VirtGraph* del *Virtuoso Jena Provider* (*virtuoso.jena.driver*), el contenido o grafo RDF almacenado en el servidor y lo almacena en el objeto *set*. Entonces es posible ejecutar sentencias SPARQL sobre dicho grafo. En este caso se ejecuta la sentencia siguiente:

```
SELECT ?term WHERE {?term a <https://socialsec.github.io/voc/got/#Term>.}
```


Esta sentencia devuelve la URI de los recursos que representan términos del GOT, es decir de la clase *Term* definida en el vocabulario.

Después se transforma la lista de términos recibidos en una lista de Java, y se comprueba si alguno de los términos coincide con el valor del metadato *dataTerm*. Si el resultado que devuelve la función es afirmativo, el método de validación del servicio WS3 pasaría a validar otro metadato y así sucesivamente hasta comprobar que no haya ningún error.

4.5.4.4 Servicio WS4: Solicitud de trámite administrativo

Este servicio recibe una solicitud de trámite administrativo por parte de una institución instructora (*sending institution*) y confirma la recepción de la misma por parte del organismo de enlace nacional (*liason body*) de la institución receptora. La contestación será un código (o *ticket*) que la institución de envío podrá utilizar en el futuro para asociar cualquier posible respuesta con el expediente correcto. La descripción del servicio se muestra en la tabla 52.

```
openapi: 3.0.0
info:
  version: 1.0.0
  title: ESSIM (Liason Body)
  description: Services offered by a Liason Body
servers:
  - url: URL_BASE_LIASON_BODY
paths:
  /WS4/request:
    post:
      summary: Accept administrative process request and generate a ticket for future
      references
      operationId: processRequest
      tags:
        - WS4
      requestBody:
        description: Request information
        required: true
        content:
          application/json:
            schema:
              $ref: '#/components/schemas/RequestSchema'
      responses:
        '200':
          description: Correct response
          content:
            application/json:
              schema:
                properties:
                  ticket:
                    type: string
                    example: '14109699384970'
          default:
            description: unexpected error
            content:
              application/json:
                schema:
                  $ref: '#/components/schemas/Error'
components:
  schemas:
    RequestSchema:
      required:
        - 'processId'
        - 'data'
        - 'metadata'
      properties:
        processId:
          description: The id of the administrative process
```

<pre> type: string data: description: Data included in the request type: string metadata: description: Data included in the request type: string Error: required: - code - message properties: code: type: integer format: int32 message: type: string </pre>

Tabla 52. Descripción del servicio WS4 con notación OpenAPI

Como se indica en la especificación del servicio, ofrece un método POST *request* que recibe como datos una estructura JSON con los siguientes elementos:

- *processId*: Identificador del proceso o trámite administrativo solicitado.
- *data*: Datos de solicitud. Son los valores de los campos del formulario de solicitud (en formato JSON).
- *metadata*: Metadatos ESSIM que describan cada uno de los datos o campos de la solicitud (en formato JSON-LD).

Es conveniente destacar la diferencia en el formato de los datos y metadatos. Los datos de la solicitud se representan en formato JSON mientras que los metadatos se representan en formato JSON-LD, es decir representan un grafo RDF, tal y como se ha previsto en la propuesta ESSIM presentada en esta tesis.

En la tabla 53 puede verse un ejemplo de datos de entrada de una solicitud que se presenta utilizando el formulario oficial P5000 de la figura 46. En el campo *processId* se utiliza el valor “P5000” para indicar que es una solicitud de información sobre periodos a tener en cuenta para el reconocimiento de derecho y cálculo de la pensión.

En la sección *data* se incluyen los valores de los campos del formulario P5000 implicado en la solicitud. Por simplicidad sólo se muestran los contenidos de tres de los campos del formulario: *date-sent*, *family-name* y *place-of-birth*.

Por último, la sección *metadata* contiene una estructura de objetos JSON-LD, cada uno de los cuales describe alguno de los campos incluidos en la sección *data*. No es necesario describir todos los campos, sólo aquellos que puedan ser susceptibles de ser malentendidos por la institución receptora de la solicitud. En el ejemplo, de los tres datos de la sección *data*, sólo se describen con metadatos los campos *family-name* y *place-of-birth*, y no se describe el campo *date-sent*, por considerar que es evidente su significado y formato, ya que aparece en el propio formulario. La

forma de relacionar los metadatos con el dato o campo que describen es a través del metadato *essim:identifijer*.

```
{
  "processId": "P5000",
  "data": [
    {
      "date-sent": "30/05/2014",
      "family-name": "Martínez",
      "place-of-birth": {
        "town-of-birth": "Barcelona",
        "region-of-birth": "Cataluña",
        "country-of-birth": "España"
      }
    }
  ],
  "metadata": [
    {
      "@context": {
        "cvmap": "http://data.europa.eu/cv/",
        "essim": "https://socialsec.github.io/voc/essim/#",
        "eurlex": "https://socialsec.github.io/voc/eurlex/#",
        "got": "https://socialsec.github.io/voc/got/#",
        "ir": "https://socialsec.github.io/voc/ir/#",
        "iso639-1": "http://id.loc.gov/vocabulary/iso639-1/",
        "od": "https://socialsec.github.io/voc/od/#",
        "person": "https://joinup.ec.europa.eu/sites/default/files/distribution/2013-10/person-vl.00.rdf#",
        "xsd": "http://www.w3.org/2001/XMLSchema#"
      },
      "@type": "essim:Field",
      "essim:creatorInstitution": {
        "@id": "ir:3656"
      },
      "essim:dataSet": {
        "@type": "cvmap:Mapping",
        "cvmap:coreVocURI": {
          "@id": "person:familyName"
        },
        "cvmap:mappingRelation": {
          "@id": "cvmap:hasCloseMatch"
        }
      },
      "essim:dataTerm": {
        "@id": "got:IC001"
      },
      "essim:dateCreated": {
        "@value": "2014-05-30",
        "@type": "http://www.w3.org/2001/XMLSchema#date"
      },
      "essim:dateModified": {
        "@value": "2015-02-24",
        "@type": "http://www.w3.org/2001/XMLSchema#date"
      },
      "essim:description": {
        "@value": "Oficial name of a person",
        "@language": "en"
      },
      "essim:format": {
        "@value": "Alphabetic text",
        "@language": "en"
      },
      "essim:identifijer": {
        "@value": "family-name"
      },
      "essim:language": {
        "@id": "iso639-1:es"
      },
      "essim:mandate": [
        {
          "@id": "eurlex:32004R0883"
        }
      ],
      "essim:references": [
        {
          "@id": "od:P5000"
        }
      ]
    }
  ]
}
```

```

    },
    {
      "@id": "od:E101"
    }
  ],
  "essim:relationWithTerms": {
    "@type": "essim:Relation",
    "essim:relationTerm": {
      "@id": "got:IC004"
    },
    "essim:relationDescription": {
      "@value": "The Surname value has officially substituted the previous names
of the person",
      "@language": "en"
    }
  },
  "essim:senderInstitution": {
    "@id": "ir:2800"
  },
  "essim:socialSecurityCategory": {
    "@id": "ir:CSS21006"
  }
},
{
  "@context": {
    "cvmap": "http://data.europa.eu/cv/",
    "essim": "https://socialsec.github.io/voc/essim/#",
    "eurlex": "https://socialsec.github.io/voc/eurlex/#",
    "got": "https://socialsec.github.io/voc/got/#",
    "ir": "https://socialsec.github.io/voc/ir/#",
    "iso639-1": "http://id.loc.gov/vocabulary/iso639-1/",
    "od": "https://socialsec.github.io/voc/od/#",
    "person": "https://joinup.ec.europa.eu/sites/default/files/distribution/2013-
10/person-v1.00.rdf#",
    "xsd": "http://www.w3.org/2001/XMLSchema#"
  },
  "@type": "essim:Field",
  "essim:creatorInstitution": {
    "@id": "ir:3656"
  },
  "essim:dataSet": {
    "@type": "cvmap:Mapping",
    "cvmap:coreVocURI": {
      "@id": "person:placeOfBirth"
    },
    "cvmap:mappingRelation": {
      "@id": "cvmap:hasCloseMatch"
    }
  },
  "essim:dataTerm": {
    "@id": "got:IC007"
  },
  "essim:dateCreated": {
    "@value": "2014-05-30",
    "@type": "http://www.w3.org/2001/XMLSchema#date"
  },
  "essim:dateModified": {
    "@value": "2015-02-24",
    "@type": "http://www.w3.org/2001/XMLSchema#date"
  },
  "essim:description": {
    "@value": "State the place of birth",
    "@language": "en"
  },
  "essim:format": {
    "@value": "Alphanumeric",
    "@language": "en"
  },
  "essim:identifier": {
    "@value": "place-of-birth"
  },
  "essim:language": {
    "@id": "iso639-1:es"
  },
  "essim:mandate": [
    {
      "@id": "eurlex:32004R0883"
    }
  ]
}

```

```

    },
    {
      "@id": "eurlex:32009R0987"
    }
  ],
  "essim:references": [
    {
      "@id": "od:P5000"
    },
    {
      "@id": "od:P9000"
    }
  ],
  "essim:relationWithTerms": {
    "@type": "essim:Relation",
    "essim:relationTerm": {
      "@id": "got:IC009"
    },
    "essim:relationDescription": {
      "@value": "Usually one nationality matches the place of birth",
      "@language": "en"
    }
  },
  "essim:senderInstitution": {
    "@id": "ir:2800"
  },
  "essim:socialSecurityCategory": {
    "@id": "ir:CSS21006"
  }
}
]
}

```

Tabla 53. Ejemplo de datos de una solicitud

Este servicio debería incluirse (ver figura 48, diagrama BPMN) en el proceso llamado *Appeals process* que debería encargarse de gestionar la solicitud y preparar la respuesta. El mecanismo utilizado debería ser similar al flujo seguido por la institución remitente (*sending institution*):

- Comprobar la existencia del acuerdo entre las dos instituciones mediante el servicio WS1.
- Resolver la solicitud y preparar un documento con la información solicitada.
- Generar los metadatos para el documento de respuesta mediante el servicio local WS2.
- Validar los metadatos en el organismo de enlace mediante el servicio WS3.

La implementación de un prototipo para probar este servicio se ha realizado también en Java, en el mismo servidor en el que se ha instalado el servicio WS2.

4.5.4.5 Servicio WS5: Respuesta a un trámite

Mediante este servicio web la institución a la que se envió una solicitud, responde con el resultado de la misma. Como el servicio WS4, éste también se debe instalar en el organismo de enlace nacional (*Liason Body*). Por tanto, en un proceso de solicitud/respuesta, la institución solicitante invocará el servicio WS4 alojado en el servidor del organismo de la institución receptora (URL_BASE_LIASON_BODY_SENDER/WS4), y posteriormente, una vez resuelta, la institución receptora invocará el servicio WS5 alojado en el servidor de la emisora (URL_BASE_LIASON_BODY_RECEIVER/WS5). La descripción del servicio se muestra en la tabla 54.

```

openapi: 3.0.0
info:
  version: 1.0.0
  title: ESSIM (Liason Body)
  description: Services offered by a Liason Body
servers:
  - url: URL_BASE_LIASON_BODY
paths:
  /WS5/reply:
    post:
      summary: Receives the result of an administrative process request identified by
a ticket
      operationId: processReply
      tags:
        - WS5
      requestBody:
        description: Reply to a request
        required: true
        content:
          application/json:
            schema:
              $ref: '#/components/schemas/ReplySchema'
      responses:
        '200':
          description: Correct response
          content:
            application/json:
              schema:
                $ref: '#/components/schemas/ResponseSchema'
        default:
          description: unexpected error
          content:
            application/json:
              schema:
                $ref: '#/components/schemas/Error'
components:
  schemas:
    ReplySchema:
      required:
        - 'ticket'
        - 'message'
        - 'data'
        - 'metadata'
      properties:
        ticket:
          description: Ticket identifying the request
          type: string
        message:
          description: Message about the result of the request
          type: string
        data:
          description: Data included in the reply document to the request
          type: string
    ResponseSchema:
      required:
        - response
      properties:
        eessim_error_code:
          type: string
          description: Standarized error code
          example: '12345'
        eessim_error_desc:
          type: string
          description: Error description
          example: 'Ticket not found'
    ok:
      type: boolean
      description: Resul

```

Tabla 54. Descripción del servicio WS5 con notación OpenAPI

Como se indica en la especificación del servicio, ofrece un método POST *request* que recibe como datos una estructura JSON con los siguientes elementos:

- *ticket*: Identificador recibido cuando la solicitud fue enviada mediante el servicio WS4.
- *message*: Un mensaje que informa sobre el resultado de la solicitud.
- *data*: Datos de la respuesta. Son los valores de los campos del documento o formulario de respuesta (en formato JSON).
- *metadata*: Metadatos ESSIM que describan cada uno de los datos o campos del documento o formulario de respuesta (en formato JSON-LD).

```
{
  "ticket": "14109699384970",
  "message": "Request attended successfully",
  "data": [
    {
      "field-1": "Content of field 1",
      "field-2": "Content of field 2",
      "field-n": "Content of field n"
    }
  ],
  "metadata": [
    {
      "@context": {
        "essim": "https://socialsec.github.io/voc/essim/#"
      },
      "@type": "essim:Field",
      "essim:creatorInstitution": {},
      "essim:dataTerm": {},
      "essim:dateCreated": {},
      "essim:description": {},
      "essim:format": {},
      "essim:identifier": "field-1",
      "essim:socialSecurityCategory": {}
    },
    {
      "@context": {
        "essim": "https://socialsec.github.io/voc/essim/#"
      },
      "@type": "essim:Field",
      "essim:creatorInstitution": {},
      "essim:dataTerm": {},
      "essim:dateCreated": {},
      "essim:description": {},
      "essim:format": {},
      "essim:identifier": "field-n",
      "essim:socialSecurityCategory": {}
    }
  ]
}
```

Tabla 55. Ejemplo simplificado de datos de una respuesta a una solicitud

En la tabla 55 puede verse un ejemplo de datos de entrada de una respuesta a la solicitud del servicio WS4 explicada en el apartado anterior. En el campo *ticket* aparece el identificador que se recibió como respuesta al método *request*. En la sección *data* se incluyen los valores de los campos

del documento de respuesta, identificados como *field-1*, *field-2*, ... *field-n*. La sección *metadata* contiene una estructura de objetos JSON-LD, cada uno de los cuales describe alguno de los campos incluidos en la sección *data*. En el ejemplo, por simplicidad, sólo aparecen los metadatos ESSIM obligatorios, pero sin ningún valor concreto. La relación de los metadatos con el dato o campo que describen se realiza es a través del metadato *essim:identifier*.

4.6 Conclusiones de las pruebas

Con la primera prueba se demostró que es posible crear el conjunto de metadatos propuesto en el capítulo 3 con una herramienta que podría ser instalada en cualquier institución de enlace de un estado. A menudo se atribuye a una excesiva complejidad el papel secundario que han jugado hasta ahora los metadatos en las organizaciones (Manso-Callejo et. al, 2010). La aplicación aportada es un ejemplo en sentido contrario, permite al usuario crear metadatos sin conocer detalles técnicos de implementación RDF, además de un control directo de la información que se está creando. Los valores asignados a los elementos están garantizados por la utilización de los vocabularios controlados, si bien sería necesario actualizar y completar su información en la mayoría de los casos. De eso se hablará más extensamente en el capítulo 5.

También quedaría pendiente de discusión el punto del ciclo de vida en que se deben generar los metadatos, también tratado por (Manso-Callejo et. al, 2010), así como los diversos aspectos de su mantenimiento.

En la segunda prueba se demostró que la solución es compatible con una arquitectura general de intercambio de datos. El estudio de caso empleado podría calificarse como el más representativo de la problemática de Seguridad Social. Los servicios web descritos proporcionan un flujo mínimo independiente de un posterior desglose a más bajo nivel.

Por otra parte, la utilización de una tecnología de servicios web REST y SPARQL se eligió por constituir estándares de facto para los intercambios por internet. También para la especificación o descripción de los servicios se utilizó uno de los lenguajes más utilizados en este campo, como es OpenAPI.

A pesar de la garantía que significa la utilización de vocabularios controlados, ambas pruebas utilizaron formas de verificación capaces de disipar las dudas sobre la calidad de las herramientas de generación que plantea (Niso, 2004).

Por último, en cuanto a la aportación de los metadatos en sí mismos, los ejemplos utilizados ponen de manifiesto el papel que puede jugar la especificación propuesta en la definición de un dato.

Los campos utilizados en los ejemplos se referían a apellido y a lugar de nacimiento. Ambos son usados habitualmente en los trámites administrativos relacionados con la Seguridad Social. Sin embargo, en ambos casos, especialmente en el del lugar de nacimiento, el campo puede presentar un alto grado de incertidumbre a menos que:

- Presente una estructura estandarizada para su tratamiento

Esta información la proporciona el metadato *dataSet*, haciendo referencia a su definición en el *Core Vocabulary* correspondiente.

- Se conozca la lengua en que está expresada (muchas localidades tienen nombres diferentes en las distintas lenguas).

Esta información la proporciona el metadato *language*.

Los restantes elementos de ESSIM ayudarán a determinar el origen e historia del dato (institución que lo ha creado, fecha, cambios, etc.), información que puede ser también relevante en el caso de que surja alguna duda en el momento de la identificación del ciudadano.

Capítulo 5. Conclusiones y trabajos futuros

5.1 Recapitulación de objetivos

El capítulo 1 recogía el contexto general en el que se desarrolla esta tesis. Los datos ahí recogidos muestran claramente que el fenómeno migratorio está en pleno crecimiento. Ese hecho se traduce en un número creciente de acuerdos internacionales en materia de Seguridad Social que necesita para su implementación práctica que se establezcan intercambios de datos entre las instituciones de Seguridad Social de distintos países.

El objetivo general se sitúa en el contexto de esos intercambios de datos y tomando como referencia la definición del concepto de interoperabilidad del EIF. Se ha partido de una perspectiva general de la problemática para después profundizar en los objetivos establecidos en el capítulo 1.

A continuación, se sintetizan las conclusiones para cada uno de los cuatro objetivos propuestos.

- **Objetivo 1: Análisis de los intercambios de datos internacionales de Seguridad Social desde el punto de vista de la interoperabilidad.**

En el capítulo 2 se ha analizado la situación de los intercambios electrónicos de información entre instituciones de Seguridad Social en todas sus dimensiones.

En el campo de la interoperabilidad política y legal se ha visto cómo existen gran cantidad de acuerdos que dan soporte a los intercambios (sección 1.2), si bien su puesta en práctica presenta un porcentaje de éxito muy bajo (sección 2.7).

En lo que se refiere a la **interoperabilidad organizacional**, el desarrollo de los sistemas de Seguridad Social en todo el mundo conlleva una complejidad en lo que a organizaciones se refiere, de la que no escapa ningún esquema nacional. El problema se estudia en profundidad en la sección 2.3.3.2, en el que se cita el artículo publicado en “El Profesional de la Información” (Delgado et al, 2012), proponiéndose una solución basada en la creación de un Directorio Maestro de Instituciones.

Respecto a la **interoperabilidad técnica**, cuenta con herramientas sólidas y suficientemente divulgadas como para facilitar la solución de la mayoría de los problemas que puedan presentarse (sección 2.3.3.4). El factor clave es alcanzar un acuerdo sobre las tecnologías concretas a utilizar y la forma de ponerlas en práctica.

En la **dimensión semántica**, la investigación reveló la necesidad de contar con un modelo conceptual común (esquema de metadatos), en el que se describa en términos de conceptos la información que se intercambia, sus propiedades y las relaciones entre estos conceptos, lo que da respaldo a la propuesta general de esta tesis.

- **Objetivo 2: Vocabularios existentes en el contexto estudiado**

Se han revisado los glosarios, taxonomías, tesauros y ontologías existentes, tanto en el ámbito la Administración Pública en general como en de la Seguridad Social, que pudieran contribuir a paliar los problemas semánticos existentes (sección 2.3.3.3 y 2.4.3).

No se ha encontrado ningún vocabulario o esquema de metadatos que se ajuste a la problemática concreta de los intercambios de datos de Seguridad Social. Algunos ejemplos podrían ser utilizados como apoyo, pero necesitarían de una actualización o adaptación. Especialmente relevantes son los casos del GOT y los *Core Vocabularies* de la UE.

- **Objetivo 3: Propuesta de una nueva especificación.**

Se ha propuesto una nueva especificación de metadatos (ESSIM) que permite acompañar los datos intercambiados con un conjunto de metadatos especialmente diseñados para evitar cualquier ambigüedad en su significado.

A pesar de la existencia de instituciones de Seguridad Social en todos los países del mundo, las infraestructuras tecnológicas y los recursos en las diferentes instituciones, así como sus limitaciones, dan lugar a una heterogeneidad que aconseja que la propuesta se base, hasta dónde sea posible, en estándares reconocidos.

La especificación propuesta por esta tesis (sección 3.2.3) se basa en los elementos de *Dublin Core*, con las adaptaciones concretas al medio de los intercambios entre instituciones de Seguridad Social que fueron descritas en los congresos ISD 2013 (Delgado et al., 2013b), en ATICA 2013 (Delgado, 2013d) , en ECEG2014 (Delgado et al., 2014) y en ATICA 2018 (Delgado, 2018). Es importante resaltar que la nueva especificación se basa en la utilización de vocabularios controlados, estudiados en el artículo publicado en LHT (Delgado et al., 2013 C), entre otros.

- **Objetivo 4: Propuesta de un esquema de arquitectura de intercambios**

Se ha propuesto en el capítulo 4 un esquema de arquitectura de intercambios entre instituciones de Seguridad Social basada en servicios web que demuestra la viabilidad y aplicabilidad de la especificación propuesta. El esquema se presentó en el congreso ATICA 2014 (Viñán et. Al, 2014).

La arquitectura propuesta refleja el sistema de intercambio de datos (figura 47, Arquitectura de intercambio propuesta) mediante cinco servicios web que se ajustan a los requerimientos más generales (tabla 46, servicios web/métodos incluidos en la arquitectura).

5.2 Aportaciones originales

La aportación fundamental de la tesis es la publicación de la especificación ESSIM, especialmente concebida para el intercambio de datos entre instituciones de Seguridad Social. Con la publicación de esta especificación se realizan las siguientes contribuciones:

- **En el campo de la interoperabilidad.**

Se aporta una solución para los problemas que actualmente se están produciendo en los intercambios de datos de Seguridad Social. Las medidas implementadas hasta el momento

generalmente se centran en los aspectos organizacionales y técnicos, con lo que se consigue que los datos lleguen finalmente a su destino, pero sin evitar las confusiones en su significado.

La mejora planteada en esta tesis se basa en recomendaciones internacionales y está en línea con las experiencias actuales más importantes. A diferencia de otras propuestas sobre interoperabilidad, en su mayoría basadas en aspectos técnicos, se centra en la interoperabilidad semántica. Incluso en esta dimensión, los pocos encontrados en el campo semántico, sólo llegan a un nivel sintáctico, no propiamente semántico.

En el capítulo 3 se demuestra que es esencial que un mismo dato tenga la misma definición en todos los documentos a intercambiar, especialmente si se trata de documentos electrónicos. De ahí surge la necesidad de establecer un conjunto de metadatos de referencia, publicarlo y mantenerlo de forma accesible, tanto para su utilización por las aplicaciones, como para su consulta por los funcionarios que van a tratar los documentos desde el punto de vista administrativo.

- **En el campo de la estandarización.**

Se presenta una nueva especificación de metadatos basada en un estándar ISO como es *Dublin Core*, con vocación de convertirse en un estándar para un dominio determinado como es la administración electrónica y la Seguridad Social.

En la sección 2.5 se ha pasado revista a las distintas definiciones del concepto de metadato existentes en la literatura especializada, así como a su evolución a lo largo del tiempo. Con respecto a otras especificaciones existentes, ESSIM presenta la novedad de centrarse en el dato y no en el documento o en otro tipo de recurso, con lo que se retoma el concepto original de los metadatos (datos que describen datos) y se garantiza la interpretación de la información transmitida en su parte más próxima a su significado.

En la nueva especificación se proponen seis nuevos elementos para el dominio concreto de los intercambios de datos de Seguridad Social, justificándose la necesidad de cada uno de ellos. A su vez, los valores de los datos descritos por los metadatos se basan en vocabularios controlados.

- **Implicaciones prácticas.**

La movilidad actual de los trabajadores hace que el problema de fondo se presente en prácticamente cualquier país del mundo, pero muy especialmente en los que están implicados en los movimientos migratorios por razones laborales. La Unión Europea y MERCOSUR serían los ejemplos actuales más claros de escenarios donde se realiza un intercambio intenso de información y dónde aparecen los problemas estudiados en la tesis.

Hay que tener en cuenta que dentro de la protección social hay que incluir el derecho a recibir asistencia sanitaria en un país distinto del de nacionalidad, en el caso de que exista convenio de reciprocidad. En este apartado concreto se está incluyendo todas las personas que se desplazan, por ejemplo, por motivos turísticos. Así mismo, se incluye toda la facturación que esto ocasiona e intercambio de información entre los países implicados.

- **Implicaciones sociales**

La mejora de la protección social y de los derechos en materia de Seguridad Social de los ciudadanos que se desplazan, mediante la automatización de los acuerdos internacionales en este ámbito, mejorando las comunicaciones entre las instituciones de diferentes países. No hay que olvidar que la Seguridad Social tiene un fuerte impacto en el equilibrio de mercado y en la justicia social.

5.3 Líneas de trabajo futuro

5.3.1 Ámbito profesional

La propuesta de esta tesis está directamente relacionada con actividades que se están llevando a cabo en diversos ámbitos profesionales del entorno de la Seguridad Social. Por ese motivo, se espera que las líneas de trabajo iniciadas tengan continuidad tanto a través de la AISS como de la UE, así como en los múltiples intercambios de datos en este contexto que tengan lugar en el futuro.

- **A través de la AISS**

La propuesta presentada en esta tesis está relacionada directamente con los trabajos realizados por el autor como miembro de la red de expertos de la AISS, que da su apoyo a instituciones de todo el mundo para la implementación de esquemas de Seguridad Social a través de las directrices (ISSA, 2015) y especificaciones adicionales, entre otros temas, para la interoperabilidad de sistemas de Seguridad Social. La coordinación de los trabajos presentados en esta tesis con otras iniciativas internacionales, no sólo hace posible la aplicación de las soluciones presentadas, sino que deja abierta una línea de trabajo que permitirá desarrollar y perfeccionar las soluciones semánticas.

Como parte del trabajo futuro, se espera que ESSIM sea presentado como un estándar ISO por la AISS en un futuro próximo. Una vez se haya establecido el estándar, será necesario definir un formato de campos que facilite el intercambio de los conjuntos de metadatos.

Para facilitar esa labor, en el anexo 1 se incluye un borrador del documento a presentar a ISO, utilizando el formato habitual para estos estándares.

- **A través de la Unión Europea**

En el momento de redactarse esta tesis, el Proyecto de EESSI ha entrado oficialmente en fase de producción, con un periodo transitorio para que las distintas instituciones empiecen a intercambiar mensajes que se extiende hasta julio de 2019. Sin embargo, los temas semánticos no sólo no se han resuelto, sino que siguen siendo objeto de debate en todas las reuniones. Diversos documentos tratados en el desarrollo del proyecto han sido referenciados en los capítulos anteriores por su relación directa con la materia tratada.

En particular, el grupo de modelado de datos está pendiente de un nuevo mandato, en función de los primeros resultados de la fase de producción.

Dado el objetivo con el que ha sido creado, las nuevas actividades de este grupo deben guardar relación con el contenido de esta tesis y pueden dar continuidad a sus propuestas

5.3.2 Proyectos de investigación

- **Actualización y extensión de los vocabularios controlados**

Para que la estandarización propuesta sea eficaz y consiga su fin, debe apoyarse en una serie de vocabularios controlados y espacios de nombres, de tal forma que su contenido identifique de forma inequívoca y sin posibilidad de confusión los datos que definen. A pesar de que los valores de los elementos de ESSIM están basados en vocabularios controlados existentes, el contenido de estos vocabularios necesita ser revisado, completado y extendido a todas las regiones del mundo.

- **Glosarios de Términos del proyecto TESS**

En la sección 2.3.3.3 se ha descrito el Glosario de Términos creado dentro del subproyecto Build 6 del proyecto TESS. Este glosario contenía las definiciones de los conceptos relacionados con los intercambios de datos de Seguridad Social al amparo de los Reglamentos Comunitarios. En el momento de su creación, se recogieron los términos de los entonces vigentes formularios E. Con la aprobación de los Reglamentos 883 (OJEU, 2004) y 987 (OJEU, 2009), dichos términos pueden haber cambiado. Por otra parte, el hecho de tratarse actualmente de documentos electrónicos estructurados puede dar lugar a nuevos requerimientos en la información recogida y sus atributos.

Asimismo, para que pueda ser una base sólida para un estándar mundial, debería contrastarse con la información intercambiada fuera de la Unión Europea, especialmente en aquellos casos que puedan ser culturalmente más diferentes.

Con motivo de la prueba 1 descrita en la sección 4.3, se preparó una versión en XML del actual GOT que puede ser muy útil como punto de partida de una nueva investigación.

Más adelante, podría desarrollarse una ontología partiendo del glosario siguiendo los métodos evolutivos propuestos por Hilera et al. (2010).

- ***Institution Repository***

En la sección 2.3.3.2 se ha descrito el Directorio Maestro de Instituciones desarrollado por el proyecto EESSI. Una nueva versión de este directorio, conocido como *Institution Repository*, está a punto de publicarse en el momento de redactarse este capítulo. Dicha versión va a ser más completa y sofisticada que la anterior, pero, como en el caso del GOT, su ámbito se limita en estos momentos al Espacio Económico europeo.

Contar con un directorio de instituciones a nivel mundial sería un avance importante no sólo para resolver los problemas de interoperabilidad organizacional, sino como alimentar los metadatos con los valores correctos que permitan la definición unívoca de los datos.

- ***Core Vocabularies***

En cualquier especificación de metadatos, es recomendable utilizar modelos de datos estandarizados siempre que estos existan. La Unión Europea, dentro del programa ISA, ha creado una serie de artefactos para su reutilización. Se trata de modelos de datos simplificados, reusables y extensibles. Según se ha dicho en la sección 2.4.3, hasta el momento se han desarrollado cuatro (*Public Services, Location, Business y Person*).

Por encima es de estos vocabularios debería completarse el modelo para un dominio concreto, como en el caso de la Seguridad Social. Una vez más, sería deseable que fuera aplicable en cualquier parte del mundo.

- ***EUR-Lex***

Es un servicio open data que proporciona acceso gratuito a las leyes de la Unión Europea y otros documentos públicos. Cada uno de sus documentos está registrado con metadatos analíticos, como referencia de su publicación, fecha o palabras claves.

La legislación aplicable es relevante para el tratamiento a realizar con los datos intercambiados, hasta el punto de ser necesarios unos datos u otros en función de los requisitos legales vigentes. Por ese motivo, todo intercambio debería contar con la referencia legal que le da validez, bien sea el acuerdo internacional que le da respaldo o bien sea legislación o convenios multilaterales cuando se trate de comunidad de estados como es el caso de la Unión Europea o Mercosur.

Una extensión a nivel mundial de EUR-Lex, así como el desarrollo de una herramienta más sofisticada para facilitar el acceso a su información, serían de gran ayuda para mejorar la calidad de los intercambios.

- ***Official Documents***

Un vocabulario donde figuren todos los documentos existentes y susceptibles de ser intercambiados puede ser útil para obtener referencias cruzadas y verificar relaciones entre los documentos, los datos que figuran en los documentos y la función de cada uno de estos datos en los diferentes procedimientos.

Este diccionario no existe en estos momentos de una manera formal y debería ser creado desde cero, empezando por la construcción de un inventario mundial de documentos de Seguridad Social susceptibles de ser intercambiados.

- **Enfoque hacia *Linked Data***

En las conclusiones del artículo publicado en la revista LHT (Delgado et. Al, 2013C) se proponía que las instituciones de Seguridad Social considerasen la posibilidad de alcanzar el nivel 2 de la política de interoperabilidad (*Formal semantic interoperability*) descrita por *Dublin Core* (2018), incluyendo el uso de RDF para documentar sus metadatos, diseño persistente, así como uso y mantenimiento de URL enlazadas con vocabularios y esquemas de datos externos, armonizando sus recursos con recursos de terceras partes, etc.

Las políticas de *open government* o gobierno abierto son calificadas por Marcos-Martin (2011) como estrategias de alto calado, que promueven tres pilares en la actuación de los poderes públicos: la transparencia en la gestión pública, el fomento de la participación y la colaboración de la sociedad en el diseño y ejecución de las políticas públicas. En este contexto la apertura de datos u *open data* aboga por que las administraciones públicas pongan a disposición de la sociedad la información pública que tienen en su poder, con el fin de que cualquier persona u organización pueda a partir de la misma crear nuevas informaciones y servicios.

En esa línea, podría estudiarse la utilización de vocabularios abiertos establecidos por diferentes comunidades o instituciones y relacionarlas para ser utilizadas por sistemas de información heterogéneos.

- **Desarrollo de la arquitectura de intercambio**

Los servicios web descritos en el modelo de arquitectura de forma esquemática (sección 4.5), deberían ser desarrollados con todo rigor para construir una arquitectura sólida que facilite los intercambios de datos entre instituciones de Seguridad Social en el ámbito internacional, cubriendo las cinco dimensiones de interoperabilidad, sin olvidar las exigencias de seguridad y protección de datos.

Referencias

- AC (2012). European Commission, Administrative Commission for the coordination of social security systems. Note of the countries Austria, Germany, Netherlands, Switzerland and Liechtenstein of 24 September 2012. EESSI: Necessary improvements for implementation on national level (AC 370/12REV). Texto completo en anexo 2.
- AC (2012b). Comisión Europea. Nota de la delegación española en la Comisión Técnica de 28 de octubre de 2012. EESSI: mejoras necesarias para su implantación. Aspectos semánticos (AC 12-446es). Texto completo en anexo 2.
- AC (2013). European Commission, Administrative Commission for the coordination of social security systems. Note from the Secretariat EESSI: Mandate for the Ad-hoc group on data modelling (AC 522/13REV). Texto completo en anexo 2.
- AC (2015). European Commission, Administrative Commission for the coordination of social security systems. Note from the Data Modelling Expert Group of 18 November 2015. EESSI: Data Modelling Expert Group recommendations on the basic principles to be followed in the data modelling work (AC 738/15). Texto completo en anexo 2.
- Activa (2018). Secretaría de Estado de la Seguridad Social. 1.815.092 afiliados extranjeros en enero. Revista Activa. <https://revista.seg-social.es/2018/02/20/>. Accedido 21-2-2018.
- ADMS (2013). European Commission. Asset Description Metadata Schema. <https://joinup.ec.europa.eu/>. Accedido 24-5-2018
- AENOR (2011). Asociación Española de Normalización y Certificación. Conjunto de elementos de metadatos de *Dublin Core*. UNE-ISO 15836. <http://www.aenor.es/aenor/normas>. Accedido 30-3-2018.
- AGLS (2010). National Archives of Australia. Using metadata to describe information and records. <http://www.naa.gov.au/records-management/>. Accedido 24-5-2018.
- AISS (2013). AISS. Directrices de la AISS sobre Tecnologías de la Información y de la Comunicación. Ginebra 2013. ISBN 978-92-843-3202-1.
- AISS (2014). AISS. Manual sobre la extensión de la cobertura de la seguridad social a los trabajadores migrantes. ISBN: 978-92-843-9184-4.
- Anastasiou D., Vázquez L.M. (2010). Localisation Standards and Metadata. In: Sánchez-Alonso S., Athanasiadis I.N. (eds) Metadata and Semantic Research. MTSR 2010. Communications in Computer and Information Science, vol 108. Springer, Berlin, Heidelberg.
- ANZLIC (2001). ANZLIC Metadata Working Group on Metadata. ANZLIC Metadata Guidelines: Core Metadata Elements for geographic data in Australia and New Zealand, version 2 (2001). <http://www.ga.gov.au/>. Accedido 25-5-2018.
- Apache (2018). Apache Jena. The Apache Software Foundation. <https://jena.apache.org>. Accedido 11-2018.

- Avato, J., Koettl, J. y Sabates-Wheeler, R. (2010). Social security regimes, global estimates and good practices: The status of social protection for international migrants. *World Development*, Vol. 28, No. 4.
- Baek, J-E., Sugimoto, S. (2011). Facet Analysis of Archival Metadata Standards to Support Appropriate Selection, Combination and Use of Metadata Schemas. *International Conference on Dublin Core and Metadata Applications*, [S.l.], p. 1-11, sep. 2011. ISSN 1939-1366. <http://dcpapers.dublincore.org/pubs/article/view/3620/1846>. Accedido 26-5-2018.
- Batlle-Montserrat, J., Abadal, E., y Blat, J. (2011). Benchmarking del e-gobierno local: limitaciones de los métodos de evaluación comparativa. *El profesional de la información*, 2011, mayo-junio, v. 20, n. 3, pp. 251-259.
- BOE (2013). Esquema nacional de interoperabilidad. Resolución de 19 de febrero de 2013, BOE de 4 de marzo de 2013.
- Brickley, D., Miller, L. (2014) FOAF Vocabulary Specification. <http://xmlns.com/foaf/spec/>. Accedido 7-2018.
- Bueno de la Fuente, G. (2018). Universidade do Minho. Departamento de Sistemas de Informação. Análisis de la interoperabilidad entre los sistemas de apoyo a la formación de TecMinho. <https://e-archivo.uc3m.es/handle/10016/9089>. Accedido 22-12-2018.
- CARICOM (2001). Tratado de Chaguaramas revisado por el que se establece la Comunidad del Caribe con inclusión del Mercado único y la economía de la CARICOM. Comunidad y Mercado Común del Caribe. <https://caricom.org/documents/>. Accedido 26-5-2018.
- Carter, C. H. (2009). Las personas primero: reformas de los servicios de seguridad social para mejorar las condiciones de las personas. *ISSA International Conference on ICT in Social Security*. Seville. 2009.
- CBSS (2018). Social Security Crossroads Bank. <https://www.ksz-bcss.fgov.be/en>. Accedido 17-3-2018.
- CEN (2012). CEN-CENELEC-ETSI Smart Grid Coordination Group. Smart Grid Reference Architecture. <https://ec.europa.eu/energy/>. Accedido 13-3-2018.
- CEPAL (2007). Naciones Unidas. Comisión Económica para América Latina y el Caribe. Libro blanco de interoperabilidad de gobierno electrónico para América Latina y el Caribe. <http://repositorio.cepal.org/>. Accedido 27-5-2018.
- CEPAL (2015). Naciones Unidas. Comisión Económica para América Latina y el Caribe. Libro blanco de interoperabilidad de gobierno electrónico para América Latina y el Caribe. Interoperability Focus: About. <http://www.ukoln.ac.uk/interop-focus/about>. Accedido 27-5-2018.

- Chappell, Dave (2004). Enterprise Service Bus. O'Reilly: June 2004. ISBN 0-596-00675-6.
- Christian, Eliot (2001). A metadata initiative for global information discovery. Government Information Quarterly. 18 (2001) 209-221.
- CIM (2011). Convenio Multilateral Iberoamericano de Seguridad Social. <http://www.seg-Social.es/>.
Accedido 29-5-2018
- Clark, T. y Jones, R. (1999). Organisational Interoperability Maturity Model for C2. https://resources.sei.cmu.edu/asset_files/WhitePaper/. Accedido 3-6-2018.
- Coyle, K. (2012). Linked Data Tools: Connecting on the Web. Library technology reports: expert guides to library systems and services 0024-2586 ; v. 48, no. 4. Library technology reports, ISSN 0024-2586. American Library Association, 2012.
- COM (2017). Comisión europea. Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social europeo y al Comité de las regiones. Marco Europeo de Interoperabilidad – Estrategia de aplicación. <https://ec.europa.eu/transparency/>. Accedido 29-5-2018.
- Craig, A. (2001). The find-it! Illinois controlled vocabulary: Improving access to government information through the jessica subject tree. Government Information Quarterly, 18(3), 195-208.
- Criado, j., Gascó, M. y Jiménez, C. (2010). Bases para una Estrategia Iberoamericana de Interoperabilidad. Documento para la consideración de la XII Conferencia Iberoamericana de Ministros de Administración Pública y Reforma del Estado. Buenos Aires, Argentina, 1-2 de julio de 2010. <http://siare.clad.org/>. Accedido 29-5-2018.
- DAMA (2009). The Data Management Association. The DAMA Guide to the Data Management Body of Knowledge. Technics Publications, LLC. ISBN 978-0-9771400-8-4.
- DB (2018) System Properties Comparison AllegroGraph vs. Jena vs. Virtuoso. DB-Engines. <https://db-engines.com/en/system/AllegroGraph%3BJena%3BVirtuoso>. Accedido 10/2018.
- DCMI (2016). DCMI Metadata Terms. Dublin Core Metadata Initiative. <http://dublincore.org/>.
Accedido 07/2018.
- DCMI (2018). Dublin Core Metadata Initiative. <http://www.dublincore.org/specifications/>. Accedido 24-3-2018.
- De Wispelaere, F., Pacolet, J. (2017). Coordination of social security systems at a glance. 2017 Statistical Report. European Union, 2017.
- DDC (2013). Deutsche Post AG. Death Data Community. <https://deathdatacommunity.rentenservice.com/>. Accedido 24-7-2017.

- Delgado, F., Hilera, JR., Ruggia, R. (2012). Soluciones para el intercambio electrónico de información de Seguridad social a nivel internacional. *El profesional de la Información*. 2012, Julio-agosto.
- Delgado, F., Otón, S., Ruggia, R., Hilera, JR. Y Gutiérrez, JM. (2013). Proceedings of the 15th international conference on enterprise information systems vol. 2 (ICEIS 2013, Angers, Francia, 4-7 Julio 2013).
- Delgado, F., Oton, S., Ruggia, R., Hilera, JR. y Barchino, R. (2013b). Improving information system interoperability in social sector through advanced Metadata. 22nd International Conference on Information Systems Development (ISD2013), Sevilla, Spain, September 2-4, 2013.
- Delgado, F., Hilera, JR. y Ruggia, R. (2013c). Proposal of a controlled vocabulary to solve semantic interoperability problems in social security information exchanges. *Library Hi Tech*, Vol. 31 Iss: 4, pp.602 - 619.
- Delgado, F. (2013d). Modelos y estándares para la Interoperabilidad en los intercambios de información entre sistemas de Seguridad Social. V Congreso Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones Avanzadas (ATICA 2013). Huancayo (Perú), 2-4 de Octubre de 2013.
- Delgado, F., Hilera, J., Ruggia, R. y Otón, S. (2014). A Semantic Interoperability Standard for social security information exchange. 12th European Conference oneGovernment (EGEC 2014).
- Delgado, F (2017). Improving Management of Social Security International Agreements. Intercultural Relations and Migration Processes. Nova Science Publisher, New York. ISBN: 978-1-53610-778-4.
- Delgado, F. (2018). Propuesta de estándar para mejorar los aspectos semánticos en los intercambios de datos de Seguridad Social. IX Congreso Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones Avanzadas y Accesibilidad (ATICA 2018). Ciudad de Guatemala (Guatemala), 7 al 9 de Noviembre de 2018. ISBN: 978-84-17729-63-9.
- Delgado, F., Hilera, J.R., Ruggia, R., Otón, S., Amado-Salvatierra, H.R. (2018b). Using microdata for international e-Government data exchange: The case of Social Security domain. Enviado a *Journal of Information Science*. Enviado el 19/12/2018, pendiente de revisión.
- Dimechkie , K. (2013). Support Free Movement of Persons & Migration in West Africa. International Labour Organization. <https://www.aau.org/>. Accedido 25-5-2018.
- Ding, Y. and Schubert, F. (2002). Ontology research and development. Part 1 - A review of ontology generation. *Journal of Information Science*. Vol 28, Issue 2, pp. 123 – 136. First Published April 1, 2002.
- DOUE (1971). Unión europea. Reglamento (CEE) 1408/71 del Consejo de 14 de junio de 1971, relativo a la aplicación de los regímenes de Seguridad Social a los trabajadores por cuenta ajena, a los trabajadores por cuenta propia y a los miembros de sus familias que se

- desplazan dentro de la Comunidad. Diario Oficial de la Unión Europea, DOUE número 149 de 5/7/1971, páginas 2 a 50.
- DOUE (1972). Unión europea. Reglamento (CEE) 574/72 del Consejo, de 21 de marzo de 1972, por el que se establecen las modalidades de aplicación del Reglamento 1408/71 relativo a la aplicación de los regímenes de seguridad social a los trabajadores por cuenta ajena y a sus familiares que se desplacen dentro de la Comunidad. Diario Oficial de la Unión Europea, DOUE número 74 de 27/3/1972, páginas 1 a 83.
- DOUE (2003). Comisión europea. Decisión nº 189 de junio de 2003 dirigida a sustituir por una tarjeta sanitaria europea los formularios necesarios para la aplicación de los Reglamentos (CEE) Nº 1408/71 y (CEE) nº 574/72 del Consejo en lo que respecta al acceso a la asistencia sanitaria durante una estancia temporal en un Estado miembro distinto al Estado competente o de residencia. Diario Oficial de la Unión Europea, DOUE número 276 de 27/10/2003, páginas 1 a 3.
- DOUE (2010). Comisión Europea. Decisión Nº E2 de 3 de marzo de 2010 relativa a la instauración de un procedimiento de gestión de los cambios introducidos en los datos de los organismos definidos en el artículo 1 del Reglamento (CE) nº 883/2004 del Parlamento Europeo y del Consejo que se enumeran en el directorio electrónico que forma parte integrante de EESSI. Diario Oficial de la Unión Europea, DOUE de 10/7/2010, C 187/5 y 6.
- Dublin Core (2009). Dublin Core Metadata Initiative. <http://dublincore.org/metadata-basics>. Accedido 29-5-2018.
- Dublin Core (2018). Dublin Core Metadata Initiative. <http://dublincore.org/interoperability-levels/>. Accedido 27-2-2018
- Durán-Valverde, F. (2012). Coordinate social protection programmes in the context of the creation of social protection floors. ISSA, 13th Int. Conf. ICT in Social Security, Brasilia, April 2012.
- e-GIF (2005). e-Government Interoperability Framework. UK e-Government Unit <http://xml.coverpages.org/egif-UK.html>. Accedido 29-5-2018.
- EC (1997). European Commission, Directorate-General V (employment, industrial relations and social affairs). Information note on the TESS programme. Version 2.1, 12/09/97. Texto completo en anexo 2.
- EC (2004). European Commission. European Interoperability Framework for Pan-European eGovernment Services. Official Publications of the European Communities, 2004.
- EC (2005). European Commission. Decision No 202 of 17 March 2005 on model forms necessary for the application of Council Regulations (EEC) No 1408/71 and (EEC).
- EC (2018). EESSI Public Directory of European Social Security Institutions. European Commission. http://ec.europa.eu/employment_social/social-security-directory/. Accedido 07-2018.

- EESSI (2010). European Commission. New Social Security Institution. EESSI Newsletter, 2010, n. 3, pp. 1. <http://ec.europa.eu/social/>. Accedido 14-5-2016.
- EESSI (2012). European Commission. Electronic Exchange of Social Security Information (EESSI). <http://ec.europa.eu/social/main.jsp?catId=869>. Accedido 30-5-2018.
- EIF (2004). European Commission. Interoperable Delivery of European eGovernment Services to public Administrations, Businesses and Citizens (IDABC). EIF - European Interoperability Framework for pan-European eGovernment services. <http://ec.europa.eu/idabc>. Accedido 30-5-2018.
- EIF (2010). European Commission. Interoperability Solutions for European Public Administrations - ISA. European Interoperability Framework (EIF). 2010. <http://ec.europa.eu/isa/documents/>. Accedido 7-8-2017.
- EIF (2017). Comisión europea. Marco europeo de Interoperabilidad. Estrategia de aplicación. [https://ec.europa.eu/isa2/eif en](https://ec.europa.eu/isa2/eif_en). Accedido 23-8-2017.
- EIS (2012). EUROPEAN COMMISSION, Directorate General for Informatics. Supporting the European Interoperability Strategy Elaboration, Final Report Phase 1 – 02/07/2009. <http://ec.europa.eu/idabc/servlets/>. Accedido 30-5-2018.
- ENI (2007). Gobierno de España. Ministerio de la Presidencia. Esquema Nacional de Interoperabilidad. <https://administracionelectronica.gob.es/ctt/eni#descripcion>. Accedido 30-5-2018.
- EPOC (2004). The European Union's Judicial Cooperation Unit. EPOC IV Project. <http://www.eurojust.europa.eu/Practitioners/operational/EpocIV/>. Accedido 9-4-2018.
- Ercegovac, Z. (1999). Introduction. Journal of the American Society for Information Science 50(13), 1165–1168 (1999).
- EU (1993). European Commission. European Procurement Handbook for Open Systems. <https://publications.europa.eu/en/publication-detail/>. Accedido 14-3-2018.
- EU (2015). European Commission. e-Government Core Vocabularies handbook. Publications Office of the European Union. <https://ec.europa.eu/isa2/sites/>. Accedido 7-2018.
- EU (2017). European Commission. Core vocabularies. <https://joinup.ec.europa.eu/asset/>. Accedido 30-5-2018.
- EU (2018). EUR-Lex. Access to European Union law. European Union. <https://eur-lex.europa.eu>. Accedido 7-2018.
- EU (2018b). European Commission. EESSI Public Directory of European Social Security Institutions. http://ec.europa.eu/employment_social/social-security-directory. Accedido 6-11-2018.

- Eurojust (2002). The European Union's Judicial Cooperation Unit. <http://www.eurojust.europa.eu/>. Accedido 9-4-2018.
- FEAF (2013). Federal Enterprise Architecture Framework Federal Enterprise Architecture Framework. 29-1-2013. <https://obamawhitehouse.archives.gov>. Accedido 14-3-2018.
- Frantz, R., Corchuelo, R., Roos-Frantz, F. y Sawicki, S. A technology proposal to realise enterprise application integration. International. Journal of Computer Research; Hittington. Tomo 22, N.º 3, (2015): 307-328.
- Fraunhofer (2008). Fraunhofer Institute for Software and Systems Engineering. Study on multilingualism. European Commission – IDABC. 2008.
- Fraunhofer (2009). Fraunhofer Institute for Software and Systems Engineering. Semantic Interoperability Centre Europe: Study on Methodology. <http://www.semic.eu/semic-eu-study-on-methodology-v1.2.pdf>. Accedido 7-8-2015.
- GCC (2006). The Unified Law of Insurance Protection Extension for the Gulf Cooperation. Organization: General Organization for Social Insurance. Saudi Arabia. 2006 . ISSA Good Practice. https://www.issa.int/es_ES/good-practices/. Accedido 31-5-2018.
- GEM (2018). U.S. Department of Education's Gateway to Educational Materials. Gateway to Educational Materials. <http://guides.ucf.edu/>. Accedido 31-5-2018.
- GIF (2007). United Nations Development Programme. e-Government Interoperability: Guide. <http://www.apdip.net/projects/gif/>. Accedido 4-10-2016.
- Gómez Hermosillo, Rogelio (2011). La efectividad de las redes de protección social: El rol de los sistemas integrados de información social en México. Banco Interoamericano de desarrollo. <https://publications.iadb.org/bitstream/>. Accedido 26-5-2018.
- GOSIP (1991). National Institute of Standards and Technology. Government Open Systems Interconnection Profile Users' Guide, Version 2. 1991.
- Guijarro, L. (2007). Interoperability frameworks and enterprise architectures in egovernment initiatives in Europe and the United States. Communications Department. Technical University of Valencia.
- Harpring, P. (2010). Introduction to Controlled Vocabularies: Terminology for Art, Architecture, and Other Cultural Works. Published on www.getty.edu in 2010 by the Getty Research Institute, Los Angeles. ISBN 978-1-60606-026-1. <http://www.getty.edu/research/publications>. Accedido 31-5-2018.
- Hartmann, U. (1990). Open Systems Standards: Status of International Harmonization and European Activities. Proceedings of the 6th International Conference on the Application of Standards for Open Systems, pp. 18.26, California: IEEE Computer Society Press. 1990.

- Hellberg, A-S., Grönlund, A. (2013). Conflicts in implementing interoperability: Re-operationalizing basic values. *Government Information Quarterly*. 30 (2013) 154 – 162.
- Hilera, J.R., Pagés, C., Martínez, J. J., Gutiérrez, J. A. y De-Marcos, L. (2010). An evolutive process to convert glossaries into ontologies. *Information Technology and Libraries*, 29(4), 195-204.
- Hípola, P. (1992). Sistemas EDI: los ordenadores aumentan su protagonismo en las comunicaciones de datos. *El Profesional de la información*. Noviembre 1992. ISSN 1386-6710.
- Hirose, K., Milos, N. y Tamagno, E. (2011). Social security for migrant workers: a rights-based approach. International Labour Organization, Decent Work Technical Support Team and Country Office for Central and Eastern Europe. Budapest: ILO, 2011. ISBN: 9789221255208; 9789221255215.
- HL7 (2017). Health Level Seven International. <http://www.hl7.org>. Accedido 31-5-2018.
- IC (2008). CEN, CENELEC, ETSI and ANSI. Conference on Interoperability: Key to International Business. Warsaw (Poland) 6-7 February 2008. <http://www.interoperabilityconference.org/>. Accedido 7-8-2015.
- IEEE (1990). IEEE Standard Computer Dictionary. <http://ieeexplore.ieee.org/document/182763/>. Accedido 13-3-2018.
- IEEE/LOM (2017). Information technology -- Learning, education and training -- Metadata for learning resources. <http://www.aenor.es/normas/>. Accedido 31-5-2018.
- IF (2017). The Government of the Hong Kong Special Administrative Region. The HKSARG Interoperability Framework. <https://www.ogcio.gov.hk/>. Accedido 14-3-2018.
- ILO (2011). International Labour Organization. Sharing Innovative Experiences: Successful Social Protection Floor Experiences . <http://www.ilo.org/wcmsp5/groups/>. Accedido 31-5-2018.
- ILO NATLEX (2014). Database of national labour, social security and related human rights legislation. Geneva, International Labour Office. <http://www.ilo.org/dyn/natlex>. Accedido 31-5-2018.
- Indarte (2011). Indarte, S. y Pazos, P. (2011). Estándares e interoperabilidad en salud electrónica: Requisitos para una gestión sanitaria efectiva y eficiente. CEPAL, 2011. <http://www.cepal.org/es/publicaciones/>. Accedido 21-8-2017.
- INSPIRE (2017). European Commission. Infrastructure for spatial information in Europe. <https://inspire.ec.europa.eu/>. Accedido 31-5-2018.
- INSS (1993). Reglamentos 1408/71 y 574/72 de la Comunidad Europea en materia de Seguridad Social. Normativa complementaria y Decisiones de la Comisión Administrativa. Instituto Nacional de Seguridad Social. ISBN: 84-87140-04-1.

- INTEROPS (2011). Le standard d'interopérabilité des organismes de protection sociale. <http://www.interops.fr/>. Accedido 14-6-2017.
- ISA (2011). European Commission. Interoperability Solutions for European Public Administrations - ISA. Towards Open Government Metadata. 2011. <https://joinup.ec.europa.eu/>. Accedido 31-5-2018.
- ISA (2014). European Commission. Interoperability Solutions for European Public Administrations - ISA. <http://ec.europa.eu/isa/documents/>. Accedido 22-8-2017.
- ISO (1998). ISO 639-3:1998 Codes for the representation of names of languages -- Part 2: Alpha-3 code. International Organization of Standardization.
- ISO (2002). International Organization for Standardization. Language codes - ISO 639. <https://www.iso.org/iso-639-language-codes.html>. Accedido 29-3-2018.
- ISO (2007). ISO 639-3:2007 Codes for the representation of names of languages -- Part 3: Alpha-3 code for comprehensive coverage of languages. International Organization of Standardization.
- ISO (2009). International Organization for Standardization. ISO 15836:2009. Information and documentation. The Dublin Core metadata element set. Geneva, Switzerland.
- ISO (2017). International Organization for Standardization. ISO/IEC 11179, Information Technology - Metadata registries (MDR). <http://metadata-standards.org/11179/>. Accedido 13-6-2017.
- ISSA (2012). Technical Commission on Information and Communication Technology. Interoperability in social security. 13th Int. Conf. ICT in Social Security, Brasilia, April 2012.
- ITF (2003). Australian Government Information Management Office. Interoperability Technical Framework for the Australian Government, Version 2, 2003. <https://www.finance.gov.au/>. Accedido 1-6-2018.
- Janowski, T., Pardo, T.A., Davies, J. (2012). Government Information Networks – Mapping Governance cases through Public Administration concepts. Government Information Quarterly, 29 (2012) S1-S10.
- Kim, H., MacDonald, RH, y Andersen, DF. (2013). Simulation and Managerial Decision Making: A Double-Loop Learning Framework. Public Administration Review 73(2): 291-300.
- Kounowski, G. (2012). Integration of social security programmes and the role of ICTs: case studies. 13th Int. Conf. ICT in Social Security, Brasilia, 2012.
- Lee-Archer, B., Brailey, C., Le Noir, M. and Ziehm, O. (2007). For the good of the global economy Social protection for the migrant worker. IBM Institute for Business Value.

- LC (2018). Library of Congress. Linked Data Service. <http://id.loc.gov/vocabulary/iso639-1.html>.
Accedido 8-11-2018
- LISI (2018). BestManufacturing Practices. LISI Model: Levels Of Information Systems Interoperability (LISI) Reference Model <http://www.bmpcoe.org/library/>. Accedido 22-12-2018.
- Löbel, G., Schmid, H. y Müller, P. (1982). Glosario de Informática. ISBN: 9788423402809.
- Manouselis, N. y Costopoulou, C. (2006). Quality in metadata: a schema for e-commerce. Online Information Review, Vol. 30 Issue: 3, pp.217-237. <https://doi.org/>. Accedido 1-6-2018.
- Manso-Callejo M., Wachowicz M. y Bernabé-Poveda M. (2010.) The Design of an Automated Workflow for Metadata Generation. In: Sánchez-Alonso S., Athanasiadis I.N. (eds) Metadata and Semantic Research. MTSR 2010. Communications in Computer and Information Science, vol 108. Springer, Berlin, Heidelberg.
- MARC21 (2010). Biblioteca del Congreso. Formato MARC 21 para registros bibliográficos. NIPO: 032-15-027-1. <http://www.bne.es/es/Micrositios/Guias/>. Accedido 9-4-2018.
- Marcos-Martin, C. y Soriano-Maldonado, SL. (2011). Reutilización de la Información del sector público y open data en el contexto español y europeo. Proyecto Aporta. El profesional de la información, 2011, mayo-julio, v. 20 n. 3.
- Mehta, B. (2014). RESTful java patterns and best practices. Packt Publishing Ltd. ISBN 978-1-78328-796-3.
- McGillivray (2010). Strengthening social protection for African migrant workers through social security agreements. Background report prepared for the extension of social security coverage to African migrant workers (MIGSEC) project. Ginebra, Oficina Internacional del Trabajo.
- Mercosur (1991). Tratado para la constitución de un mercado común (Tratado de Asunción). <http://www.mercosur.int/innovaportal/>. Accedido 3-3-2018.
- Morris, E., Levine, L., Meyer, C., Place, P. and Plakosh, P. (2004). System of Systems Interoperability (SOSI): Final Report. Carnegie Mellon University and Software Engineering Institute. <https://resources.sei.cmu.edu/>. Accedido 22-12-2018.
- Munk, S. (2002). Interoperability in the Infosphere. Challenges, Problems, Solutions. Obrana a Strategie, 2002.
- MusicBrainz (2018). Open music encyclopedia. <https://musicbrainz.org/>. Accedido 24-3-2018.
- NBII (2018). Stanford University. National Biological Information Infrastructure. <https://wayback.archive-it.org/>. Accedido 24-3-2018.

- NISO (2004). National Information Standards Organization. Understanding Metadata. NISO Press, 2004. ISBN: 1-880124-62-9.
- Noh, Y. (2011). A study on metadata elements for web-based reference resources system developed through usability testing", Library Hi Tech, Vol. 29 Issue: 2, pp.242-265, <https://doi.org/10.1108/>. Accedido 8-12-2017.
- Novakouski, M., Lewis, GA. (2012). Interoperability in the e-Government Context. SEI, Carnegie Mellon University, 2012.
- NPS (2016). Data exchange on overseas beneficiaries between social security institutions for the prevention of illegal pension payments. National Pension Service. Republic of Korea. ISSA Good Practices, 2016. https://www.issa.int/es_ES/good-practices/. Accedido 3-6-2018.
- NTI (2013). Ministerio de Hacienda y Administraciones Públicas. Secretaría de Estado de Administraciones Públicas. Normas técnicas de interoperabilidad. <http://administracionelectronica.gob.es/>. Accedido 3-6-2018.
- NZ e-GIF (2008). New Zealand E-Government Interoperability Framework v. 3.3. <http://ict.govt.nz/>. Accedido 7-8-2015.
- OISS (2007). Convenio Multilateral Iberoamericano de Seguridad Social. Organización Iberoamericana de Seguridad Social, 2007. <http://www.oiss.org/>. Accedido 3-6-2018.
- OIT (1982). Organización Internacional del Trabajo. C157 - Convenio sobre la conservación de los derechos en materia de seguridad social, 1982 (núm. 157). <http://www.ilo.org/dyn/>. Accedido 1-3-2018.
- OIT (1983). Organización Internacional del Trabajo. R167 - Recomendación sobre la conservación de los derechos en materia de seguridad social, 1983 (núm.@167). <http://www.ilo.org/dyn/>. Accedido 1-3-2018.
- OIT (2006). Organización Internacional del Trabajo. Marco multilateral de la OIT para las migraciones laborales. <http://www.ilo.org/wcmsp5/>. Accedido 1-3-2018.
- OIT (2012b). Organización Internacional del Trabajo. Ensuring social security benefits for Ukrainian migrant workers: Policy development and future challenges. Kiev, ILO Decent Work Technical Support Team and Country Office for Central and Eastern Europe. <http://www.ilo.org/wcmsp5/>. Accedido 4-6-2018.
- OIT (2014). Organización Internacional del Trabajo C102 - Social Security (Minimum Standards) Convention, No.102, 1952. <http://www.ilo.org/>. Accedido 4-6-2018.
- OJEU (2004). European Union. Regulation (EC) No 883/2004 of the European Parliament and of the Council of 29 April 2004 on the coordination of social security systems. Official Journal of the European Union, L166, 30-4-2004.

- OJEU (2009). European Union. Regulation (EC) No 987/2009 of the European Parliament and of the Council of 16 September 2009 laying down the procedure for implementing Regulation (EC) No 883/2004 on the coordination of social security systems. Official Journal of the European Union, L284, 30-10-2009.
- OMB (2011) . The White House. The Office of Management and Budget (OMB). OMB Circular A-133 Compliance Supplement 2011. <https://obamawhitehouse.archives.gov/>. Accedido 4-6-2018.
- OMG (2014). Object Management Group. Ontology definition metamodel. <https://www.omg.org/>. Accedido 4-6-2018.
- OMG (2017). Object Management Group. <http://www.omg.org>. Accedido 20-9-2017.
- OMG2 (2017). Object Management Group. Specifications updated recently. <http://www.omg.org/>. Accedido 4-6-2018.
- OpenAPI (2018) The OpenAPI Specification. OpenAPI Initiative. The Linux Foundation. <https://www.openapis.org>. Accedido 11-2018.
- OpenLink (2018). Virtuoso Universal Server Documentation. OpenLink. <http://docs.openlinksw.com/virtuoso>. Accedido 10-2018.
- PAI (2015). European Commission. EESSI Public Directory of European Social Security Institutions, Public access Interface. http://ec.europa.eu/employment_social/. Accedido 17-3-2018.
- Park (2004). Park, J. y Ram, S. (2004). Information Systems Interoperability: What Lies Beneath? ACM. Transactions on Information Systems, Vol. 22, No. 4. 2004.
- Pautasso, C., Zimmermann, O., & Leymann, F. (2008). Restful web services vs. big'web services: Making the right architectural decision. Proceedings of the 17th International Conference on World Wide Web, 805-814.
- Pekka, R. (2013). Data modelling in EESSI – project. EESSI Working Party in Vilnius (Lithuania), 14-11-2013. Documento completo en anexo 2.
- Peristeras, V., Tarabanis, K. and Goudos, SK. (2009). Model-driven eGovernment interoperability: A review of the state of the art. Computer Standards & Interfaces 31 (2009) 613-618.
- Pons, D., Hilera, JR. y Pagés-Arévalo, C. (2011). ISO/IEC 19788 MLR: Un Nuevo Estándar de Metadatos para Recursos Educativos. IEEE-RITA 6 (2011): 140-145.
- POSIT (1995). National Bureau of Standards. Federal Information Processing Standards Publication: profiles for open systems internetworking technologies. <https://archive.org/details/>. Accedido 13-3-2018.
- Prytherch, R. (1999). Harrod's Librarians' Glossary and Reference Book. Gower Publishing Company. 1999.

- Ratha, D., Shaw, W. (2007). South-South migration and remittances (Working paper, No. 102). Washington, DC, World Bank.
- RAE (2018). Real Academia Española. <http://dle.rae.es/>. Accedido 22-3-2018.
- RDF (2004). W3C Semantic web. Resource Description Framework. <http://www.w3.org/RDF/>. Accedido 4-6-2018.
- RNCPS (2006). Répertoire National Commun de la Protection Sociale. 2006. <http://www.securite-sociale.fr/Repertoire-National-Commun-de-la-Protection-Sociale-RNCPS>. Accedido 4-6-2018.
- Sanati, F. and Lu, J. (2010). Life-event modelling framework for e-government integration. *Electronic Government an International Journal* 7(2):183-202 · January 2010.
- Schema (2018). Schema.org vocabulary. Schema.org community. <https://schema.org>. Accedido 7-2018.
- Schutte, M. (2009). The semantic web for knowledge and data management: Technologies and practices. *Online Information Review*, 33(3), 621-623. doi:10.1108/14684520910970059.
- SEMIC (2012). Semantic Interoperability Community Core Person Vocabulary. <https://joinup.ec.europa.eu/release/>. Accedido 7-2018.
- SEMIC (2018). Semantic Interoperability Community. Core Vocabularies. <https://joinup.ec.europa.eu/collection/>. Accedido 7-2018.
- Shadbolt, N., Hall, W., and Berners-Lee, T. (2006). The semantic web revisited. *Intelligent Systems, IEEE*, 21(3), 96-101.
- Sheth, AP. and Kashyap, V. (1993). So Far (Schematically) yet So Near (Semantically). In *Proceedings of the IFIP WG 2.6 Database Semantics Conference on interoperable Database Systems (Ds-5) (November 16 - 20, 1992)*. D. K. Hsiao, E. J. Neuhold, and R. Sacks-Davis, Eds. IFIP Transactions, vol. A-25. North-Holland Publishing Co., Amsterdam, The Netherlands. 1993. <https://corescholar.libraries.wright.edu/>. Accedido 4-6-2018.
- Sicilia, MA. (2006). Metadata, semantics and ontology: providing meaning to information resources. *International Journal of Metadata, Semantics and Ontologies*, vol. 1 No 1, pp. 83-6.
- SOAP (2007). W3C. Simple Object Access Protocol. <http://www.w3.org/TR/soap/>. Accedido 4-6-2018.
- Tambouris, E., Tarabanis, K., Peristeras, V., y Liotas, N. (2007). Study on Interoperability at Local and Regional Level. Final Version – Version 2.0. eGovernment Unit DG Information Society and Media, European Commission. 2007.

- TF (2007). European Commission. Task Force on electronic data Exchange. Nota TF 063/07: Build 6 Outcomes and the EESSI Project. Texto completo en anexo 2.
- TOGAF (2008). The Open Group. The Open Group Architecture Framework. <http://pubs.opengroup.org/architecture/>. Accedido 4-6-2018.
- Tolk, A. (2003). Beyond Technical Interoperability – Introducing a Reference Model for Measures of Merit for Coalition Interoperability. 8th International Command and Control Research and Technology Symposium. <http://www.dtic.mil/>. Accedido 1-6-2018.
- UDDI (2004). OASIS. Universal Description, Discovery and Integration. <http://www.oasis-open.org/committees/>. Accedido 4-6-2018.
- UKOLN (2013). UKOLN. Interoperability Focus: About. <http://www.ukoln.ac.uk/interop-focus/>. Accedido 4-6-2018.
- UN (2015). United Nations. International migrant stock 2015. <http://www.un.org/>. Accedido 28-2-2018.
- Verborgh, R., Harth, A., Maleshkova, M., Stadtmüller, S., Steiner, T., Taheriyani, M., & Van de Walle, R. (2014). Survey of semantic description of rest APIs. In rest: Advanced Research Topics and Practical Applications (pp. 69-89). Springer New York.
- Viñán, M., Otón, S., Delgado, F., Ruggia, Raul. (2014). Interoperabilidad en Instituciones de Seguridad Social. Actas del VI Congreso Internacional sobre Aplicación de Tecnologías de la Información y Comunicaciones Avanzadas (ATICA2014). Universidad de Alcalá. Alcalá de Henares (España) 29 - 31 de octubre de 2014. ISBN: 978-84-16133-42-0.
- W3C (2008). World Wide Web Consortium. Extensible Markup Language. World Web Consortium, 2008. <http://www.w3.org/XML/>. Accedido 4-6-2018.
- W3C (2009). World Wide Web Consortium. Web Services. <https://www.w3.org/WebServices>. Accedido 12-4-2018.
- W3C (2012). World Wide Web Consortium. ISA Programme Person Core Vocabulary. <https://www.w3.org/ns/person>. Accedido 9-4-2018.
- W3C (2013). World Wide Web Consortium. Registered Organization Vocabulary. <https://www.w3.org/TR/vocab-regorg>. Accedido 4-6-2018.
- W3C (2014). RDF 1.1 Concepts and Abstract Syntax. World Wide Web Consortium (W3C). <https://www.w3.org/TR/>. Accedido 7-2018.
- W3C (2015). World Wide Web Consortium. ISA Programme Location Core Vocabulary. Second version in w3.org/ns space - 2015-03-23. <https://www.w3.org/ns/locn>. Accedido 9-4-2018.

- W3C (2015a) World Wide Web Consortium. RDFa 1.1 Primer - Third Edition. Rich Structured Data Markup for Web Documents. <https://www.w3.org/TR/rdfa-primer/>. Accedido 7-2018.
- W3C (2018). World Wide Web Consortium. LargeTripleStores. <https://www.w3.org/wiki/>. Accedido 10-2018.
- W3C (2018a). World Wide Web Consortium. HTML Microdata. <https://www.w3.org/TR>. Accedido 07-2018.
- W3C (2018b). World Wide Web Consortium. JSON-LD 1.1 A JSON-based Serialization for Linked Data. Draft Community Group Report. <https://json-ld.org/>. Accedido 7-2018.
- WS-Security (2004). OASIS. Advancing open standards for the information society. <https://www.oasis-open.org/>. Accedido 17-3-2018.
- WSDL (2007). World Wide Web Consortium. Web Services Description Language. <http://www.w3.org/TR/wsdl>. Accedido 4-6-2018.
- WSBPEL (2007). OASIS. Web Services Business Process Execution Language. <http://docs.oasis-open.org/>. Accedido 4-6-2018.
- Zachman, J. (2008). Zachman International Enterprise Institute. The Zachman Framework. <https://www.zachman.com/about-the-zachman-framework>. Accedido 4-6-2018.
- Zeigler, B., Murzy, A., Yilmaz, L. (2006). Artificial Intelligence in Modelling and Simulation. In: Encyclopedia of Complexity and System Science. Springer, Germany.

Anexos

Anexo 1: propuesta de estandar

Standard proposal

Exchange of Social Security Information Metadata

Warning

This document is a standard proposal to be submitted to the International Organization for Standardization.

He is waiting for his presentation to the international organization.

Introduction

Since the end of the 90s, some industrial, academic and public administration organizations have developed information technology standards in their respective domains, in order to address the different challenges presented by interoperability problems between different organizations.

In the context of Social Security systems at the international level, the challenge is presented in the exchanges of data that support the implementation of international agreements. The magnitude and social impact of the problem justifies the development of a solution that can be applied by any Social Security institution, in order to improve the results obtained so far by the exchange of data in this area.

When data has to be processed by different organizations, semantic interoperability practices are essential to develop common definitions and interpretations.

The linguistic heterogeneity in the countries involved in the exchange of Social Security data is the most obvious reason that supports this assertion. The lack of terminological standardization is another major cause, because Social Security operations use a wide range of terms, which despite having the same or similar name, can be interpreted in a different way. A term should always refer to a single concept, otherwise the relationships between concepts and definitions could become very difficult to maintain. Likewise, the lack of a nomenclature policy leads to disparities and conflicts in the interpretation of the same terms.

These difficulties enter into the field of semantic interoperability, whose main objective is to allow the information resources are connected and are automatically understandable and, consequently, reusable by the computer applications that did not take part in their creation. The fact that the information generated by a computer has to be processed by another system interpreting its meaning correctly, causes a series of additional complications that affect both the source of the information and its receiver.

In similar scenarios, the European Interoperability Strategy proposes a solution consisting of a definition of metadata. A metadata schema of this type should specify the information to be exchanged in conceptual terms, its properties and the relationships between the concepts.

As a precedent, several standards defined in circumstances that bear certain similarities could be quoted.

- ISO 15836: 2017. The Dublin Core metadata element set is a vocabulary of fifteen properties or semantic definitions for the description of resources. It is described in the standard as part of a broader set of metadata vocabulary and technical specifications maintained by the Dublin Core corporation (DCMI).

The Dublin Core metadata is being used as the basis for descriptive systems for various interest groups, such as:

- o Government institutions.
- o Scientific sector of the research.
- o Authors of web pages.
- o Businesses that require activities in the field of research.
- o Corporations with complex knowledge management systems.

- AS5044: 2010. AGLS is the Australian standard, based on Dublin Core and used to describe electronic Government resources. The resources can be documents, images, sound, video, physical objects, people and services. It takes thirteen of its elements in a mandatory way, leaving the rest in a recommended or optional way. It includes a series of extensions up to a total of 60 elements.

- IEEE / LOM (Learning Object Metadata). It is a standard used for the description of learning objects. It was developed and formalized through the IEEE and the Learning Technology Standards Committee. Closely related is the ISO / IEC 19788 Metadata Learning Resource, published in January 2011.

Its set of elements has a correspondence with those of Dublin Core.

- ISO / IEC 11179. Establishes the characteristics of metadata to describe data and the management and administration of this metadata in a metadata record. It is applied to the formulation of representations of data, concepts, meanings and relationships among them to be shared by people and machines, independently of the organization that generates the data. It is not applicable to the physical representation of the data as bits or bytes at the machine level.

The benefits expected from the use of the new metadata standard are focused on solving the semantic interoperability problems that currently arise in the exchanges of data between Social Security institutions in the international arena. In particular:

- o Ensure faster processing of exchanged data.
- o Provide the officials of the sending institutions with identification and other data that the receiving institutions can obtain from their clients.
- o Make it possible to verify the information entered as regards its structure.
- o Make it possible, after receiving a structured electronic message, to identify the persons or any information associated with them and the processing of the information.

Exchange of Social Security Information Metadata (ESSIM)

1.1.Scope

ESSIM provides a set of metadata that should be attached with every single of the data exchanged by Social Security institutions in the international arena.

The specific domain referred to is:

- Data exchanged by competent Social Security institutions that have signed international agreements. The competence in matters of Social Security will depends on the content of the agreements.
- Data related to the concepts that the International Labor Organization recognizes as belonging to Social Security.

1.1.Exclusions

The scope of application of ESSIM does not include the exchange of data that, despite being managed by Social Security institutions, do not correspond specifically to concepts of this domain. This would be the case of data related to human and material resources or on economic aspects of other domains.

1.2 Areas not addressed

EESI does not intend to address organizational or technical interoperability aspects, except in those aspects that are common with semantic interoperability.

2. Normative references

The main reference standard is the ISO 15836: 2017 standard (set of Dublin Core metadata elements).

In addition, the ISO 639 standard is used to identify the languages used in data exchanges.

3. Terms and definitions

For the purposes of this document, the terms and definitions included in ISO 15836: 2017 (DC) are applied, in addition to the following:

- **Metadata:** data about data. In the proposed standard, they try to determine unambiguously the meaning of the data they accompany.
- **Data:** The content of each of the fields or set of fields of an electronic form with its own and recognized meaning.

4. Symbols and abbreviations

The following abbreviations are used in the description of the standard.

AGLS: Australian Government Locator Service

DC: Dublin Core

DCMI: Dublin Core Metadata Initiative

EESSI: Electronic Exchange Social Security Information

GOT: Glossary of Terms

ISA: Interoperability Solutions for European Public Administrations

TESS: Telematics for Social Security

XML: eXtensible Markup Language

5. Content of the ESSIM specification

This specification is based on Dublin Core and includes eleven of its fifteen elements. Six of them are considered as mandatory and seven as optional. In some cases, the name given to the element in DC has been modified in order to adjust more to the intended use in ESSIM.

The values of the data described by the metadata are based on controlled vocabularies.

The European Union has developed some tools for this purpose, such as the Glossary of Terms, the Institution Repository and Core Vocabularies. The extension and application of these tools in a global environment would be fundamental for the application of the standard, although the development of that proposal is considered outside the scope of this specification.

5.1 ESSIM elements

5.1.1 Creator institution

Element specification	
Element name	creatorInstitution
Label	Creator Institution
Definition	Social Security institution primarily responsible for making the data
Obligation	Mandatory
Value range	Institution Repository (Institution)
Multiplicity	1..1
Refines Dublin Core element	Creator

5.1.2 Data Set

Element specification	
Element name	dataset
Label	Data Set
Definition	Core Vocabulary where a similar data concept could be found
Obligation	Optional
Value range	Core Vocabulary (Mapping)
Multiplicity	0..1
Refines Dublin Core element	Subject

5.1.3 Data Term

Element specification	
Element name	dataTerm
Label	Data Term
Definition	Term included in the Glossary of Terms (GOT) that defines the meaning of the data field
Obligation	Mandatory
Value range	Glossary of Terms (Term)
Multiplicity	1..1
Refines Dublin Core element	Title

5.1.4 Date Created

Element specification	
Element name	dateCreated
Label	Date Created
Definition	A point of time associated with the starting point in the lifecycle of data
Obligation	Mandatory
Value range	Literal
Multiplicity	1..1
Refines Dublin Core element	Date

5.1.5 Date Modified

Element specification	
Element name	dateModified
Label	Date Modified
Definition	Date when data was last modified or changed
Obligation	Optional
Value range	Literal
Multiplicity	0..1
Refines Dublin Core element	Date

5.1.6 Description

Element specification	
Element name	description
Label	Description
Definition	An effective description is 25-30 words long or approximately 150 characters
Obligation	Mandatory
Value range	Literal
Multiplicity	1..*
Refines Dublin Core element	Description

5.1.7 Format

Element specification	
Element name	format
Label	Format
Definition	Format of the data
Obligation	Mandatory
Value range	Literal
Multiplicity	1..1
Refines Dublin Core element	Format

5.1.8 Identifier

Element specification	
Element name	identifier
Label	Identifier
Definition	Identifier of the data (field)
Obligation	Optional
Value range	Literal
Multiplicity	0..1
Refines Dublin Core element	Identifier

5.1.9 Language

Element specification	
Element name	language
Label	Language
Definition	Language of the data when not numeric
Obligation	Conditional
Value range	ISO 639
Multiplicity	0..1
Refines Dublin Core element	Language

5.1.10 Mandate

Element specification	
Element name	mandate
Label	Mandate
Definition	Legal reference of the exchanges where data is involved
Obligation	Optional
Value range	EUR-Lex (Regulation)
Multiplicity	0..*
Refines Dublin Core element	Relation

5.1.11 References

Element specification	
Element name	references
Label	References
Definition	Official documents where data is present
Obligation	Optional
Value range	Official Documents (Document)
Multiplicity	0..*
Refines Dublin Core element	Relation

5.1.12 Relation with Terms

Element specification	
Element name	relationWithTerms
Label	Relation with Terms
Definition	Description of relationship between the data and other terms, different of the dataTerm
Obligation	Optional
Value range	- relationTerm: Glossary of Terms (Term) - relationDescription: Literal
Multiplicity	0..*
Refines Dublin Core element	Relation

5.1.13 Sender Institution

Element specification	
Element name	senderInstitution
Definition	Social security institution responsible for sending the data
Obligation	Optional
Value range	Institution Repository (Institution)
Multiplicity	0..1
Refines Dublin Core element	Publisher

3.2.4.14 Social Security Category

Element specification	
Element name	socialSecurityCategory
Label	Social Security Category
Definition	Branch of social security involved in the exchange of data
Obligation	Mandatory
Value space or data type	Institution Repository (Category)
Multiplicity	1..1
Refines Dublin Core element	Coverage

5.2 Associated vocabularies

5.2.1 Literals

In accordance with the recommendations established by the W3C in regard to strings of characters, numbers or dates.

This is the case of the elements "dateCreated" and "dateModified" (dates), as well as the elements "description", "format", "identifier" and "relationDescription" (character strings).

5.2.2 Glossary of Terms (GOT)

Created within the sub-project Build 6 (TESS project) with a view to the institutions improving the processing of the data received from abroad. It contains the definitions of the concepts related to the exchanges of Social Security data.

It is used in the "dataTerm" and "relationTerm" elements.

5.2.3 Institution Repository

Developed by the EESSI project, this directory allows an institution to be able to contact other institutions thanks to the facilities of the services directory and the mechanisms established both nationally and internationally.

In the ESSIM specification, the Institution Repository is used by the "creatorInstitution", "senderInstitution" and "socialsecurityCategory" elements.

5.2.4 Core Vocabularies

Resources defined by the European Commission within the ISA program. It is suggested that the Dataset element refer to this model whenever the field described is included in any of the ones developed so far.

5.2.5 ISO 639

The Dublin Core Metadata Initiative includes in its specification of terms recommended for use with DC, the ISO 639-2 and ISO 639-3 versions as possible languages. As a result, it is possible to reuse this publicly available vocabulary, and reference it from the "language" element of ESSIM.

5.2.6 EUR-Lex

EUR-Lex is an open data service that provides free access to the laws of the European Union and other public documents. This database is updated daily and contains more than three million entries with texts that can be dated until 1951. Each of these documents is registered with analytical metadata, as a reference for its publication, date or keywords.

The EESSI “Mandate” element is designed to store this type of information, given that the applicable legislation is relevant to decide on the process applied to the data exchanged.

5.2.7 Official Documents

Vocabulary where all existing documents and susceptible to be exchanged. This dictionary does not exist at the moment in a formal way.

Anexo 2: textos completos referenciados

AC 370/12REV

Note of the countries Austria, Germany, Netherlands, Switzerland and Liechtenstein of 24 September 2012. EESSI: Necessary improvements for implementation on national level.

AC 370/12REV
EESSI improvements
AUSTRIA, GERMANY, NETHERLANDS, SWITZERLAND, LIECHTENSTEIN – 24.09.2012

Orig.: EN

**ADMINISTRATIVE COMMISSION
FOR THE COORDINATION OF SOCIAL SECURITY SYSTEMS**

Subject: EESSI: Necessary improvements for implementation on national level

**Note of the countries Austria, Germany, Netherlands, Switzerland and Liechtenstein of 24
September 2012**

Referring to the findings of the business analysts, the outcome of the testing and implementation activities the countries Austria, Germany, Netherlands, Switzerland and Liechtenstein would like to propose a few major improvements to the EESSI implementation.

An efficient and affordable implementation on a national level will hardly be possible, if the following suggestions for improvements are not considered.

The Implementing Regulation defines that the exchange of data between member states should, where possible, be processed by electronic means. This general idea implies the following basic principles, which should be applied for all flows and SEDs:

1. Flows, SEDs and corresponding data models must be correctly translated into XML schema definitions (XSDs) in order to achieve a proper validation of data (e.g. definition of libraries of data types, KeyRef relations and facets).
The objective must be: reuse, uniqueness and systematic structuredness of data types. This should be achieved on the basis of proven industrial standards and best practice examples.

2. The message originator can not send a message before having successfully validated the message.
3. The definition of SEDs must be in compliance with all international and national data protection regulations and acts.

A working group should be installed to complete business rules for legal, organisational and technical issues of data structures for data exchange between member states.

AC 12-446es

Nota de la delegación española en la Comisión Técnica de 28 de octubre de 2012. EESSI: mejoras necesarias para su implantación. Aspectos semánticos.

EUROPEAN COMMISSION

EMPL/01994/12 - EN

AC 446/12
EESSI - Necessary improvements for its establishment
SPAIN – 28.10.2012

OR.: ES

**ADMINISTRATIVE COMMISSION
FOR THE COORDINATION OF SOCIAL SECURITY SYSTEMS**

Subject: EESSI - Necessary improvements for its establishment

Semantic aspects

Note from the Spanish delegation in the Technical Commission of 28 October 2012

1. Introduction

In note AC 387/12, discussed at the 54th meeting of the Technical Commission, the Secretariat presented an initial inventory of alternative solutions for the EESSI project. In the presentation of the note, reference was made repeatedly to the European Interoperability Strategy (EIS) as a guide for proposing improvement measures.

The main aim of the EIS, according to its final document, is to establish a general overview of the problem which makes it possible to put forward specific actions to improve public services. In its conclusions, it cites semantic interoperability as the priority on which future actions should focus.

Furthermore, note AC 370/12 REV on the necessary improvements for implementing the EESSI at national level points to a series of problems relating to the SEDs, the definition of XML schemas and the validation of messages, directly related to the difficulties associated with semantic interoperability.

The aim of this note is to examine in further detail the semantic aspects of the EESSI project and propose a discussion of possible solutions to the problems which arise in this field.

2. Semantic interoperability in the EESSI

In the above-mentioned context, semantic interoperability means interoperability which ensures that the meaning of the exchanged information is not lost in the process and is maintained and understood by the persons using it, and by the applications and institutions involved. One of the first objectives suggested by the EIS to ensure this dimension of interoperability is to “agree on dictionaries, semantic core components and taxonomies”. The EESSI does not yet have a dictionary of terms. Although it has a glossary of business terms, its aim is to describe the terminology used in the project, not in the content of the exchanged information.

The only existing precedent is the Glossary of Terms from Build 6 in the former TESS project. At the time, Build 6 defined all the data of the repository for information. First of all, it provided generic definitions. Then it added a specific definition for each Member State so as to facilitate processing by the receiving institution. The note from Task Force 63/07 presented various points for consideration with regard to the possible use of the results of Build 6 for the development of the EESSI project.

More recently, in note 356/11, the Polish Presidency proposed creating a dictionary of personal identification numbers. Once the information was compiled in note 579/11, consideration had to be given to various points relating to the meaning of the fields, the information contained and the expected attributes, thereby confirming the need to examine the semantic aspects of the project in further detail.

The hoped-for benefits of this initiative can easily be extrapolated to a definition of all exchangeable fields. In other words, and as mentioned in the above-mentioned note, this initiative should:

- ensure the faster processing of all the cases dealt with under the regulations on coordination of the social security systems;

- provide officials of the issuing institutions with data (identification and other types of data) which the receiving institutions can obtain from their clients;
- make it possible to check the information introduced in aspects concerning its structure;
- make it possible, after receiving an SED, to identify the persons or any information associated with them and to enable the processing of information.

Furthermore, during the current development of the EESSI project, information on the content of the fields can be found on various occasions. The following is a non-exhaustive list:

Master Directory

- The Excel sheet for data entry contains the description, format and length of the fields.
- Notes CASSTM 08-178rev and 11-241: contain the extended description, format, length, obligatory nature and lists of accepted codes.

Documents generated by the Ad hoc groups on SED and flows

- Note CASSTM 09-287 and its three annexes contain information on the details which make up each of the SEDs, lists of values, articles of the Regulation to which they refer, etc.
- Notes 08-011 to 08-16, and 08-31 with their annexes contain diagrams of flows, tables with a description of details and SEDs which use them, etc.
- Flow tables relate the flows to the articles of the Regulation, the SEDs, the dispatch address of each SED, and description.
- Correlation tables relate the SEDs of each flow to their equivalent E form.
- Guidelines contain instructions on using SEDs and flows.
- Explanatory notes with instructions on completing some fields of the corresponding SED.

Technical documentation

- UML model: the conclusions of the ad hoc groups have been converted into the class diagrams which form the UML model of the project. These diagrams contain the classes into which the data for each SED have been divided, with a description, the lists of values allowed in the field, etc. It can be assumed that using the UML model and by means of the Modelling Extraction Tool, the XML schemas, HTML documentation and Xforms documents which shape the Webic displays are generated automatically.
- XML schemas: for each SED, the structure and sequence of the elements which comprise it, names of elements and lists of allowed values are included.
- HTML documents: descriptions of flows (name, version and corresponding SEDs), SED (name, version and details) and details (name, version, allowed values, restrictions and obligatory nature, etc.).

Messaging

- Document ICD2: describes the format of the messages and header elements.
- Document PDD: describes the system messages and the protocols governing exchanges.

3. Analysis of the situation

The information summarised in the point above shows a situation which calls for a systematic analysis of all the documentation accumulated over recent years. The size of the project, its extension over time and the changes it has undergone since the start are factors which have led to failed efforts to achieve the desired result. A great number of documents have been generated. Some of them are already obsolete, while others are completely or partly in force but often forgotten.

As regards the semantic problems, the solution generally proposed for systematic treatment is the use of controlled vocabulary, i.e. in order of increasing complexity: glossaries, thesauruses, taxonomies and ontologies. Already in 2006, note 353/06 from the Secretariat on “Working methods to include an electronic tool for assisting the ad hoc groups in their work” mentioned the possibility of using ontologies. Although its use would be a very important contribution towards guaranteeing the quality of electronic exchanges, the definition and maintenance of an ontology involves a very high level of complexity which is often off-putting.

This is perhaps why the EIS aims merely to reach an agreement on the definition of metadata. For a project with the EESSI characteristics, it would be necessary to add that these metadata must be easily accessible for all participants in the project, i.e. for all the institutions of the Member States participating in the exchanges.

It is essential for the same data item to have the same definition in all the SEDs, flows and sectors, irrespective of the use which will be made of it or of the group which worked on its description. Consequently, the corresponding metadata should be clearly established, published and kept in accessible form both for use by applications and for consultation by officials who will be handling the documents for administration purposes.

By way of example, we could cite the *Esquema Nacional de Interoperabilidad* (National Interoperability Scheme) developed in Spain, based on Community standards in the field of interoperability, in particular in the ISA programme. From the published standards, we could cite the “*Norma Técnica de Interoperabilidad del Documento electrónico*” (Technical standard for interoperability of the electronic document). This standard sets out the minimum technical conditions

required for the standardised exchange of electronic documents. Among other things, it indicates the obligatory minimum metadata and their format.

The usual solution is to develop a tool to allow the creation and maintenance of metadata in an office computer environment for the subsequent generation of corresponding XML schemas. In short, the solution would be to supplement the description given by the UML model with a data-definition layer accessible to end users, focussing in particular on the semantic problems which could occur in exchanges. This would mean that the information on metadata would be accessible to end users and available for processing by national applications.

4. Conclusions

The note from Task Force 63/07 concludes that the generic definitions of Build 6 could be used for developing the SEDs. The terms defined in this project, over 400 of them, could be the starting point for the creation of an EESSI glossary of terms.

Furthermore, note 579/11 picks up on Sweden's suggestion that this PINS dictionary should be part of a more complex system. The matter is still open for discussion. It could well be the start of a dictionary of data which would further develop the glossary of terms.

This base could be completed by the essential point: the definition of metadata by means of a tool developed for that purpose. As proposed in the previous point, it could be the compromise solution for solving most of the semantic problems which could arise.

Lastly, thought could be given to a second phase of EESSI and the use of metadata defined for exchanges extending to outside the European Union, returning to the proposal made by the Netherlands in note CASSTM 234/08.

AC 522/13REV

Note from the Secretariat. EESSI: Mandate for the Ad-hoc group on data modelling.

EUROPEAN COMMISSION

A.C.522/13REV
EESSI – Mandate for the Ad-hoc group on data modelling
SECRETARIAT – 04.12.2013

Orig.: EN

**ADMINISTRATIVE COMMISSION
FOR THE COORDINATION OF SOCIAL SECURITY SYSTEMS**

Subject: EESSI – Mandate for the Ad-hoc group on data modelling

Note from the Secretariat of 04 December 2013

At the 58th Technical Commission meeting it was agreed that an ad-hoc group on data modelling would be required in order to discuss possible improvements in the current electronic model, especially from an information architecture point of view. Further to this meeting, several Member States have shown interest in participating in this working group while others expressed potential interest, subject to further clarification and detailing of the mandate of this working group.

The possible enhancements in terms of data modelling were further discussed in the EESSI Workshop in Vilnius on 13-14 November 2013, where the participant supported the need for further action in this field and the set up of a working group (AC note 548/13).

The draft mandate for this data modelling and has been presented to the Technical Commission to comment on it (AC note 522/13 from 22 November 2013). The Technical Commission expressed its support for the draft mandate, subject to the comments made at its 59th meeting (AC note 509/13), which have been accommodated in the below revised mandate presented in the Annex of this AC note 522/13_REV.

The Administrative Commission is asked to formally approve the mandate of this group, in accordance with article 6 paragraph 5 of the Rules of the Administrative Commission for the Coordination of Social Security Systems attached to the European Commission (OJ C 213/06.08.2010, p.20).

Annex

Mandate of the Data Modelling Ad-hoc group

1. Legal basis

The Data Modelling Ad hoc group is established in accordance with article 6 paragraph 5 of the Rules of the Administrative Commission for the Coordination of Social Security Systems attached to the European Commission (OJ C 213/06.08.2010, p.20).

2. Context

The Technical Commission agreed to the need of a data modelling ad-hoc group that would assist with the enhancements in the electronic data model of EESSI, especially from a information architecture point of view.

It has been acknowledged that this work should take as input for its work the SEDs, as defined by the SEDs Ad-hoc groups and approved by the Administrative Commission, in order to ensure that the electronic implementation of the business needs in terms of data items that are exchanged is converted in an efficient electronic data model, based on re-usable components.

The working group, while also taking into account the already existing work performed in implementation of the electronic business model, as well as the planned work within the framework of the S.A.F.E. consortium.

The working group will suggest enhancements, improvements or rebuild of the model, to the extent that these are required by existing shortcomings identified in the current model and are in line with best practices in the field and the relevant project constraints.

3. Objectives of the Ad-hoc group

- ✓ Assist with the enhancement and steering of the EESSI information architecture.
- ✓ Contribute to the development of data modelling approach that would include best practices, principles and standards.

4. Tasks

- Make a detailed inventory of the problems with the electronic model (including the information architecture), especially from a Member State (MS) perspective.
- Make concrete proposals for the principles for the data modelling. This should be based on standards and good practices of implementing data models and will ultimately lead to the

elaboration of the EESSI data-exchange European Standard. The following non-restrictive list of items could be considered:

- Quality of the conceptual data modelling
 - Quality of the schema documentation
 - Data types and improvement of schemas' structures (data elements, grouping, common data library, cross-sectorial re-usable elements)
 - Naming policy
 - Scope of the data model (common and sectorial part of the model)
 - Analyzed and documented relationships between items
 - Semantic correctness
 - Security aspects
-
- Identify where concrete interaction with the business audience (in particular the SEDs Ad-hoc groups) is required for driving and validating the work on data modelling.
 - Reflect on cross-sectorial synergies for ensuring re-usability of the data structures and modelling practices. These would be translated into concrete proposals to be raised to the business community. These would be submitted for first consideration to the SEDs Ad-hoc groups in order to further analyse these proposals before any further actions are taken.

5. Deliverables

- The Ad-hoc group shall prepare an interim report for the attention of the Technical Commission by the end of June 2014.
- The Ad-hoc group shall prepare a final report for the attention of the Technical Commission outlining the results of its activities by the end of December 2014.

5. Organisation and working methods

1. The Rapporteur and the Deputy Rapporteur of the Ad-hoc group shall be appointed by the Administrative Commission.
2. The Ad-hoc group shall consist of volunteers nominated by the Member States' delegations in the Technical Commission.
3. The volunteered persons are expected to be experts in data modelling items with a good understanding of the EESSI related business needs in terms of data exchange.
4. The members of the Ad-hoc Group shall not act as members of their delegation, but in the Union interest and on the basis of their personal expertise on this matter.
5. The Ad-hoc Group shall also take into consideration the input of the delegations which are not represented by a member.

6. The Ad-hoc group is responsible for documenting the outcomes of its meetings and the work of the group, prepare relevant papers or reports. The Ad-hoc group shall share such documentation with the DG EMPL EESSI project team and the EESSI Executive Board as and when necessary.
7. The work of the Ad-hoc Group shall be coordinated by the Rapporteur and supported, to the extent possible, by the Secretariat.
8. The Ad-hoc group will work closely with the DG EMPL EESSI project team and will provide input in the work on the data modelling for DG EMPL, who is responsible for this work.
9. The Rapporteur, supported by the Secretariat, shall ensure the necessary cooperation with the Executive Board and with other relevant Ad-hoc Groups of the Administrative Commission, in particular with the SEDs Ad-hoc groups. To this end, the ad-hoc group will define where it shall be needed to contact the SEDs Ad-hoc groups for clarification or input on business related issues.
10. The current mandate brings no alteration to the mandate of the SEDs Ad-hoc groups who remain responsible for defining the business needs in terms of data items that are to be exchanged by the Member States in order to solve cross-border cases as set forth by the Social Security Coordination Regulations.
11. The Rapporteur shall report on the outcomes of its work to the Technical Commission.
12. To allow it to carry out its work, the Ad-hoc Group shall rely as much as possible on the use of flexible working methods such as informal meetings, telephone or video conferences, exchange of e-mails, virtual meetings, newsgroups, etc.
13. The Secretariat of the Ad-hoc group will provide logistic support for the organisation of meetings (i.e. room reservation).
14. The European Commission will do its best efforts to ensure, in accordance with its own applicable rules, the reimbursement of the costs incurred by the Ad-hoc group members when travelling from their home country to the place where the meeting of the group is held. However, this does not represent a commitment to ensure such reimbursements.
15. In light of the deliverables presented under part 5 of this mandate, continuation or closure of the Ad-hoc group will be considered once the current tasks outlined in part 4 have been successfully completed.

AC 738/15

**Note from the Data Modelling Expert Group of 18 November 2015.
EESSI: Data Modelling Expert Group recommendations on the basic
principles to be followed in the data modelling work.**

AC 738/15

**EESSI – Data Modelling Expert Group
recommendations on the basic principles
to be followed in the data modelling work**

SECRETARIAT – 18.11.2015

Orig. EN

**ADMINISTRATIVE COMMISSION
FOR THE COORDINATION OF SOCIAL SECURITY SYSTEMS**

Subject: EESSI – Data Modelling Expert Group recommendations on the basic principles to be followed in the data modelling work

Note from the Data Modelling Expert Group of 18 November 2015

In accordance with its mandate set forth in note AC 522/13 REV, the Data Modelling Expert Group (DMEG) has been working to ensure that it puts forwards actions and recommendations that will ensure that the business needs regarding the data items that will be exchanged in the EESSI system will be converted and implemented in an efficient electronic data model, based on re-usable components.

To this end, the group, working together or consulting with various stakeholders, in particular DG EMPL, the business experts in the Ad-Hoc Groups established for the definition of data to be exchanged electronically (SEDs AHGs) and the business experts involved in the work of the SAFE consortium, has carried out an extensive analysis of current situation. Based on this, the group has undertaken a thorough reflection on changes that will have to be implemented to achieve the objective above. Some of these changes only require IT related activity and intervention whereas some will require some adaptations and changes in the data items that will have to be exchanged and therefore require business support.

As a result of this work, the DMEG puts forward a number of recommendations, in the forms of basic principles that will have to be followed in any further implementation of the electronic data model in EESSI, as per Annex I.

The DMEG asks the Technical and Administrative Commission to support these basic principles for the target data model of the EESSI system as outlined in Annex I of this note.

These principles will be fundamental for any further steps of implementation and DMEG will continue working with the SEDs AHGs and DG EMPL to ensure that these principles are applied and any required changes discussed and agreed with the relevant parties.

Annex 1

BASIC PRINCIPLES for data modelling work

- Recommendations from the Data Modelling Expert Group-

Introduction

Given the different linguistic, cultural, legal, technical and administrative environments in the social security domain, there are significant challenges to ensuring that the precise meaning and formats of exchanged information in EESSI is understood and preserved.

Analysing the content model of the paper SEDS currently in use shows that similar or identical concepts are used by all sectors of social security. What can be seen by this aforementioned analysis is, that there are differences in the naming conventions, in the content structure and the principle content used for those concepts in an applied business context. The consequence of this diversity is a huge effort to process and evaluate information based on this entities especially if the business architecture relies on electronic data exchange. Additionally this diversity introduces risk of misinterpretation and erroneous handling of information.

The core challenge for a transition into the electronic data exchange in EESSI is to reduce the diversity of (redundant) business entities in order achieve interoperability between (1) technical components for the routing and processing of messages, (2) national applications for the management of business cases as well as (3) clerks using these applications for understanding and preserving this information.

The result of this reduction process is a reliable data model which supports the following objectives:

- A reliable data model enables faster EESSI integration for Member States
- A reliable data model decreases the overall software and hardware costs of Member State EESSI implementations
- A reliable data model reduces future EESSI service upkeep/development and maintenance costs on both EU and Member State level
- A reliable data model reduces efforts to change or extend the information architecture of EESSI
- A reliable data model increases the understanding of information exchange in Social Security sector across Europe

In the following the data modeling expert group recommends necessary requirements which should be applied to reach the aforementioned goal of a reliable data model.

The requirements for the data model

A data model based on standardization, alignment and unification of building blocks is the key to reach interoperability of the EESSI system at different levels. Standardization and unification reduces the number of necessary building blocks for the message exchange system and reduces structural differences of information items sharing the same meaning across several business cases.

Based on an extensive analysis on the current content model of the necessary business messages in EESSI the DMEG group defines the following requirements for the target EESSI data model:

1. It must allow to conduct business according to rules specified in EC regulations 883/2004 and 987/2009;
2. It must be based on the current content model developed by the AD hoc groups for processing paper SEDs;
3. It must reflect the fact that the business architecture is based on electronic data exchange;
4. It must support human-readable and machine process able representations of business information;
5. It must take into account the fact that data precision/quality/availability is different in different sectors;
6. It should avoid preventable development and maintenance costs, provided other conditions are met.

As a consequence of these requirements and taking into account the structure of the EESSI Business architecture, basic principles have to be applied in order to develop the EESSI target data model.

The basic principles for the data model

The Data Modeling Expert Group (DMEG) recommends, based on the analysis of the business architecture the current content model of the SEDS and the aforementioned requirements the following business related objectives for the EESSI information model:

- 1. The EESSI data model must define a set of common building blocks which are used without any further changes by all business use cases in the same way in all SEDs and which share the same meaning and format in all sectors (EESSI standards).**

These EESSI standards guarantees a uniform usage of this building blocks across all sectors and their associated business messages. This is of special importance if critical information is processed in different sectors of the social security system. The generation of EESSI standards will be performed by (1) defining the union of information items developed by all sectors related to this building block and (2) by reducing the possible information items due to the electronic nature of the message exchange system. The impact on the current content model documented in the paper SEDS will be the following:

- (1) The naming of specific information items will be unified across all sectors.
- (2) Some sectors will face (minor) additional information items in the content model of specific building blocks. These additional information will be always optional.

(3) All sectors will be aligned due to the fact that the EESSI system is based on electronic data exchange and supported by applications. This means that sometimes not the information itself will be exchanged but a reference to that information provided by specific electronic system.

2. The EESSI data model must define a comprehensive set of sector specific building blocks that are derived from single common building blocks but that share the same meaning and format and are used consistently across a single sector (EESSI alignment and unification).

This EESSI alignment and unification procedure guarantees a uniform meaning of concepts across all sectors in the social security domain. This is of special importance if information related to the same concept is processed across different sectors of the social security system. The generation of EESSI standards will be performed by (1) defining the union of information items developed by all sectors related to this building block, (2) by reducing the possible information items due to the electronic nature of the message exchange system and (3) by applying sector specific restrictions on the specific building block. The impact on the current content model documented in the paper SEDS will be the following:

- (1) The naming convention of EESSI standards will be used consistently in all sectors.
- (2) The naming conventions introduced by the paper SEDS for sector specific information items will be used consistently in all sectors (where concepts do not overlap).
- (3) All SEDS in a sector will be aligned due to the fact that the EESSI system is based on electronic data exchange and supported by applications. This means that sometimes not the information itself will be exchanged but a reference to that information provided by specific electronic system.
- (4) The diverse implementation of a specific concept in different SEDS of a sector will be avoided.

3. The generation of electronic business messages must follow a standard methodology

The EESSI data modeling exercise based on a standard methodology is a prerequisite for the development of a reliable data model in EESSI. This is of special importance for the transition of the current paper SEDS into the electronic data exchange system, since the business requirements expressed through the current paper SEDS have to be preserved but a cross sectorial alignment is needed. The impact on the processes related to data analysis and data modeling will be the following:

- (1) The electronic representation of information is the key focus.
- (2) Localized information (e.g. translations) related to data will be stored and maintained.
- (3) The implementation of the data model will be supported by a consistent tool chain.
- (4) The change management of the data model is based on standard procedures.

The DMEG asks the Technical and Administrative commission to support these basic principles for the target data model of the EESSI system as outlined in this paper. These principles will be fundamental for any further steps of implementation.

EC (1997)

Information note on the TESS programme

INFORMATION NOTE ON THE TESS PROGRAMME

(previously known as TESS / SOSENET¹ Programme)

*A programme for the development of telematic services
for the co-ordination of European Social Security Schemes*

Master Plan Reference : IM.03.02

1. General objectives and Underlying Principle

The free movement of people constitutes one of the four pillars of the European Community. The Treaty of Rome guarantees that nationals of Member States may work anywhere in the Community. This implies the abolition of any discrimination based on nationality, particularly concerning entry into the labour market.

The fundamental differences between the social security systems of the Member States are liable to produce undesirable effects for the worker when he/she crosses intra-community frontiers. The rules for co-ordination, as laid down in (EEC) Regulations No 1408/71 and No 574/72, are meant to ensure that existing disparities do not become an obstacle to the free movement of people.

For more than 30 years, Community legislation has been providing a high standard of social security protection for workers and their families moving within the European Union and, since 1 January 1994, within the European Economic Area.

2. Statement of the Problem

In order to apply this legislation, a large number of social security institutions need to exchange information about insured persons. Up to now, this is still conducted mainly by clerical means. The clerk in one institution completes a paper form which is then sent to another institution and handled there manually. These procedures are staff intensive and therefore expensive and they are subject to delays and other inefficiencies (e.g. no automated control or assistance).

This situation calls for modernisation. Today, the new Information Technologies and, in particular, the use of Telematic services allow the exchange of information by electronic means.

However, the technology will only result in speeding up the exchange of forms. The result of all this effort will be very poor (e.g. three days will be saved from a three-year procedure). In order to reduce three years to three months, new technologies have to be combined with an in depth analysis of the problems encountered in the current exchange of information and the implementation of complementary administrative rules and procedures.

3. Objectives of the TESS Programme

¹ TESS stands for Telematics in Social Security and SOSENET stands for Social Security Network.

In order to speed up the processing of the information, the modernisation has to apply a *strategy : to organise the availability of an accurate information set wherever and whenever it is required*. This is the means of reaching the objective set by the Member States to the TESS Programme : *to accelerate and simplify the administrative procedures in order to improve the recognition of the rights acquired or in course of being acquired, the award and payment of benefits derived from the application of the (EEC) Regulation No 1408/71*.

Moreover, with this, the TESS Programme will foster superior objectives: *improving mobility of people* and, at the same time, *reinforcing the Single Market by making the disappearance of internal frontiers more tangible*. Furthermore, in line with the White Paper on Growth, Competitiveness and Employment, this will contribute towards laying *the foundation for the information society* and the development of a *Community employment market*

4. Strategy for a Staged Implementation

The overall objective set for the TESS implementation strategy is to create and build-up a reference system for telematic services. This strategy will act as a catalyst for their extension to all institutions concerned.

The first stage of the programme was exploratory. Two projects were undertaken from 1991 to 1994. As a result, it was possible to identify the major problems related to the exchange of data with a view to verifying the feasibility of modernisation using electronic mail technology, and principally X400 services, as well as through EDI (Electronic Data Interchange) and the EDIFACT message format (a world-wide standardized format).

The studies required to carry out this first stage were financed by the European Commission and a consortium of private companies in the context of a technological research and development programme. The validation of these studies by seven institutions in seven Member States was financed by the institutions themselves.

The second stage of the programme consisted of the establishment on a small scale of various telematic services. These concentrated on the sector of firstly old-age pensions of migrant workers, and then sickness benefits in kind.

In the old-age pension sector, the main problem to be solved is to facilitate the reconstitution of a migrant worker's career in order to determine and calculate his/her pension based on the insurance period in each Member State concerned. For the sector embracing sickness benefit in kind, the main problem to be solved is to facilitate the reimbursement of the benefit in kind advanced by the Insurance Body in the State where the person has resided or stayed.

Other objectives at this stage of the programme were the definition of potential co-operation between telematic services and social security card and to prepare the transfer of responsibility for operating the telematic services to the institutions themselves.

Today, the achievements so far are:

- the definition of an information exchange model which will be applicable in all of the social security sectors of all Member States;

- a prototype among 7 institutions of the old-age pensions sector of 7 Member States which validates a first part of the model and is used for operational exchange of data in order to prepare the calculation of an old-age pension of migrant workers;
- the creation of a social security glossary of terms describing the meaning of all the terms exchanged, their use and their equivalent in all Member States; EDIFACT messages have been developed on this basis.

The following steps in course of being performed or to be performed are:

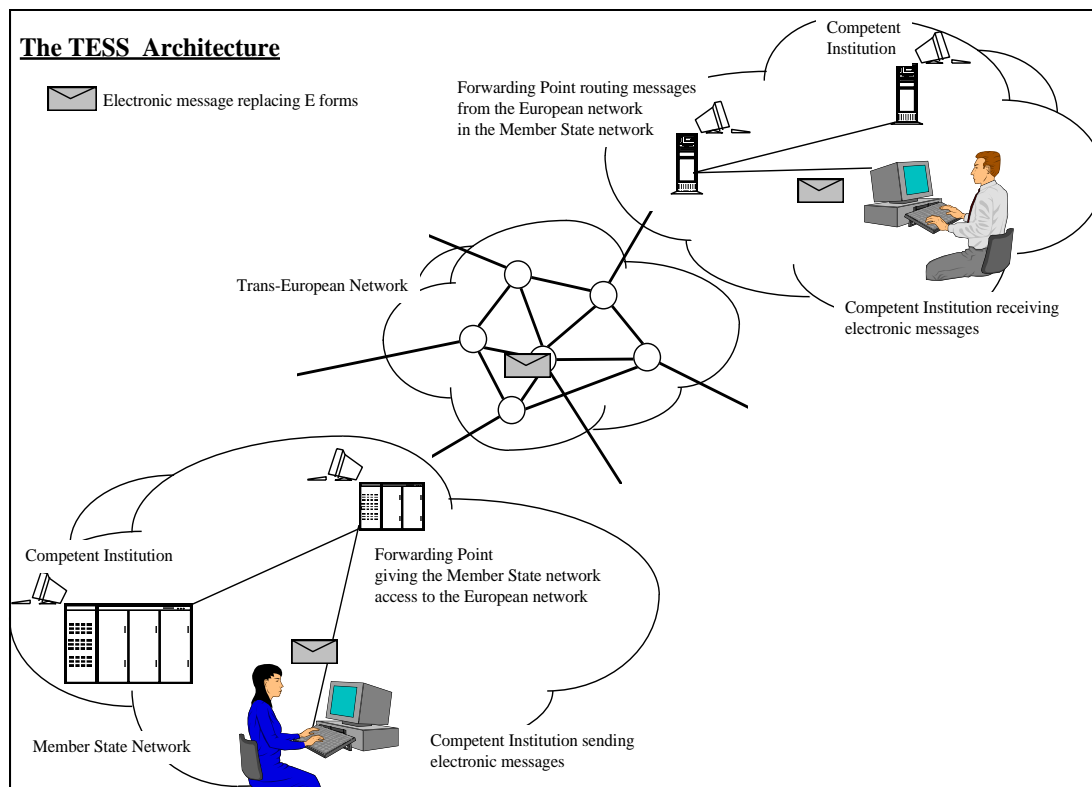
- the development of a reference system among a limited number of institutions (approximately 30) from the sectors of old-age pension and sickness benefits in kind;
- investigations on the possible solutions for modernising the E 111 and other similar forms;
- definition of the necessary mechanisms for management of the developed telematic services and the financing of their expansion.

The studies will be completely financed by the European Commission, although their implementation, which means the purchase or hire of hardware and software, together with the development of interfaces for the applications, will be the responsibility of the institutions themselves.

The third stage, which is planned for 1998, will consist of the extension of the system to all institutions concerned. This will involve a progressive powering up of the network.

5. Overall Architecture Underlying the TESS Services

The operational principle underlying the TESS services is simple, since it is based on the concept of the postal system used for letters and other paper documents? It is illustrated by the plan reproduced below.



By way of illustration, let us take the case of an official in a social security institution who has to calculate the amount of the old-age pension to which a worker is entitled. In the case a worker has spent a part of his/her working life in another Member State, the official may have particular need of information.

The official therefore creates an electronic E form from the database using various programmes. After a number of controls particularly to ensure data security and privacy protection, the form is sent through the national network to the point of transfer, the "Forwarding Point".

At the Forwarding Point, the form is addressed to the correspondent in the Member State to which the form is being sent. It is then transmitted to the competent institution, which receives it through its information system.

Depending on the case, the request for information will either be dealt with by an official and the response sent back the same way, or it will be automatically processed by a computer application and the response transferred back, still subject to controls for security and privacy of the data.

In order to be implemented, this system has to resolve enormous organisational and technical problems. Social security systems are organized in a very different way from one Member State to another. Moreover, computer systems in the institutions are also extremely varied and are still rarely compatible.

This is accordingly a long process which requires a lot of energy and a high level of involvement from all of the actors. Co-operation is the guiding principle in this field of activity.

Pekka, R. (2013)

Data Modelling in the EESSI-project

EESSI Working Party in Vilnius (Lithuania), 14-11-2013.

Memo: Data Modelling in the EESSI-project

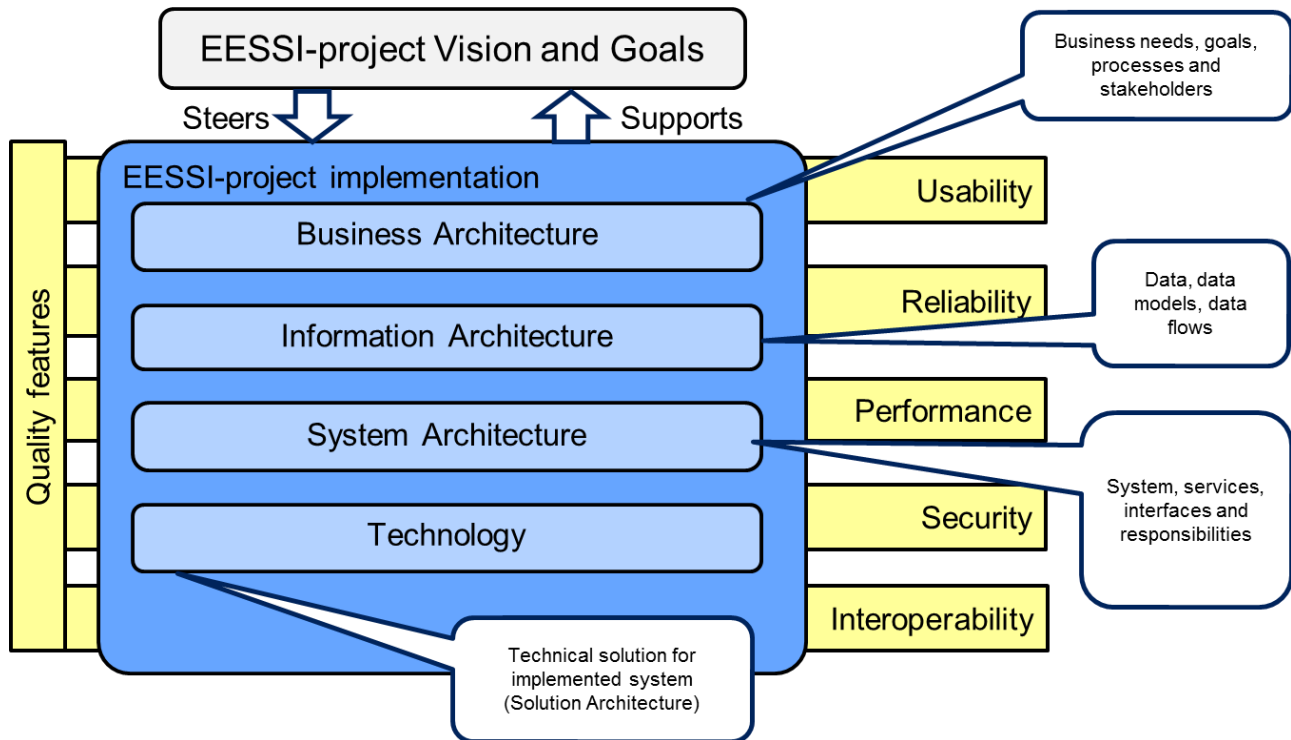
Introduction

The EESSI Data model is very important both to national EESSI application development and when developing the EESSI RI. Therefore, well executed data modelling is one of the key elements for success in the EESSI-project. This will be highlighted especially in terms of changes in data model. If the model has not been designed to take changes into account, it will be difficult and expensive for the Member States and the Commission to adjust the model in the future.

The purpose of this document is to create an overview of the concepts of information architecture and data modelling in the central EESSI-project but also for the Member States. This document briefly describes how data modelling fits in the ICT-project and the document is based on the notes presented earlier by several member states and discussions with representatives from the Nordic countries.

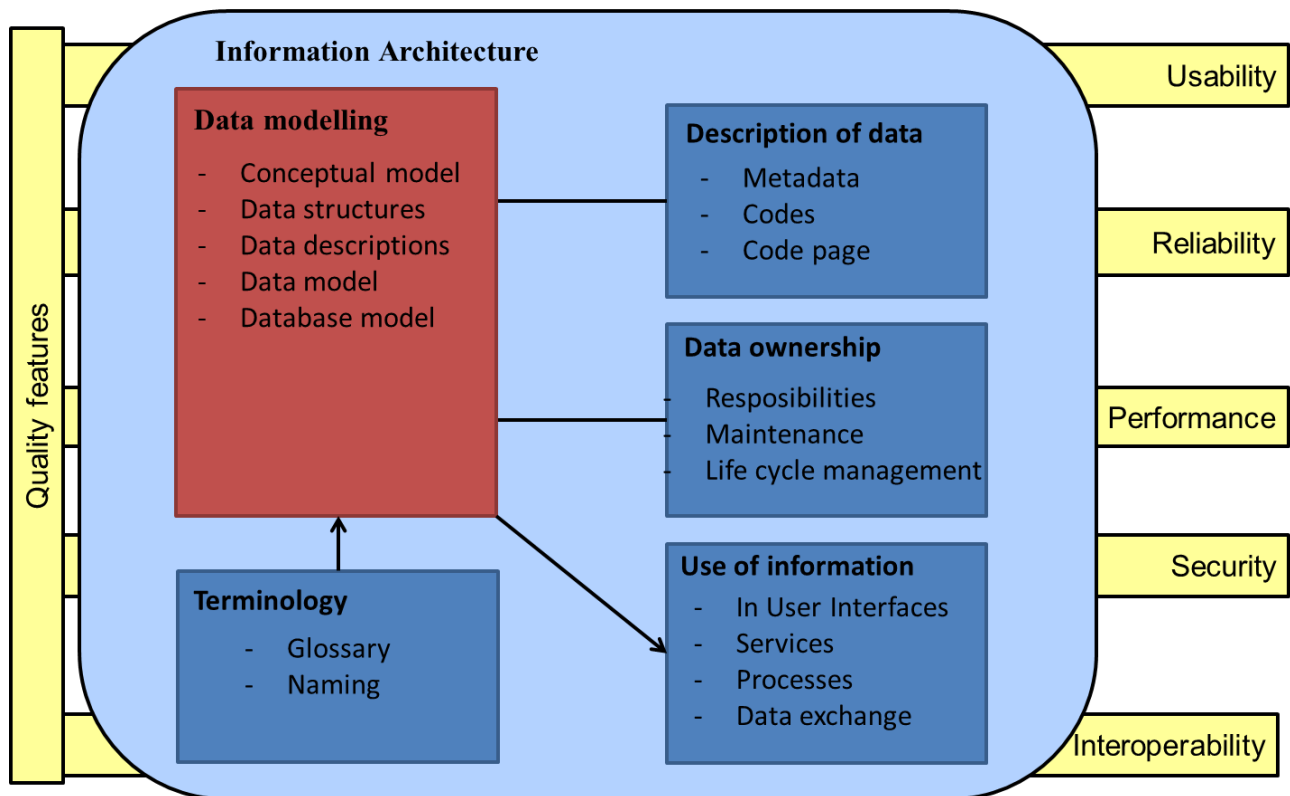
Architecture viewpoints in general

Different architecture viewpoints which should be taken into account in projects are described below. Data modelling is part of the information architecture.



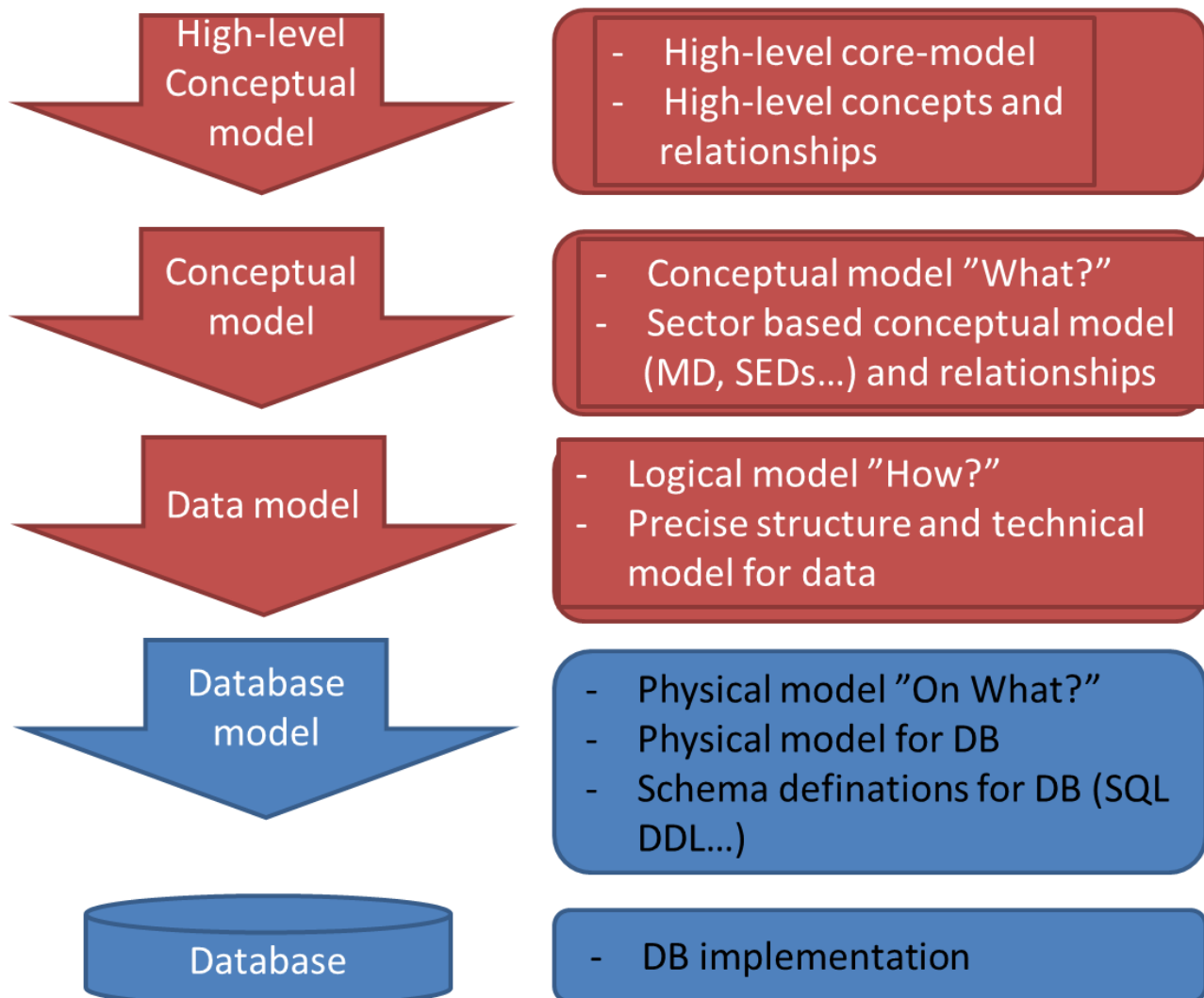
Information Architecture

Information architecture is a view that describes the use of information, the structure of information, the classification of information and the exchange of information between systems. The same cross-cutting quality aspects that apply to ICT-project also apply to information architecture.



Data modelling

The most important part of the information architecture is data modelling including data analysis. The purpose of data modelling is to describe the information in a structured and formal way (conceptual model). The work is usually done top down: from conceptual model to data model and eventually to database model. The diagram below defines the layers inside the information architecture.



Technical items such as a database schema are something that is derived from higher abstraction levels and as such, they also reflect all the changes on those levels.

In EESSI we can have one core model which describes the high-level conceptual model (high-level concepts). Sub-systems (e.g. Master Directory) or divisions (SEDs=business model) can use the common information from the core model.

From the logical to the physical model we use specific syntaxes, e.g. XML Schema for SEDs or SQL DDL for databases.

While the primary goal is to model the SEDs in XML Schemas, it is important that the same logical model can be used by national applications and their databases. DG EMPL is encouraged to provide reference definitions for widely used DBMS, e.g. IBM DB/2 and Oracle 11g.

One big decision will be to decide what kind of data model EESSI will offer. There are different implementation strategies for EESSI data modelling.

- 1) One data model for EESSI – This will be hard to maintain because all changes will affect everything (in national and international domain)
- 2) Sectorial domains – This will limit the effect of the changes to the sectorial domains (sub-systems/business sectors)
- 3) Document based domains – This is the current situation, the model is hard to maintain and national implementation work is hard (e.g. building national archiving system – you have to search the meaning of the information for each element individually).
- 4) Mix of the three above

All these implementation strategies are possible to implement but the question is which model is easiest to maintain and which strategy offers Member States the best tool for implementation in national applications. **The model should contain one core model where the information which will not change (library for core elements like first name, last name, institution name...) is described and sectorial models where the information which are sector based information (e.g. in this sector we use this information concerning person) is described. There is also a need for document based domains (e.g. elements inside country specific SEDs).**

Data modelling in EESSI-project

This chapter describes where we are now and what should be done in the future concerning data modelling, collecting the key points from the following notes:

- 413/13 - Improved schemas needed
- 370/12 EESSI - Necessary improvements for implementation on national level
- 446/12 EESSI - Necessary improvements for its establishment semantic aspects
- 645/11 - The improvement of SEDs and the acceptance procedure
- ANNEX TO NOTE AC 645/11 – Memo on schemas

Paper SEDs and flow definitions are based on the legislation and processes concerning EC regulation 883/2004, which is a good base to begin with for work with the EESSI data model. The actual data modelling can be executed in the DG EMPL – project or this can be executed outside the project (Member States or external company). However DG EMPL – project is responsible for data modelling and should finance the actual work.

Before the actual modelling can start there must be an agreement with Member States on what are the principles for modelling. Even if the modelling is executed outside the DG EMPL – project, the project must have an information architect in place that is responsible for the whole EESSI information architecture. One of the main tasks is to coordinate the information architecture work with EESSI.

There are huge problems in the current data model concerning quality of schemas. Re-usability of elements, structure of schemas, lack of naming policy and semantic correctness are the worst things in current model. Components should have been coordinated, the structure of the components should have been analysed and relationships between components should have been verified for semantic correctness. One of the causes for these problems is the lack of the naming policy. The missing naming policy is causing conflicts between the terms and the definitions. A term should always represent a single concept, otherwise the associations between concepts and definitions will become very hard to maintain. If the same term is to be used to define multiple concepts, it should be prefixed by an identifier (technically a namespace) that bounds it to the context. It is important to remember that the tool for data model is NOT the key thing – mostly the modelling is hard work and communication between stakeholders (expert groups). The data analysis phase will be the base for working technical data model.

Also, procedures for acceptance of schemas need to be in place within the EESSI-project.

Data modelling task list in a nutshell for EESSI-project (proposal)

Task	Responsibility
Allocate Information architecture skills for the project	DG EMPL
Approval of version 3 of pSEDs	MSs/AC
<p>Documented list of concepts used in pSEDs (Data analysis for pSEDs)</p> <ul style="list-style-type: none"> - Select tool for data analysis (e.g. Excel) - Use of same structure with same element - Descriptions for data elements - Grouping of the elements 	EG EMPL and Expert Groups
Decide who will make the data modelling	DG EMPL
<p>Common agreement of the principles of data modelling: Data modelling workshop</p> <ul style="list-style-type: none"> - Common agreement on notation used in EESSI-project - Common agreement on scope of the data model (one data model for EESSI, sectorial data models for EESSI....) - Common agreement on code-pages (e.g. will DG EMPL deliver common rules for transliteration or will all Member States use non-Cyrillic letters or is it responsibility of sender to make the transliteration) - Common agreement on life cycle management for EESSI business data (national legislation will steer this but there must be rules e.g. regarding minimum retention period for SEDs.) 	DG EMPL and MSs
Acceptance of the principles	EB / MSs (Expert Group)
<p>Data analysis</p> <ul style="list-style-type: none"> - Select tool for data analysis (e.g. Sparx EA) - Data descriptions (data elements, grouping, common data library) to the tool - Security aspects (e.g. data classification: new EU-regulation concerning security) 	DG EMPL/MSs expert resources/Externalresources

- Conceptual model (relationships between entities)
- Release of the data analysis document

Acceptance of the data analysis document

EB/TC

Data modelling

- Select tool for data modelling (usually it's possible to use same tool as data analysis)
- Create Structures, grouping, common data library (data types, lengths, mandatories)
- Utilization of described data elements and structures in SEDs

Establish procedure for acceptance of the model

DG EMPL

Release of the data model

DG EMPL

Acceptance of data model and acceptance process

TC

Useful information concerning data modelling

http://pubs.opengroup.org/architecture/togaf9-doc/arch/chap10.html#tag_10

http://en.wikipedia.org/wiki/Data_modeling

Nota TF 063/07

Task Force on electronic data Exchange: Build 6 Outcomes and the EESSI Project.

TF 063/07
Build 6 outcomes & the EESSI project
SECRETARIAT – 24/09/2007

ORIG: EN

**TASK FORCE
ON ELECTRONIC DATA EXCHANGE**

Subject: Build 6 Outcomes and the EESSI Project

Note from the Secretariat of 24 September 2007

The purpose of this report is to provide a consolidated summary of the outcomes of Build 6 and examine how far the Build 6 outcomes of 1997 can help with the development & deployment of EESSI.

Summary

Build 6 was a pilot project in the field of Pension and Health care and developed the electronic messages, data model and glossary of terms by applying an intersectorial approach.

For the EESSI project, it must be said that the reusability of the final outcomes of Build 6 are low, but to some extent several of the interim outcomes of Build 6 could be applied. Build 6 is based on Regulation 1408/71 and was influenced by EDIFACT – and as both these building blocks are now outdated - it does not make much sense to utilise Build 6 per se without further considered reflection.

1. It may be worth considering the horizontal intersectorial approach when consolidating the work of the various Ad hoc groups on business flows
2. It is however important to bear in mind that the EDIFACT philosophy was to cram as many exchanges as possible into a single standard (but heavily overloaded) message. This had the benefits of greatly simplifying the standardisation process at the cost of the implementation effort. Now, Web Services tends to adopt a more granular approach and enjoys the benefits of layered modelling introduced in the 90's.
Therefore, though in the main most definitions of terms etc, the identification of flows may not have changed dramatically under the new Regulation 883/2004 and its implementation regulation - using the final results (implementations) of Build 6 is not really appropriate.

3. It is probably more important to focus on the intermediate results than on the final (EDIFACT-based) implementation. It may be useful to undertake an exercise to bring the intermediate result into a modern modelling structure.

DETAIL

1. Origin of Build 6, its objectives and approach

The Build 6 Project was part of a “Master Plan for the second stage of the TESS Programme (1994-1998)” and agreed upon in 1995 by the Technical Commission and the Administrative Commission. Its objective was the development of pilot projects in the field of Pension and Health care with the view to build a “Reference System” for large scale deployment. Other projects defined in the Master Plan and still active are Build 3+ and Build 4.

Running from 1995 to 1997, Build 6 objectives were the development of the electronic messages, data model and glossary of terms necessary for the execution of the Master Plan.

The project has applied an intersectorial approach in order to identify commonalities between all paper forms (E forms as well as the bilateral ones) and rationalise the manner in which data and flows should be structured and exchanged when moving from paper to electronic. As it was the most comprehensive Electronic Data Interchange (EDI) project of that time in Europe, Build 6 had to develop its own methodology, utilising the best practices identified and recognized at that period in time.

2. Outcome of Build 6 re-usable by EESSI

While the project has delivered what was expected, it has to be concluded that not all outcome of re-usable by EESSI. Indeed while the level of re-usability of some final outcomes (see below) can be low, one should consider that the re-usability level of several interim outcomes which can be higher by far.

It has however to be borne in mind that the Build 6 outcomes relate to the application of EC Regulation 1408/71.

Remark: the simplified document flow diagram of Annex 8 identifies the key (interim and final) deliverables of Build 6 and their interrelationship.

2.1 Electronic messages

As it can be understood from Annex 4 below, the definition of the messages has been very much influenced by the EDIFACT context and history, this part of the Build 6 outcome is probably of little use for EESSI.

On the other hand, the work done by Build 6 on the 4 “Business Functions” (see Annex 2) is likely to be re-used for the creation e.g. of 4 SED:

- | |
|---|
| <ol style="list-style-type: none">1. Collecting information2. Authorising to supply/attesting the right to a benefit3. Deciding a reimbursement |
|---|

4. Deciding to award a benefit.

In that approach, as much as possible intersectorial building blocks of data could be developed and the amount and size of sectorial-dependant building blocs of data would be reduced to the minimum. By doing so, maintenance efforts would be minimized and consistency in the data structure maximized. Each of the 4 messages would then include an intersectorial part and a sector-dependent one.

For defining the content of these 4 Business Functions, Build 6 has worked out 43 “Exchange of Data” (see Annex 1) that are similar to the “Business Protocol” defined in the Guidance Note CA.SS.TM. 222/07.

Also the fact that these information flows have also been described from a data viewpoint (see section 0) is also potentially of great interest.

2.2 Data model

Starting from the current forms (as per 1995); Build 6 has identified all information concepts used by the forms, created an information repository (see Annex 5). It is likely that this Information Repository could be re-used for defining the data content of each of the Business Protocols.

Furthermore, the Build 6 data model and its 4 diagrams modelling all data related to e.g. the Person and his/her Institution, (see Annex 6) could also usefully be re-used when creating intersectorial building blocks of data for electronic messages.

2.3 Glossary of Terms

Build 6 has furthermore defined each data from the information repository (as per 1995) with a generic definition, followed by a definition specific to each Member states with a view to enable an institution to best process the data received from abroad. This definitory work has produced the Glossary of Term (see Annex 7).

The generic definitions could also be re-used for the development of the SED and the specific definitions could be reviewed in due time, by the Member States for ensuring mutual understanding between two IT systems from two different Member States.

Technical Annexes (1 to 8)

Annex 1: 43 Exchange of Data (ED)

The Guidance Note 222/07 defines “Business Protocol” as a “special set of rules that the end points (in an electronic communication connection) use when they communicate. It is composed of two things: a) the data and b) the flow of that data (also called the business flow). The data will serve to create the Structured Electronic Documents (SED) while the business flows will stipulate how that data will flow between the various points.”

The Build 6 concept that best matches with this definition of “Business Protocol” is “Exchange of Data” “which represents all the flows logically linked that are needed to fulfil a specific Social Security task” involving several institutions from different countries.

By analysing the current cross-border exchange of information (as per 1995), Build 6 has identified 43 sectorial “Exchange of Data”.

• Exchanges of Data in Health Care and Determination of Applicable Legislation Sector
- EDH01: Posted worker: declaration of posting
- EDH02: Posted worker: extension of posting
- EDH03: Right to option for person employed by diplomatic missions and consular posts, and auxiliary staff of the European Communities
- EDH04: Reference period for sickness and maternity benefits for people moving to another state
- EDH05: Certificate concerning employed person in international transport
- EDH06-A: Entitlement to sickness and maternity insurance benefits in kind for persons residing in a country other than the competent country
- EDH06-B: Registration of members of the employed or self-employed person's family and updating of lists
- EDH06-C: Entitlement to benefits in kind during a stay in a member state
- EDH06-D: Entitlement of unemployed persons and the members of their family to sickness and maternity insurance benefits
- EDH06-E: Entitlement to benefits in kind for pension claimants and members of their family
- EDH06-F: Registration of pensioners and updating of lists
- EDH06-G: Granting of benefits in kind to members of the family of pensioners
- EDH07: Notification of admission in hospital and/or discharge from hospital
- EDH08: Granting of major benefits in kind
- EDH09: Retention of the right to sickness or maternity benefits currently being provided
- EDH010: Suspension or withdrawal of sickness and maternity benefits in kind for persons residing outside the country where they are insured
- EDH011: Health care reimbursement
- EDH012: Reimbursement on lump sum basis of medical expenses incurred abroad during a temporary stay

- EDH013: Death allowances
- EDH014: Accident at work and occupational disease benefits in kind for workers residing outside the country where they are insured
- EDH015: Incapacity benefits in cash (due to maternity, illness, professional illness or labour accident) for persons residing or temporarily staying outside the country where they are insured
- EDH016: Request for information about the composition of the family of incapacitated person for the calculation of the benefit

• Exchanges of Data in the Pension Sector

- EDP01: Joining the voluntary or optional continued insurance scheme
- EDP02, case 1: Claim for a pension (old age pensions, invalidity benefits and survivor's benefits) in case of two institutions involved
- EDP02, case 2: Claim for a pension (old age pensions, invalidity benefits and survivor's benefits) in case of more than two institutions involved
- EDP03: Administrative check on the position of a pensioner

• Exchanges of Data in the Sector Related to Decisions 117 and 118 (Migrant Worker Legislation)

- EDD01: Registering a migrant worker
- EDD02: Changing the identification data of a migrant worker
- EDD03: Early drawing up the insurance history

• Exchanges of Data in the Unemployment Sector

- EDU01: Request for reference period for the awarding of unemployment benefits
- EDU02: Request for information about the composition of the family of the unemployed person for the calculation of the benefit
- EDU03 and EDU05: Unemployment benefit for people going abroad in search of a new job
- EDU04: Request for reimbursement of the benefit paid for people having searched a new job abroad

• Exchanges of Data in the Family Sector

- EDF01: Request for reference period for the entitlement to family allowances
- EDF02: Request for certification concerning the family composition
- EDF03: Request for certification of apprenticeship
- EDF04: Request for certification for a student
- EDF05: Request for certification for young children's allowances
- EDF06: Request for medical report
- EDF07: Request for information concerning the family allowance awarded in country of residence

• Exchanges of Data in the Non-Contributory Benefits Sector

- EDN01: Request for information about the family income for the calculation of the non-contributory benefit
--

- EDN02: Request of reference period for the entitlement to non-contributory benefits

Remark: Since Build 6 started from the paper forms and ended with electronic messages, the project has been working in an integrated manner on both the flows and the data. The definition of “Exchange of Data” has therefore not to be interpreted as excluding the data, but as including them implicitly.

Annex 2: Business Functions, activities & flows

In order to be able to describe these “Exchange of Data” in a structured and sector-neutral way (this is seen to be a pre-condition to identify the building blocks of data that are intersectorial/generic from those that are sector-related/specific), they have been further analysed from a data viewpoint and classified into 4 sector-independent groups called “Business Functions”. They are the following ones (and we think they may be very useful in an EESSI context):

1. Collecting information
2. Authorising to supply/attesting the right to a benefit
3. Deciding a reimbursement
4. Deciding to award a benefit.

Each of these “Business Functions” has then been described by listing a set of generic activities and sub-activities included in a particular business function. Between two activities, a generic information flow has been identified (see the table below).

A full description of each of the 43 “Exchange of Data” is then obtained by applying the generic activities/sub-activities and information flows to the particular case of “Exchange of Data” considered (see example of such a description in Annex 3).

Business Function Classes	from activity	to activity	Classes of Flows
Collect information (CS01)	Requesting administrative information	Collecting administrative information	CS01-F01
	Collecting administrative information	Distributing the received information	CS01-F02
	Distributing the received information	Processing the receiving information	CS01-F03
	Collecting administrative information	Processing the receiving information	CS01-F04
	Processing the receiving information	Processing the receiving acknowledgement	CS01-F05
Authorize to supply/	Requesting authorization/attestation	Providing authorization or attestation	CS02-F01

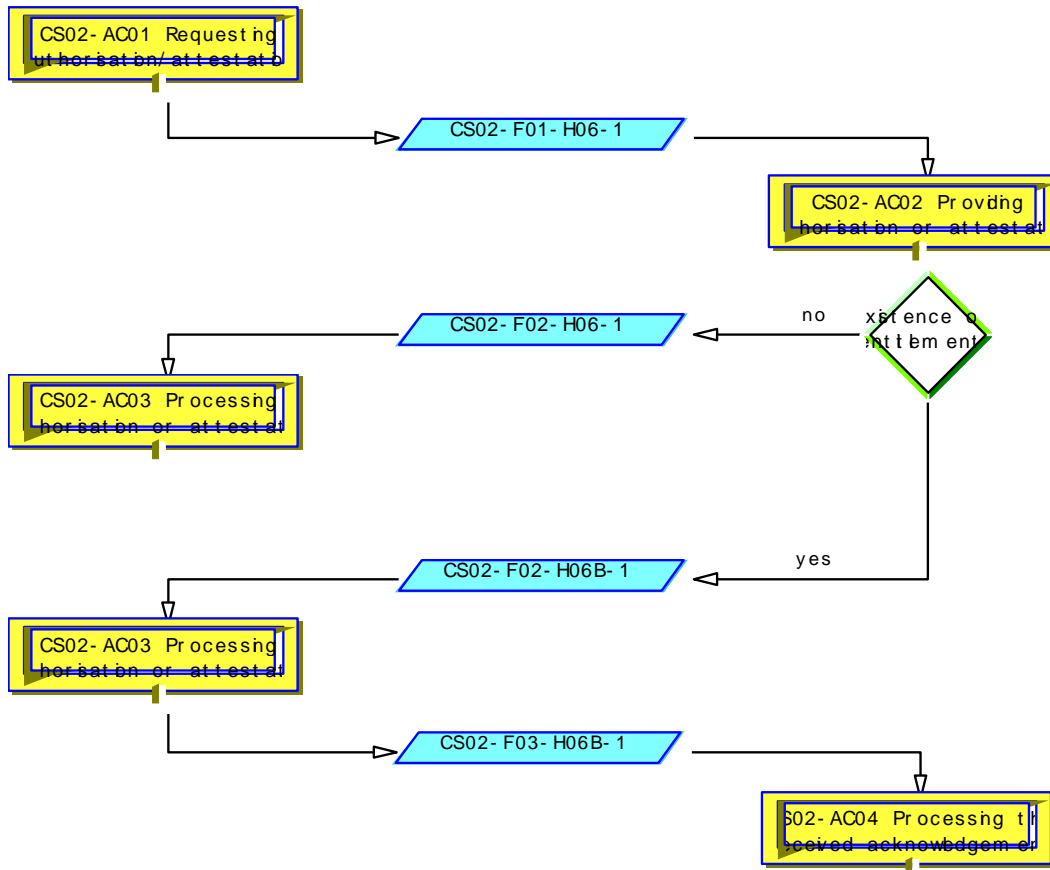
Business Function Classes	from activity	to activity	Classes of Flows
attest the right to a benefit (CS02)	Providing authorization/attestation	Processing authorization/attestation	CS02-F02
	Processing authorization/attestation	Processing the received acknowledgement	CS02-F03
Decide a reimbursement (CS03)	Requesting reimbursement	Deciding reimburse/pay to	CS03-F01
	Deciding reimburse/pay to	Processing the received decisión	CS03-F02
Decide to award a benefit (CS04)	Requesting a decision	Deciding to award a benefit	CS04-F01
	Deciding to award a benefit	Distributing the received decision	CS04-F02
	Distributing the received decision	Processing the received decisión	CS04-F03

Annex 3: Example of an ED description

8.6.6 Exchange of Data EDH06-B: Registration of members of the employed or self-employed person's family and updating of lists

Remark: Build 6 has not considered the situations where the insured person has a document certifying rights. Portable forms were out of scope of the project.

EDH06- B



Description of activities and sub-activities:

CS02-AC01 Requesting authorization/attestation:

Subactivity 1: Taking delivery: n/a

Subactivity 2: Synchronization: n/a

Subactivity 3: Internal function: Verify the entitlement

Subactivity 4: Sending to recipients: Sending one flow CS02-F01-H06-1

Subactivity 5: Waiting acknowledgement:

peculiar service: n/a

CS02-AC02 Providing authorization or attestation:

Subactivity 1: Taking delivery: Receiving one flow CS02-F01-H06-1

peculiar service: n/a

Subactivity 2: Synchronization: n/a

Subactivity 3: Internal function: Attest the entitlement

Subactivity 4: Sending to recipients:

If the entitlement exists, sending one flow CS02-F02-H06B-1
else sending one flow CS02-F02-H06-1

Subactivity 5: Waiting acknowledgement:

peculiar service: n/a

CS02-AC03 Processing authorization or attestation, in case of entitlement denied

Subactivity 1: Taking delivery: Receiving one flow CS02-F02-H06-1

peculiar service: n/a

Subactivity 2: Synchronization: n/a

Subactivity 3: Internal function: Deny the benefit

Subactivity 4: Sending to recipients: n/a

Subactivity 5: Waiting acknowledgement: n/a

CS02-AC03 Processing authorization or attestation, in case of entitlement existing

Subactivity 1: Taking delivery: Receiving one flow CS02-F02-H06B-1

peculiar service: n/a

Subactivity 2: Synchronization: n/a

Subactivity 3: Internal function: Supply the benefit and communicate the list of the persons

Subactivity 4: Sending to recipients: Sending one flow CS02-F03-H06B-1

Subactivity 5: Waiting acknowledgement:

peculiar service: n/a

CS02-AC04 Processing the received acknowledgement:

Subactivity 1: Taking delivery: Receiving one flow CS02-F03-H06B-1

peculiar service: n/a

Subactivity 2: Synchronization: n/a

Subactivity 3: Internal function: Update the inventory

Subactivity 4: Sending to recipients: n/a

Subactivity 5: Waiting acknowledgement

Annex 4: Build 6 Messages

All the sectorial information flows identified through the above-mentioned description process have also been used for creating new intersectorial messages or modifying existing messages.

Build 6 reviewed the existing EDIFACT messages. 4 messages were already in existence at some stage of standardisation and Build 6 has drafted the necessary modification requests. Two brand new messages were developed as a complement.

- The SSREGW message is exchanged between social security organisations to enable interchange and registration of a person's social security numbers. The message is also exchanged by social security organisations to compile details of person social security insurance record.
- The SSIMOD message is sent to communicate to the country of first registration the modification of identity details which took place since the last transmission related to the worker.
- The SSRECH message is sent by a social security body to another social security body giving details of a person's Social Security insurance history or of the current situation of the person regarding social security insurance.
- The SSDREQ message request transmission of information needed to complete social security activities. The information relate to persons, family, health conditions, benefits, incomes, claims, etc...
- The MEDRUC message has initially been designed for dealing with information flows between healthcare institutions and social funding institutions. Build 6 has enlarged its function to also cover the information flows between social security institutions and in order to advise which reimbursements have been made for a provided healthcare.
- The SSCLDE message is sent by a social security institution or organisation to another social security institution or organisation to provide the decision taken on a claim.

Remark: Since this approach is very much influenced by the EDIFACT context and history, this part of the Build 6 outcome is probably of little use for EESSI.

Indeed, the work done by Build 6 on the 4 “Business Functions” would have logically led to the creation of 4 electronic messages (i.e. (1) Collecting information, (2) Authorising to supply/attesting the right to a benefit, (3) Deciding a reimbursement, (4) Deciding to award a benefit.)

However, with the 4 EDIFACT messages were in existence before Build 6 started its activities, the EDIFACT community would not have accepted proposals for new messages that would overlap with existing messages. Build 6 has therefore formulated how these messages best should be adapted, in order to be as close as possible from its approach, without being 100% consistent with it.

Annex 5: Information Repository

Build 6 has not only worked on the flows but also on that data with the view to:

- Have a consistent and comprehensive data model ready to be used when creating intersectorial building blocks of data for electronic messages (See Annex 6)
- Solve semantic interoperability issues since the same data may have a different understanding in different Member States (e.g. the family name of a married women may be her name at birth or the name her husband or a combination of both) (See Glossary of Terms).

Starting from the current forms (as per 1995), Build 6 has identified all information concepts used by the forms, created an information repository (see an example below) and modelled them in order to define the entities, attributes and relationships (see example in Annex 6).

Example extracted from the Information repository

Form (in its version of 2007)

ADMINISTRATIVE COMMISSION
ON SOCIAL SECURITY
FOR MIGRANT WORKERS

E 106 (1)

CERTIFICATE OF ENTITLEMENT TO SICKNESS AND MATERNITY INSURANCE BENEFITS IN KIND FOR PERSONS RESIDING IN A COUNTRY OTHER THAN THE COMPETENT COUNTRY

Employed and self-employed persons and members of their families residing with them; members of the family of unemployed persons

Regulation (EEC) No 1408/71: Article 19.1.a.; Article 19.2.; Article 25.3.i
Regulation (EEC) No 574/72: Article 17.1 and 4; Article 27 (first sentence)

The competent institution should complete Part A of the form and send two copies to the insured person, or send them - where necessary through the liaison body - to the institution in the place of residence if the form is drawn up at that institution's request. As soon as it has received the two copies, the latter institution should complete Part B and return one copy to the competent institution.

Please complete this form in block letters, writing on the dotted lines only. It consists of 4 pages, none of which may be left out.

A. Notification of entitlement

1.	Institution of the place of residence ⁽²⁾	
1.1	Name:	
1.2	Identification number of the institution:	
1.3	Address:	
1.4	Reference: your E 107 form of	
2.	the insured person	
2.1	Surname(s) ⁽³⁾ :	
2.2	Forename(s) ⁽⁴⁾ :	Date of birth:
2.3	Previous name(s):	
2.4	Address in the country of residence:	
2.5	Personal identification number:	
2.6	The person <input type="checkbox"/> is a employed person	

Full Screen
Close Full Screen

Form Fields

IDCode	Description	FFGroup	Values	Data Element	DataElement 2	DataElement 3
E106-1.1-1	Name	E106-1		IC031		
E106-1.1-2	Code number	E106-1		IC030		

IDCode	Description	FFGroup	Values	Data Element	DataElement 2	DataElement 3
E106-1.2	Address (street, number, post code, town, country)	E106-1		IC032		
E106-1.3	Reference: your E107 form of	E106-1		IC083		
E106-10.1-1	Surname	E106-10		IC001		
E106-10.1-2	Forenames	E106-10		IC003		
E106-10.1-3	Sex	E106-10	F – M	IC005		
E106-10.1-4	Previous names	E106-10		IC004		
E106-10.1-5	Date of birth	E106-10		IC006		
E106-11.1	Name	E106-11		IC031		
E106-11.2	Address (street, number, post code, town, country)	E106-11		IC032		
E106-11.3	Stamp	E106-11		IC089		
E106-11.4	Date	E106-11		IC088		
E106-11.5	Signature	E106-11		IC089		

Data Elements (Information Concepts)

ID Code	Name	Description	Type	Values	Usage
IC001	Surname	official name of a person	Text	anything, E: both names at birth	M/I
IC002	Surname at birth	official name of a person at birth, maiden name	Text	anything	O/M/I
IC003	Forenames	first name(s) associated with surname to identify a person	Text	anything	M/I
IC004	Previous names	different names a person had during his life	Text	anything	O/M/I
IC005	Sex		Single selection	male - female	O/M/I
IC006	Date of birth		Date	-	M/I
IC007	Place of birth	locality where a person is born	Structure	place, province, department or county, country	O/M
IC008	Address of person		Structure	street, number, post code, town, country. D: no tel#	M/I
IC009	Nationality	nationality and date of naturalisation	Text	a valid nationality	M
IC010	Father	father's surname and forenames	Text	anything	
IC011	Mother	mother's surname and forenames	Text	anything	
IC012	Registration/Insurance number	unique id of insured person within institution in country	Text	A/D: Vers# B: No.Nat. DK: CPR F: pop.reg.	O/M/I

ID Code	Name	Description	Type	Values	Usage
				S/N: pers# L: AHVins#	
IC013	Pension scheme, dead was receiving	survivor's pension claim	Single selection	none - employed - self-employed - all residents	O/M/I
IC014	Relationship	to the members of the family	Text	anything	
IC030	Code number	institution number	Unknown	unknown	O/M
IC031	Name of institution		Text	anything	M
IC032	Address of institution		Structure	street, number, post code, town, country, tel	O/M
IC033	Type of insurance cover (Pension)	risk, reason justifying the right to social security benefit	Single selection	death/survivor - invalidity - old-age	
IC034	Type of benefit (Health)	for benefits in kind	Single selection	medical treatment - dental treatment - medicaments - hospitalisation - examination - other	
IC035	Address of employer		Structure	street, number, post code, town, country	N/M
IC036	Description of benefit		Text	anything	
IC037	Type of employment I		Single selection	employed - self-employed	O/M

Annex 6: Data Model Diagram

4 modelling diagrams have been developed

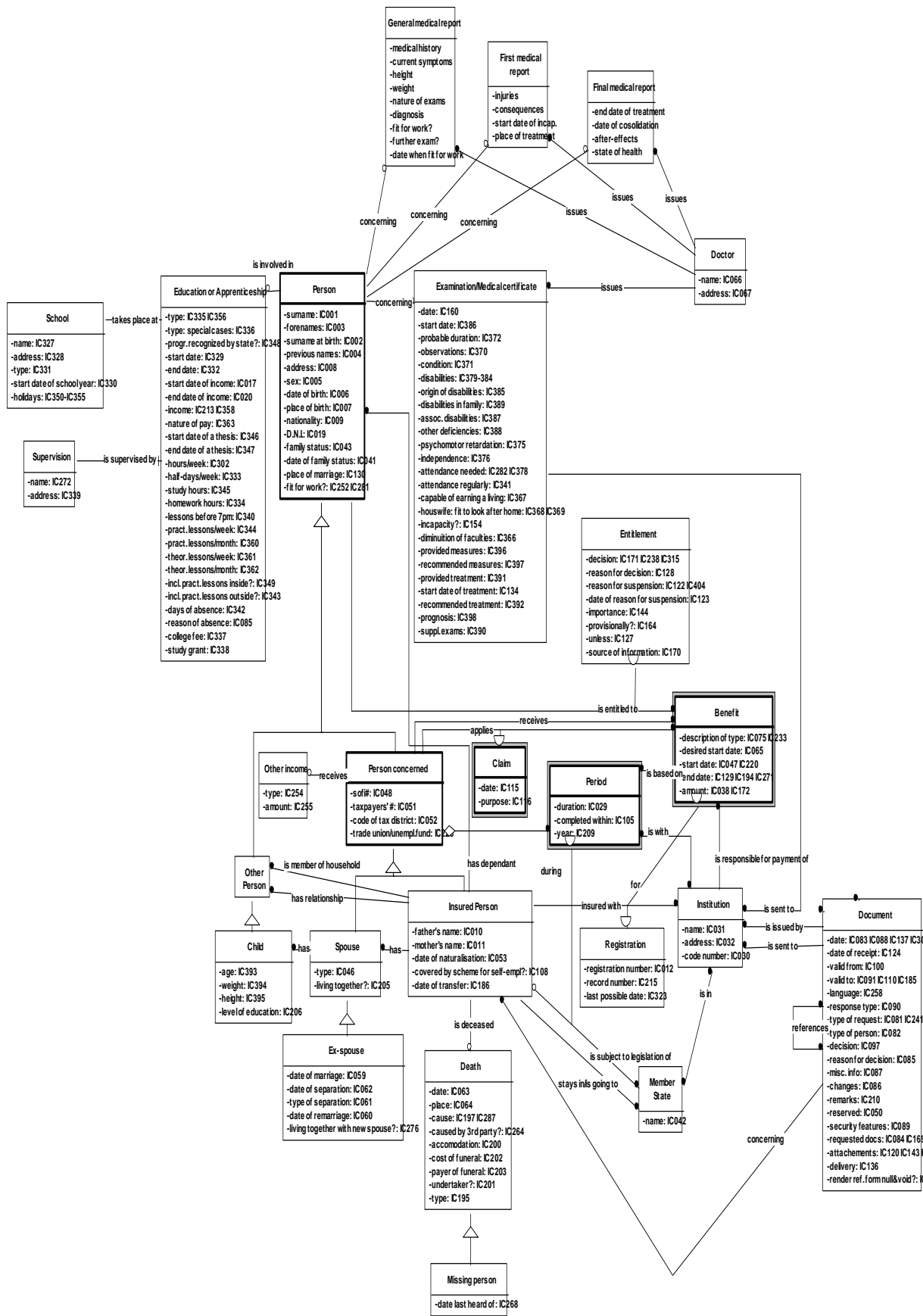
- The main diagram (see below) shows the entities that act as the central part of the model, which are Person and Institution and their relationships and parts
- The period diagram refines all aspects that deal with Periods
- The benefit diagram details the sector specific types of Benefits
- The claim diagram contains all entities relevant for the Claim of a benefit

Each data has also been described from a syntax view point (free text, maximum length, type of value, coded data ...) for further use in the design of the messages.

Remark: Thanks to the fact that the link between the E Forms and the 43 sectorial Exchanges of Data was kept, Build 6 has defined the data content of each flows. Using then the linkage between these flows and the electronic messages, Build 6 has finally defined the data content of each message.

A first observation is useful to introduce the reader in the below complex diagram:

- The concept of a Person acting as a client to social security institutions is the central part of the model and is common to all areas. It has to be distinguished between the Insured person and the Spouse or Ex-spouse of an insured person as the kind of person the institutions directly deal with (Person concerned), and Other persons, e.g. Children, which have to be considered in one way or another by the social security on behalf of the client because they are related to the insured person, are members of their household or are dependants of the insured person.



Annex 7: Glossary of terms (extract)

Furthermore, each data from the information repository has been defined with a generic definition, followed by a definition specific to each Member states with a view to enable an institution to best process the data received from abroad (see example below).

ID	Country	Name	Data element	Description	Values	Remark	Domain
T0001	GENERIC	Surname	IC001	State the official family name of a client.			ALL
T0001	BEL	Surname	IC001	Form E6XX not used Official surname given to a person			ALL
T0001	D	Surname	IC001	HC: Official <u>family</u> surname as it appears in valid personal identity papers”. When using electronic data exchange, form E110 is no longer needed. E123 is exclusively a form relating to statutory accident insurance. CB: Corresponds to the GENERIC Glossary of Terms UN: Official family name (surname) of a person.		For UN : E301-1.1-1 E303-1.1-1	ALL
T0001	IRL	Surname	IC001	State the clients family name. Note: National applications format consists of 20 characters (free text)	Free Text	E105,E206,E211, E215,E302,E402, E403,E404,E406, E407,E413 & E602 not currently used by Ireland	ALL
T0001	LUX	Surname	IC001	OAP (CPACI) : See GOT Framework if not married, same as IC002 otherwise, married name		OAP (CPACI) : wife can keep her maiden name, therefore if name and maiden name are the same, married status is not evident	
T0001	LUX	Surname	IC001	OAP (CPEP): See GOT Framework if not married, same as IC002 otherwise married name		OAP (CPEP): wife can keep her maiden name, therefore if name and maiden name are the same, marital status is not evident	

ID	Country	Name	Data element	Description	Values	Remark	Domain
T0001	LUX	Surname	IC001	OAP (E500): See GOT Framework for a married woman: married name otherwise, maiden name (same as name at birth)			
T0001	LUX	Surname	IC001	UN: See GOT Framework maiden name in the former surnames box (E 301) if married			
T0001	LUX	Surname	IC001	CB: See GOT Framework		CB: If wife, this box is completed with the maiden name for the 401/411	CB
T0001	NED	Surname	IC001	See GOT Framework		For the surname the standard NEN 1888 (the Dutch national standard concerning personal data) applies. The complete name of a person consists of several data-elements that are structured in a standard way. The data regarding the name of a person with the exception of the Christian names, consisting of a significant part and possibly a prefix.	ALL
T0001	POR	Surname	IC001	Full name as it appears on the identity card or passport.			ALL
T0001	UK	Surname	IC001	<i>That part of a person's name which is used to describe the family, clan, tribal group, marital or professional association.</i>		Further work is being done on the use and format of the surname to enable DSS to conform to the legal aspects i.e. the Data Protection Act.	
T0002	GENERIC	Surname at birth	IC002	State the official family name of a client, given at his birth. Note: The maiden name of a woman is equivalent to the surname at birth			ALL
T0002	BEL	Surname at birth	IC002	Form E6XX not used The first official surname give to a person is the surname at birth. This birth name can be replaced by another name following an adoption or a request for a change of name.			ALL

ID	Country	Name	Data element	Description	Values	Remark	Domain
T0002	D	Surname at birth	IC002	<p>HC:</p> <p>Corresponds to GENERIC Glossary of Terms</p> <p>Comment: In form E122-3.2-2 there is the concept "Previous surnames".</p> <p>This does not correspond to data element IC002!!.</p> <p>The concept "Maiden surname" in forms E101 and E102 is replaced in brief by "Previous surnames".</p> <p>OAP:</p> <p>See GOT Framework</p> <p>CB:</p> <p>Corresponds to the GENERIC Glossary of Terms</p> <p>UN:</p> <p>Official surname of a person at birth. In the UN forms : Former name</p>	<p>UN:</p> <p>see also IC004</p>	<p>For OAP:</p> <p>"Maiden surname" = "Surname at birth".</p> <p>For HC:</p> <p>"Previous name" (IC004) = "Previous surname held" is not the same as "Surname at birth", therefore the concept of the "Maiden surname" should not be replaced by "Previous surnames"</p> <p>For UN:</p> <p>Forms E303-1, 1-3, E303-1.1-3</p>	All
T0002	IRL	Surname at birth	IC002	<p>State the clients surname at birth.</p> <p>Note: National applications format consists of 20 characters (free text)</p> <p>E213-2.2.2 does not refer to Maiden Name</p>	Free Text	<p>E213,E215 &</p> <p>E413 not currently used by Ireland</p>	ALL
T0002	NED	Surname at birth	IC002	<p>{Though maiden name is a common often used term, it is hard to describe it in a clear way. This is caused by the fact that in our opinion a so called maiden name of a person don't necessarily has to be the same as the surname at birth of that person (due to f.e. adoption or change of surname).The national standard concerning personal data, as well as the GBA (Municipal basis administration) which is the formally authorised source of personal data in the Netherlands, therefore do not distinguish maiden name as a data-element}</p>		<p>The complete name of a person consists of several data-elements that are structured in a standard way.</p> <p>Since a surname is part of the complete name of a person, the standard structure also applies to surname at birth.</p> <p>For addressing a married woman our Social Security domain uses the composite data-elements surname and surname husband, but only if the woman wishes to be addressed that way.</p>	ALL
T0002	POR	Surname at birth	IC002			Not used	
T0002	UK	Surname at birth	IC002	<p><i>The surname given to a person at birth as used for the exchange of information with member states of the European Union.</i></p>		<p>This item is also currently under discussion. The validation procedures associated with this item in the UK is very stringent and the same procedures outside the UK are not necessarily applied or indeed if the same definition is understood. We are therefore looking at the use of another item (possibly base surname) to interpret the information received from the EEC members thereby complying with our current legislation.</p>	

Annex 8: Document Flow Diagram (simplified)

