D3.2 Final Requirements for Semantic Assets

Document Identification			
Status	Final	Due Date	31/12/2021
Version	1.0	Submission Date	25/02/2022

Related WP	WP3	Document Reference	D3.2
Related Deliverable(s)	D2.1, D2.4, D3.1, D3.5, D4.1, D7.1	Dissemination Level (*)	PU
Lead Participant	IHU	Lead Author	C.Berberidis (IHU)
Contributors	T. Karunaratne, E. Kontopoulos (SU); A.-R. Guzman (MTFP-SGAD); Slavko Žitnik, Dejan Lavbič (SI-MPA) S.-Iftikhar. H Shah, E. Kapantai, I. Konstantinidis (IHU)	Reviewers	Fredrik Lindén (SU)
			K. Kern Pipan and M. Jesenko (SI- MPA)

Keywords :
Requirements Elicitation, Semantic Assets, Core Vocabularies, Co-creation, Semantic Framework, Canonical Evidences, Evidence Exchange, Semantic Model, Information Desk, Public Procedures, Information Exchange Model

Document Information

List of Contributors	
Name	Partner
Christos Berberidis	IHU
Syed Iftikhar H. Shah	
Ioannis Konstantinidis	
Eleni Kapantai	
Thashmee Karunaratne	SU
Efstratios Kontopoulos	
Ana-Rosa Guzman	MTPF-SGAD
Slavko Žitnik	SI-MPA
Dejan Lavbič	

Document History			
Version	Date	Change editors	Changes
0.1	20/10/2021	IHU	Initial version of the document
0.2	12/11/2021	IHU	Structure and outline of the document proposed
0.3	16/11/2021	IHU, SU, SI-MPA, MTFP-SGAD	Structure and outline of the document revised and agreed
0.4	19/01/2021	IHU, SU	1st version of the document (Content to chapters 1-3 and 4.1)
0.5	25/01/2021	SU, SI- MPA, MTFP-SGAD	2nd version of the document (Content to chapters 4.2,4.3 and 5)
0.6	28/01/2021	IHU, SU, SI-MPA, MTFP-SGAD	Final version of the document after internal revisions
0.7	31/01/2021	IHU	Submission for initial review
0.8	07/02/2022	IHU, SU, SI-MPA, MTFP-SGAD	Document after 1st reviewer's comments
0.9	11/02/2022	IHU, SU, SI-MPA, MTFP-SGAD	Document after 2nd reviewer's comments
0.95	17/02/2022	IHU, SU, SI-MPA, MTFP-SGAD	Final version of the document after overall review and content improvements
1.0	25/02/2022	Julia Wells (Atos)	Quality check and updates for submission

Quality Control		
Role	Who (Partner short name)	Approval Date
Deliverable leader	Christos Berberidis (IHU)	25/02/2022
Quality manager	Julia Wells (ATOS)	25/02/2022
Project Coordinator	Ana Piñuela Marcos (ATOS)	25/02/2022

Table of Contents

List of Tables

List of Figures

List of Acronyms

Abbreviation / acronym	Description
AC	Authorisation Controller
ADMS	Asset Description Metadata Schema
ATU	Administrative Territorial Unit
BB	Building Block
BPMN	Business Process Model and Notation
BRegDCAT-AP	Data Catalog Vocabulary Application profile for base registries
BRIS	Business Registers Interconnection System
CCCEV	Core Criterion and Core Evidence Vocabulary
CE	Canonical Evidence
CPOV	Core Public Organization Vocabulary
CPSV	Core Public Services Vocabulary
CPSV-AP	Core Public Services Vocabulary - Application Profile
CQ	Competency Question
DBA	Doing Business Abroad
DC	Data Consumer
DCAT	Data Catalog Vocabulary
DCAT-AP	Data Catalog Vocabulary- Application profile
DBA	Doing Business Abroad
DE	Data Evaluator
DE4A	Digital Europe for All
DO	Data Owner
DoA	Description of Action
DP	Data Provider
Dx.y	Deliverable number y, belonging to WP number x
EC	European Commission
EDCI	European Digital Credentials Infrastructure
eID	Electronic Identification
eIDAS	Electronic Identification and Signature (EU regulation)
EIF	European Interoperability Framework
EIRA	European Interoperability Reference Architecture
e-SENS	Electronic Simple European Networked Services
ESL	Evidence Service Locator
EuroVoc	EU Vocabularies
FOAF	Friend of a Friend
IAL	Issuing Authority Locator
IDK	Information Desk
IEM	Information Exchange Model
IHU	International Hellenic University

Abbreviation / acronym	Description
IOP	Interoperability
ISA2	Interoperability solutions for public administrations, businesses and citizens
JSON	JavaScript Object Notation
MA	Moving Abroad
ML	Machine Learning
MOR	Multilingual Ontology Repository
NACE	Nomenclature generale des Activites economiques dans les Communautes europeennes
OOP	Once-Only Principle
OOTS	Once-Only technical system
OWL	Web Ontology Language
PKI	Public Key Infrastructure
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
RIHA	Riigi Infosüsteemi Halduse Infosüsteem
SA	Studying Abroad
SAML	Security Assertion Markup Language
SCOOP4C	Stakeholder Community Once-Only Principle for Citizens
SDG	Single Digital Gateway
SDGR	Single Digital Gateway Regulation
SEMIC	Semantic Interoperability Centre Europe
SHACL	Shapes Constraint Language
SKOS	Simple Knowledge Organisation System
TL	Task Leader
TOOP	The Once-Only Principle Project
UUID	Universally Unique Identifier
URN	Uniform Resource Name
W3C	World Wide Web Consortium
WP	Work Package
WPL	Work Package Leader
XML	Extensible Markup Language
XSD	XML Schema Definition

Glossary

Term	Explanation
Canonical Evidence	Structured data models that include a common set of attributes associated with the evidence type that can be provided by the corresponding lawfully issued evidences [37].
Co-creation	A process that brings together users and designers to work towards a shared goal.
Code list	Predefined set of terms from which some statistical coded concepts take their values [38].
Competent Authority	The competent authorities for the procedures are most often public bodies at the national, regional and local levels – such as Ministries, National Social Security and Pensions systems, Regulatory Authorities, Local authorities and others. However, the competent authority may be a private body or institution acting with a mandate from public authorities [36],[39].
Competency Question	Natural language query that expresses a pattern for a type of question the ontology should be able to answer.
Controlled Vocabulary	A consistent way to describe data. They are standardized and organized arrangements of words and phrases presented as alphabetical lists of terms or as thesauri and taxonomies with a hierarchical structure of broader and narrower terms [40].
Criterion	(1) A standard on which a judgement or decision may be based [41]. (2) A standard by which you judge, decide about, or deal with something [42]. (3) A condition or fact used as a standard by which something can be judged or considered [42].
Event	In general, an event is something that happens. It is an arbitrary classification of a space/time region, by a cognitive agent. An event may have actively participating agents, passive factors, products, and a location in space/time. For example, in computer coding, an event is when something happens that triggers the code to run.
Evidence	Any document or data, including text or sound, visual or audio-visual recording, irrespective of the medium used, required by a competent authority to prove facts or compliance with procedural requirements referred to Article 2.2.b (SDGR) [36].
Federated OOP Architecture	The European Commission is working to develop a generic Federated OOP Architecture, through different initiatives/projects supporting the interconnection and interoperability of national base registries across state borders. The aim is to provide consolidated reusable Building Blocks (BBs) for the implementation of the "once-only" principle in public services in Europe. From a methodological point of view, such an architecture will not be developed from scratch. Efforts have been made in the development of generic building blocks for European cross-border public services.
Once Only Principle	The public administrations should ensure that citizens and businesses can supply the same information only once to a public administration and administrations should be able to retrieve and share this data to serve the user, in accordance with data protection rules.
Ontology	An ontology is a formal, explicit specification of a shared conceptualisation. In computer science and information science, an ontology encompasses a representation, formal naming and definition of the categories, properties and

Term	Explanation
	relations between the concepts, data and entities that substantiate one, many or all domains of discourse [43].
Legal Person	A legal person is a registered organization, having its registered office in a Member State [36].
Natural Person	A natural person is a citizen of the Union or a human residing in a Member State [36].
Proof	(1) Fact or piece of information that shows that something exists or is true [44], (2) Something that induces certainty or establishes validity [45], (3) Evidence operating to determine the finding or judgement of a tribunal [45].
Public Service	The concept of public service is twofold: it embraces both the bodies providing services and the services of general interest they provide. Public service obligations may be imposed by the public authorities on the body providing a service (airlines, road or rail carriers, energy producers and so on) either nationally or regionally [46].
Scenario	One typical way in which a system is used or in which a user carries out some activity.
Semantic Asset	A specific type of standard which involves highly reusable metadata (e.g., xml schemata, generic data models) and/ or reference data (e.g., code lists, taxonomies, dictionaries, vocabularies).
Taxonomy	A taxonomy is a scheme of categories and subcategories that can be used to sort and otherwise organise items of knowledge or information [47].
Technical Coordinator	The Technical Coordinator is a senior technical expert who will facilitate the smooth execution of the whole development lifecycle within the DE4A project.
Thesaurus	A list of words and their synonyms about a domain of knowledge.
TOOP	The Once-Only Principle Project (TOOP) was launched by the European Commission in January 2017 as an initiative of about 50 organizations from 20 EU Member States and Associated Countries. The said project was ended on 31st March, 2021. The main purpose of TOOP was to investigate and explain the once-only principle on a cross-border pan-European scale, focusing on reducing the administrative burden of businesses [4].
Use case	A specification of one type of interaction with a system. One use case may involve several scenarios (usually a main success scenario and alternative scenarios).
User	User is anyone who is a citizen of the Union, a natural person residing in a Member State or a legal person having its registered office in a Member State, and who accesses the information, the procedures, or the assistance or problem-solving services, referred to in Article 2(2), through the gateway [36].
Vocabulary	A collection of terms for a particular purpose. Vocabularies can range from simple, such as the widely used RDF schema, FOAF and DCMI element set to complex vocabularies with thousands of terms, such as those used in healthcare to describe symptoms, diseases and treatments. Vocabularies play a very important role in linked data, specifically to help with data integration. For example, metadata vocabulary. The use of this term overlaps with that of 'ontology' within the scope of computer and information sciences [48].

Executive Summary

This deliverable exhaustively provides the final requirements as far as semantic interoperability is concerned within the context of the Digital Europe for All (DE4A) project. To achieve this, analysis of existing taxonomies and libraries of core vocabularies was performed, presented already in previous deliverable D3.1 ''*Initial Requirements for Semantic Assets*'' [1]. While the present document focuses on the final set of requirements for semantics in DE4A, it also defines a *requirements elicitation methodology* that is important to be followed by similar initiatives in the future.

Semantic interoperability faces various common issues at national and at European Union (EU) level in terms of semantic and syntactic aspects. The example of such common issues include lack of harmonised or common data structures, minimal agreements on reference data, lack of common management metadata, and absence of unique identifiers for the public services and public organisations. Moreover, the semantic assets of the EU ISA^2 (Interoperability solutions for public administrations, businesses and citizens) Programme constitute a good initiative in terms of academic rather than a practical view. The EU ISA^2 programme has now evolved into interoperable Europe [2].

To tackle semantic interoperability common issues, work package WP3 – "*Semantic Interoperability Solutions*" was tasked to **i)** analyse and extend the existing taxonomies and libraries of core vocabularies, **ii)** to design the semantic interoperability framework and **iii)** implement semantic tools.

This deliverable represents the final validated requirements for the DE4A semantic assets along with their definitions and clear justifications. A co-creation methodology was implemented in order to gather all semantic interoperability requirements from the pilots. The reader of this document will be aware about semantic assets related efforts in terms of initiatives, and projects, and standards, in the field of e-government, emerging not only at the European level but also at a worldwide level. Moreover, such efforts are a potential baseline to build a common repository of semantic models and business rules.

The main outputs of this deliverable are:

▸ the requirements elicitation methodology that is based on an agile co-creation methodology. It serves as a general methodology for semantic requirements collection and semantic interoperability solutions design

▸ the requirements for the DE4A semantic framework based on the outcomes from relevant projects and initiatives regarding semantics. These focus on the underlying semantic frameworks (organised semantic initiatives, EIF recommendations [3], TOOP [4]and SCOOP4C guidelines [5]) and requirements for the DE4A semantic assets (DE4A Pilot's canonical evidences, Information Desk architecture framework and Information Exchange Model (IEM) definitions). Existing core vocabularies from European Commission (EC) and domain specific ontology libraries identified are the ISA^2 Core vocabularies (Core Public Services Vocabulary - CPSV, Core Public Organisation Vocabulary - CPOV) and generic ontologies such as Asset Description Metadata Schema (ADMS), Dublin core, Friend of a Friend - FOAF

▸ final requirements for the DE4A semantic assets - Information Desk (IDK), Information Exchange Model and all the canonical evidences (used in all three DE4A pilots).

1 Introduction

The SDGR [6]establishes the obligation on EU Member States to facilitate access to and completion of online administrative procedures by cross-border businesses and citizens. Such facilitation also includes direct exchange or verification of lawfully issued evidences between competent authorities of different Member States by electronic means in application of the once-only and relevant-only principles. Consequently, cross-border evidences should be able to be processed in any Member State, which requires a sound semantic interoperability approach to be put into practice. Within the European context, several semantic interoperability initiatives have been aimed at delivering integrated cross-border public services by developing taxonomies and vocabularies for certain issues. Some of them are for specific-domains and rely on real field tests, and others are for the general domain (c.f. D3.1 "*Initial requirements for semantic assets*").

The ambition of the semantic interoperability solutions for DE4A is to develop a comprehensive map of modular semantic building blocks needed for delivering integrated cross-border public services, particularly in the context of the SDGR fully online public services. The main focus here is not to provide novel semantic assets for interoperability but a full semantic layer for delivering cross-border public services that will reuse, as much as possible, those that are already available. This layer will be defined by modular semantic assets and their relationships, widely described with complete requirement specifications and a practical implementation guide that is tested in DE4A pilots. The inventory of these semantic assets along with their requirements, relationships and, when applicable, implementations will be provided in a way that eases the maintainability of the coherence of the semantic layer.

 Ultimately, DE4A will help overcome the fragmented views of the public service concept (e.g., following different flavours of national, regional or local traditions) that impact on the quality and the efficiency of public service provision for cross-border users by increasing administrative burdens and making public service provision costly. Within this context, this deliverable correlates with Pilots for Citizen & Business (WP4 "Cross-border Pilots for Citizens and Business and Evaluation"), WP5 "Common Component Design & Development" and WP2 "Architecture Vision and Framework" to jointly form the agile development process within the project. The focus is to get the pilot implementations working and accepted in real life, while adhering to the agreed PSA (Project Start Architecture) from WP2 and adhering to design principles that ensure a maximum reusability beyond the project. The overall semantic interoperability setting is illustrated in Figure 1:
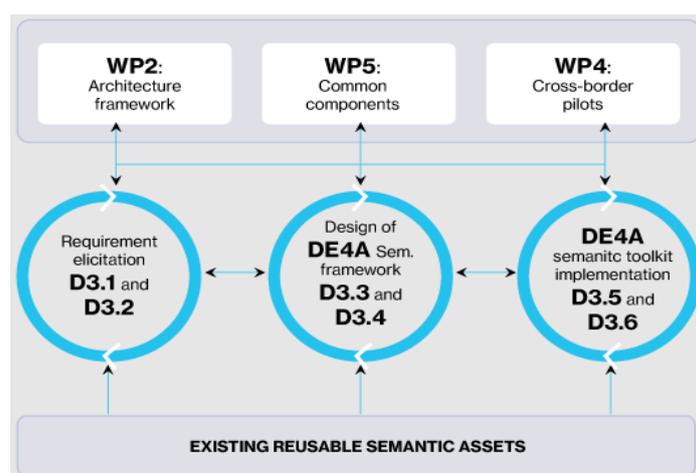
Figure 1: The semantic interoperability approach and respective DE4A deliverables

This deliverable incorporates a semantic requirements definition and looks beyond the pilots in order to close the perceived gap in semantic interoperability in the cross-border information exchange

through the once-only technical system today. Subject document and WP3 in general will build and expand on prior work from ISA (i.e., SEMIC, CPSV) and W3C and deliver semantic components, ready to use in the WP4 pilots.

1.1 Purpose of the Document

The purpose of this document is to elicit and discuss the final set of semantic interoperability requirements, which provide the basis for the development of the DE4A semantic interoperability framework. The set of requirements has been developed incrementally by accommodating the outcomes from pilots, analysis of results of other projects like TOOP [4]and other requirements of emerging vocabularies, taxonomies, dictionaries, libraries and ontologies. Additionally, the description of how the semantic requirements are elicited, using an agile co-creation methodology, by bringing together the users (DE4A pilots) and designers (WP3 team) to capture and align semantic interoperability needs to define a DE4A semantic interoperability framework, is illustrated in the subsequent sections of this document.

WP3 implements the "*reuse before adapt before develop*" principle in the development of all semantic assets. This means that, for any given feature, it must be investigated whether existing, preferably open source, solutions are available that can be reused or can be adapted to fulfil the requirements (D3.5 "*Semantic toolkit – Initial version*"[7]& D3.6 "*Semantic toolkit – Final version*"). Any technical design document for a bespoke solution in DE4A shall summarise this investigation in the introduction and explain why existing solutions were considered unfit for the purpose.

Moreover, this document also briefs the contributions in dealing with the challenge of availability and accessibility limitations of open data and semantic assets, attempting to bridge the gaps of interoperability that may occur due to the limited availability and accessibility to data. Towards this, it identifies the semantic assets and proposes a framework to get feedback by the pilots retrospectively.

1.2 Structure of the Document

This deliverable is structured as follows:

▶ **Chapter 2** introduces the adopted agile methodological framework for requirement elicitation.
▶ **Chapter 3** presents the requirements for the DE4A semantic framework that were identified following recommendations, guidelines and principles derived from EIF, SDGR and from related EU projects/initiatives, aligned with the DE4A principles. The results are presented using a mapping table and combining EIF recommendations [3] to DE4A requirements.
▶ **Chapter 4** describes the requirements for semantic assets proposed by the DE4A project. The following types of requirements are considered: (a) requirements for the canonical evidence, (b) requirements for evidence exchange, and (c) requirements for the provision of public procedures.

2 Requirements Elicitation Methodology

Agile is an iterative way to build a product, through stakeholder collaborative efforts, that is evolved on an incremental basis instead of delivering it at once near the end of a project. **Co-creation** brings together users and designers to capture and align needs to design and test a product. The combination of these two aspects helped WP3 to engage with DE4A Pilots (users) through WP4 and to capture their semantic asset requirements and experiences about semantic interoperability.

The Agile co-creation methodology is a way to manage a product by breaking it up into several stages. To take advantage of this approach, it is important that product requirements may also be defined through the agile framework [8].For example, in the Scrum lifecycle for user stories, to document product requirements, a requirement analyst describes the product to be developed through product backlog. Wherein a backlog is a set of user stories to elicit its detailed requirements. A user story is an agile way to explain the user needs about an artifact (software, semantic component, etc.) with an aim to show how artifacts will offer value to the users [8]. Moreover, product backlog items are selected as per the priorities for implementation in successive iterations which is known as Sprints. Usually, a user story in Scrum goes through the following states:

<p align="center">Co-created → Accepted → Estimated → Planned → Ongoing → Done</p>

This is only the minimal path; this basic view does not signify all the possible states, nor the loops that may arise. In the literature, we also observed two additional superstates to focus two periods i.e., maturing and selected [8]. The maturing superstate means the requirement or user story is being explained and analyse. The Selected superstate is essentially implemented in the product. We also followed other states Issue and Dismissed to manage refusal from stakeholders and defects not fixed in the current iteration. In this way, we attain Figure 2 as shown below:
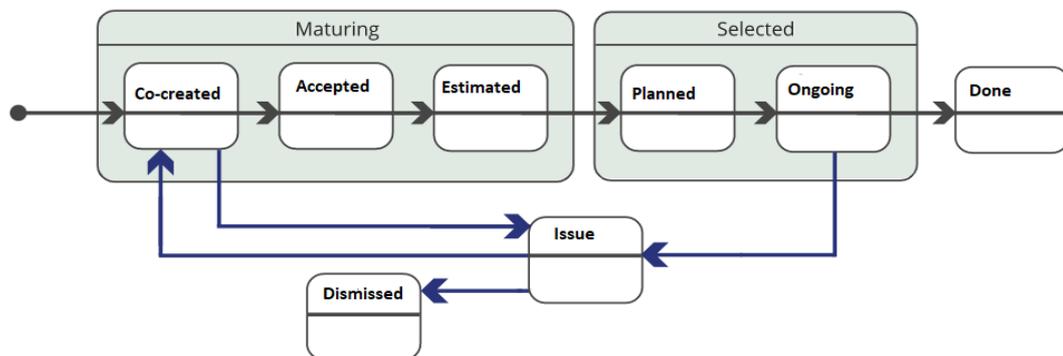
<p align="center">Figure 2: Lifecycle of a product backlog item in Scrum [8]</p>

DE4A also adopted the above-mentioned lifecycle to elicit pilot requirements and to complete the related activities of our semantic products for three pilots of the project. We used the primary mode 'Scrum' of Atlassian JIRA Agile tool for WP3 product backlog and for iterations planning, and other such activities of our work. Additionally, we applied 'Kanban' for WP3 constraint-based task management as well.

2.1 Introduction

The notion of **agile co-creation** is diverse from the traditional "push-and-pull" approaches, as it denotes that diverse parties 'create' something together, instead of one fragment developing

something for the other one to use (push-approach) or expressing a clear request or need to the other (pull-approach). When parties are expected to create together, they must be equal partners with a similar level of resources and speak a common language towards a shared goal or value.

This agile co-creation approach brings together users (DE4A pilots) and semantic asset designers & implementers (WP3 team) to capture and align semantic interoperability needs (D3.2 "*Final requirements for the semantic assets*"), design a semantic interoperability framework (D3.4 "*Design of the semantic interoperability framework*"), and implement semantic assets for delivering cross-border public services (D3.6 "*Semantic toolkit – Final version*").

In view of the above, we adopt the aforementioned approach in WP3 to design and implement the required semantic assets within DE4A. In this regard, we integrate works of existing initiatives (e.g., SEMIC, ISA[2], SDG) into an extensible multi-domain, cross-border, and cross-sector semantic interoperability framework.

2.2   Alignment with Project

The aim of this approach provides a wide methodology for agile co-creation, that is used as a template allowing DE4A WP3 partners (IHU, MPTFP-SGAD, SU, SI-MPA) to co-design the set of internal mechanisms. This co-creation is performed internally within DE4A WP3 and via liaison with expert groups on semantic assets, while ensuring the agreed dependencies between WPs of the DE4A project, particularly DE4A Pilots. The co-creation sessions (e.g., semantic workshops, focused group sessions, etc.) are carried out with the DE4A piloting EU Member States through WP4. To do so, WP3 establishes a close liaison with WP4 partners. The desired outcomes are shared with WP5 "Common Component Design & Development'', the DE4A Technical Coordinator, and with WPs (WP2-3-4-5) Technical Working Groups for the design and development of the project architecture and other common technical components as per the scope of the project.

2.3   An Agile Co-creation Methodology for Requirements Elicitation

The agile co-creation approach helps to overcome the push-and-pull schism and stimulate a vibrant, competitive co-creation process within WP3. One of the main impediments to open and collaborative innovation is the complexity of having researchers and public officials speak the same language and addressing the problems from a common perspective and with analogous resources and tools. It is a new way of driving research, with and for the market, at corporate and public-sector organisations. The following iterative steps define the methodology:

I.   Identification of Points of Contact from the relevant DE4A Pilots (Moving Abroad, Studying Abroad, Doing Business Abroad), stakeholders within WP3, including piloting Member States (partners) of the DE4A project.

II.   Design and organization of different online events (e.g., semantic assets workshops, focused group sessions, etc.) with DE4A Pilots, to bring together users (DE4A pilots) and designers, and implementers (WP3 team) to capture and align semantic interoperability. We detail such events in the forthcoming sections of this chapter. To achieve this goal, in these events, different exercises, like user stories, were performed.

III. Gathering of semantic assets needed from the Pilots by using different techniques for interactions. Such techniques include open-ended competency questions, short surveys, and brainstorming sessions with pilots.

IV.   Consolidate all DE4A pilots' semantic assets needs.

V.   Prioritise semantic assets issues to be addressed in the proposed DE4A semantic interoperability framework (D3.4 "*Design of the semantic interoperability framework*").

VI. Perform analysis, integration of existing taxonomies, vocabularies, and libraries based on relevant project stakeholders, including piloting Member states' feedback. Subsequently, WP3 partners also proposed such sematic assets, and libraries for DE4A semantic interoperability framework.

VII. Through a co-design process, WP3 team designed a DE4A semantic interoperability framework, and implemented a semantic toolkit to deliver cross-border public services within the EU.

VIII. Obtain ideas, testing, validation, and evaluation of the DE4A semantic interoperability framework, and DE4A semantic assets through testbeds at DE4A Pilots.

IX. Share the final design of DE4A semantic interoperability framework and semantic toolkit with WP5 and DE4A technical coordinator, who will facilitate the smooth execution of the whole DE4A development lifecycle.

One of the prime objectives is to create open innovation proposals and activities, like events, with DE4A Pilots, other related internal (at project level) and external stakeholders. The DE4A Pilots also have close liaison with respective civil societies and businesses as well and mobilise participation in the identification and prioritisation of concrete societal goals. For example, the DBA pilot requested companies from the Netherlands, Romania, and Sweden to participate and contribute for making doing international business easier in the pilot.

Once the innovation is conceptualised, building from both technology-push and market-pull principles, and is understandable to all DE4A Pilots, the co-creative process that may lead to its implementation can start, as long as the necessary resources are gathered. The project consortium suggested testbeds at DE4A Pilots in the form of DE4A iteration-1, iteration-2, etc. The purpose of testbeds is to create a shared arena to obtain ideas, testing, validation, and evaluation of the design of the semantic framework, and implementation of the semantic toolkit. About Quadruple Helix's model implementation, shown in Figure 3 below, it is indicated that there is an essential need to create innovation processes based on a semantic interoperability framework, that all DE4A Pilots should be an active part within WP3.

The implementation of WP3 objectives converges to a 9-step methodology (see above) that deploys our approach to open, collaborative innovation. We believe that this approach fully captures the complexity of the process and allows for innovation to occur iteratively at every step along the way, contributing to the full achievement of objectives of WP3.

Figure 3: Quadruple Helix model implementation [10]

In Figure 3 above, when we refer to culture, it describes the way WP3 partners contribute within the ecosystem. By understanding the *diverse cultures* inside the diverse set of contributors, the ecosystem can reshape behaviours, and, in turn, create a stronger intramural culture that supports the unique objectives as per the scope of WP3. Developing a systematic approach to evaluate and familiarise the cultures, the ecosystem can maximise the potential within the co-creation process to build a common repository of semantic models and business rules. By *practice*, we refer to the co-creation as well as the DE4A pilot's context needs, its public settings (rules and procedures), and finally, the current routines of interaction between actors i.e., *stakeholders*. Then the *structure* is crucial to keep the motivation and production of the co-creation stakeholders, promote interaction and innovation while linking the activities to the objectives [8]. It is the main construct of the ecosystem that promotes value sharing and value acquisition. The structure includes semantic technology, management, reports, and communications with internal and external stakeholders. Finally, with evolution, we refer to the feedback loop. Constant engagement and feedback mechanisms from the *stakeholders*, particularly DE4A pilots that are fed back to the design of the semantic framework for adjustments and enhancements, and implementation of the semantic toolkit.

Furthermore, presented below are the key ingredients driving the process of value co-creation within innovation in the public sector. These include the ability to transform the perception of WP3 stakeholders, including pilots, as recipients of solutions to equal partners in the design process of a semantic framework, and implementation of a semantic toolkit. The building of capabilities and a sense of mutual development, along with the blurring of traditional power roles, is what characterises this process.

2.4   WP3 Stakeholders

The WP3 main stakeholders include IHU (EL), Ministerio de Politica Territorial y Función Pública (ES), Stockholms Universitiet (SE), Ministry of Public Administration (SI), the Piloting (EU) the Member State, WP2, and WP5. The DE4A partners that are working on other WPs of DE4A are also stakeholders of WP3**.**

2.5   Pilot Context and Events

The DE4A pilots are essential for WP3 to get their semantic needs to define/design, and evaluate DE4A semantic framework, and to implement DE4A semantic toolkit. WP4 is responsible for dealing with the pilots. It is pertinent to mention that the Technical Coordinator and WPs (WP2-3-4-5) Technical Working Groups were also invited to participate in the following WP3 events and support the WP3 team to organise such events:

▸ Semantic workshops,
▸ Focused group meetings,
▸ Dedicated presentation sessions about semantic aspects in the DE4A General Assembly meetings,
▸ Contribution and participation in DE4A Technical Group Meetings.

2.6   Legal and Ethical Guidelines

The WP3 team fully adopts the work taking place within WP7 "Legal and ethical compliance and consensus building", and especially the outcomes reported in D7.1 *"Overview of legal and ethical requirements"* [11] that are in line with GDPR for data collection, storage, and processing requirements from the WP3 stakeholders, including DE4A pilots, and civil society**.**

3 Requirements for the DE4A Semantic Framework

To identify the semantic requirements, DE4A follows recommendations, guidelines and principles derived from EIF, SDGR and from related EU projects/initiatives, aligned with DE4A principles mentioned in the deliverable D2.1 ''Architecture Framework" [12]. The results are presented in the table below. Thus, recommendations and guidelines derived from our initial study of relevant projects and initiatives follow. The table will be extended and mapped based on detailed investigation of other such guidelines to conclude a concrete list of (potential) semantic assets specific for DE4A. To cover the subject aspects, the table, starts by mapping the above-mentioned EU projects/initiatives to DE4A semantic requirements.

*Note: In addition to the Table 1, *we describe DE4A project pilots' specific semantic requirements in the forthcoming section of this document.*

Table 1: Mapping of EIF Recommendations and TOOP and SCOOP4C Guidelines to DE4A Requirements

ID	Related Initiative	Guidelines/Recommendation	Initial Requirement/ Need	Potential Semantic Assets
SR-1	EIF	Recommendation (R) 30: Perceive data and information as a public asset that should be appropriately generated, collected, managed, shared, protected and preserved.	Need for **semantic components as assets** in DE4A for preserving, collecting, managing and requesting related information by data actors (DC and DP).	DCAT-AP
SR-2	EIF	R31: Put in place an information management strategy at the highest possible level to avoid fragmentation and duplication. Management of metadata, master data and reference data should be prioritised.	Overcome information sharing and interoperability problems and understand data assets across borders with **agreements on reference data**, in the form of taxonomies, controlled vocabularies, thesauri, code lists and reusable data structures/models.	Use of W3C standards to provide taxonomies etc. Dublin Core OP Core Metadata Element set [13] Named Authority Lists (NALs) [14] FOAF SKOS
SR-3	EIF	R32: Support the establishment of sector-specific and cross-sectoral communities that aim to create open information specifications and encourage relevant	**Base registries** should provide such identifiers, which help to differentiate persons with the same name, allows	TOOP Exchange Data Model[15] e-Documents and Semantics Building Blocks

ID	Related Initiative	Guidelines/Recommendation	Initial Requirement/ Need	Potential Semantic Assets
		communities to share their results on national and European platforms.	to keep track of the company even if it changes the name, etc.	provided by e-SENS eIDAS identification schemes
SR-4			**DE4A policy for identifiers**.	
SR-5			**Robust, coherent and universally applicable information standards** and specifications are needed to enable meaningful information exchange among piloting Member States, while considering the different linguistic, cultural, legal, and administrative environments between them.	
SR-6	EIF/DE4A	Openness principle: Ensure a level playing field for open-source software and demonstrate active and fair consideration of using open-source software, considering the total cost of ownership of the solution.	Use of **open-source software** for pilot ontology serialisation, management, editing and storage.	XML Altova [16], Vocbench 3 [17]
SR-7	EIF/DE4A	Reusability principle: Reuse and share solutions and cooperate in the development of joint solutions when implementing European public services.	**Reuse of existing standards** for data models and vocabularies, as well as of data models from other related EU initiatives	ISA^2 Core Vocabularies, TOOP Registered Organisation ontology [18]
SR-8	EIF/DE4A	User-centricity principle: Provide a single point of contact to hide internal administrative complexity and facilitate users' access to European public services.	Use of **central semantic components for identifying issuing authorities of evidence and the related evidence services**.	TOOP Criterion & Evidence Type Rule Base [19], TOOP Data Service Directory [20]

ID	Related Initiative	Guidelines/Recommendation	Initial Requirement/ Need	Potential Semantic Assets
SR-9	EIF/DE4A	Security and privacy principle: Define a common security and privacy framework and establish processes for public services to ensure secure and trustworthy data exchange between public administrations and in interactions with citizens and businesses.	Use of **semantic components for supporting the DP to check if the DC has the required authorization** for making a request.	TOOP Registry of Authorities [21]
SR-10	EIF/DE4A	(DE4A: Inclusion & accessibility), EIF multilingualism principle: Use information systems and technical architectures that cater for multilingualism when establishing a European public service. Decide on the level of multilingualism support based on the needs of the expected users.	Need for a **semantic component that will facilitate the multilingual representation of data models** for evidence to be exchanged by using multilingual controlled vocabularies.	TOOP Semantic Repository [22], EuroVoc
SR-11	EIF/DE4A	Assessment of effectiveness and efficiency principle: Evaluate the effectiveness and efficiency of different interoperability solutions and technological options considering user needs, proportionality and balance between costs and benefits.	Use of **Semantic Web standards and ontology metrics for validating ontologies**.	SHACL
SR-12	TOOP	Build a Federated OOP Architecture. Support interconnection and interoperability of national registries at the EU level.	**Support the interconnection and interoperability between DE4A piloting Member States registries** by using a standard data model / specification for base registries access and interconnection.	BRegDCAT-AP [23]
SR-13	TOOP	Development of an information exchange model for the payload specification of the messages to be exchanged between competent authorities	Need to **develop an information exchange model** for describing requests and responses for evidence exchange between piloting Member States. Additionally, such a model needs to be agnostic to any	TOOP Exchange Data Model, eIDAS SAML Attribute Profile, SEMIC Common data types, XML Schema

ID	Related Initiative	Guidelines/Recommendation	Initial Requirement/ Need	Potential Semantic Assets
			technical implementation and domain.	
SR-14	SCOOP4C	The unique personal identification code provides an opportunity to merge personal data from different registers	Need to **define an Identifier class that represents any identifier issued by any authority**, whether a government agency or not	ADMS
SR-15	RIHA	RIHA stores metadata of Estonian public sector databases, registers and information systems. Assets are available in human- and machine-readable format (XML, OWL), human-readable only format (PDF), and machine-readable only format (CSV, WSDL)	Need a **semantic repository for DE4A that stores DE4A semantic assets** so that such assets would be accessible in human and machine-readable format using different information exchange and knowledge representation languages.	TOOP Semantic Repository
SR 16	SDGR	Article 14 refers to exchange of lawfully issued evidence in an electronic format that allows automatic exchange of information. To allow automated exchange, the evidence should be structured uniformly standardized.	Need to **define a set of attributes (metadata) to enable cross-border** transfer of evidences	CCCEV, eIDAS
SR-17	SDGR and DE4A	Data minimization principle: Only the data or documents that are specifically required for the procedure by the requesting competent authority are transferred	Need to **define a minimum set of attributes** needed by the procedure to be exchanged cross-border (canonical evidence).	SDG [24]

4 Requirements for the DE4A Semantic Assets

The SDGR aims to implement an exchange system that would be able to automatically distribute the required evidence between European Union countries. To achieve the above-mentioned aim by 2023, the Member States and the European Commission need to set-up the Once Only technical system, adopting all the necessary acts towards its inextricable and efficient completion. To exchange data across borders, two major semantic interoperability requirements are to be fulfilled; 1) models for evidence to be exchanged and 2) meta models for evidence exchange through the once only technical system.

The SDGR covers a series of administrative procedures. These procedures constitute a sequence of actions that are required to provide to the users all the necessary evidence that will result in their capability to exercise their rights. In DE4A, the propositions of common data models that contribute to exchange evidence are based on the analysis of the existing data models. More specifically, the exchanged evidence is divided into following domains related to DE4A pilot use cases: Studying, Doing Business and Citizen life event services. The meta models for evidence exchange are based on the requirements of the OOTS. This chapter illustrates how the requirements are elicited in both the scenarios.

4.1   Requirements for the Canonical Evidences

The canonical form of a domestic evidence according to a common data structure and format is referred to as **Canonical Evidence (CE)**. It was required to define a canonical form for each of the evidence types required by the procedures involved in the cross-border evidence exchange that are piloted by the De4a project. However, a canonical evidence is not lawfully issued evidence and the issuing authorities are required to provide both domestic and canonical evidence and they are responsible for the accuracy of the matching between canonical and domestic evidence. In this way, canonical evidence is aimed at the automated processing of cross-border and domestic evidence for legal audits on a regular basis or at the request of interested parties. This approach follows the logic behind the multilingual standard forms set by the Regulation 2016/1191 [25]on public documents.

The canonical form for each evidence type that has been identified during the requirement elicitation is called **Canonical Evidence Type**. These types provide agreements on concepts, meanings, languages and ways to represent the information. There are requirements applicable to all the canonical evidence types, presented in Table 2 below. The requirements which are specific to each canonical evidence type will be described in the subsequent sections. We rely on the MoSCoW [26] prioritization technique (Must Have, Should Have, Could Have, Won't Have) for managing priorities.

Table 2: Overall/generic requirements for all the CEs.

Req#	Requirement Description	FUNC/ NON FUNC	Priority
CE-GEN-01	All CE models must rely on existing data exchange standards (e.g., XML/XSD).	FUNC	M
CE-GEN-02	Reuse of existing evidence models (e.g., SDG) is a prerequisite.	FUNC	M
CE-GEN-03	Reuse of core models (e.g., EDCI, Core Vocabularies, etc.) is required.	FUNC	M
CE-GEN-04	Wherever possible, reuse of existing controlled lists (i.e., codelists) should be adopted.	NON-FUNC	C

Req#	Requirement Description	FUNC/ NON FUNC	Priority
CE-GEN-05	Semantic interoperability between evidence items needs to be established on the basis of agreed formats and standards (e.g., W3C-recommended standards).	FUNC	S
CE-GEN-06	Whenever necessary, multilingual representation of information fields shall be supported.	NON-FUNC	S
CE-GEN-07	The naming of the fields should be in line with the respective namings in the adopted models, in order to maintain uniformity with existing approaches.	NON-FUNC	C
CE-GEN-08	Respective free-text descriptions and documentation shall accompany every field.	NON-FUNC	C

4.1.1 Requirements for Company Information Canonical Evidence

When doing business in a foreign country, a company may be required to provide the information (evidence) about the company such as the company name, address etc. to the service providers (data consumers- DC). The Doing Business Abroad Pilot (DBA) intends to exchange such company data through the once only technical system. Before delivering a service, reliable data on the company should be provided by the data providers. In the Doing business abroad pilot, two of such service provisions are tested, namely, starting a business in another EU Member State (UC1), and doing business in another EU Member State (UC2). The data required by both the UCs are the information about the company. The semantic requirements for Company information evidence were elicited according to the requirement elicitation methodology described in section 2.

As the first step of the semantic requirements elicitation, the deliverables of pilot planning were examined and the semantic requirements with respect to the use cases defined and explained in their planning deliverable D3.1 "*Initial requirements for semantic assets*"[1] were extracted. However, the fine-grained semantic requirements were fundamentally drawn from the needs from the specific member states who are physically conducting the pilots. The primary semantic requirements for the company information evidence are presented in Table 3 below.

Table 3: DBA pilot specific requirements elicited by the pilot descriptions

Req#	Requirement Description	FUNC/ NON FUNC	Priority
DBA_CRE_01	A useful company identifier - The data consumer needs to use a company identifier in its data request to the data provider.	FUNC	S
DBA_CRE_02	The company identifier should be exchanged in a fixed structure. See annex 5 for possible company identifiers / structures to use.	FUNC	S
DBA_CRE_03	A natural person can authenticate to services on behalf of the company.	FUNC	S
DBA_CRE_04	The data model should include mandate relationships between the person and the company.	FUNC	S
DBA_CRE_05	Company branch information should be included.	FUNC	S
DBA_CRE_06	Company registration and end dates as simple data type.	FUNC	C
DBA_CRE_07	Reuse of existing semantic models - TOOP for naming.	NON-FUNC	S

Req#	Requirement Description	FUNC/ NON FUNC	Priority
DBA_CRE_08	Code lists for company type (NACE) [27]and company status (BRIS) [28].	FUNC	C
DBA_CRE_09	Use BRIS naming convention to the data model.	FUNC	S
DBA_CRE_10	Identify common attribute lists required by all MS and other optional attributes different MS can provide.	NON-FUNC	S
DBA_CRE_11	Standard format (plain text) needed for company (postal, visitor's) address subcomponents.	NON-FUNC	C

4.1.2 General Requirements for the Moving Abroad Pilot

In the canonical evidence types for the Moving Abroad (MA) pilot, we reuse the approach behind the multilingual standard forms set by the Regulation 2016/1191 [25] on public documents and ongoing work from SDG OOP. Therefore, we identified some general requirements that take place in all MA pilot canonical evidence types and are represented in the table below:

Table 4: General requirements for Moving Abroad pilot

Req#	Requirement Description	FUNC/ NON FUNC	Priority
MA-GE-01	The following information fields are needed for specifying birth evidence: ▶ identifier: according to the DE4A policy for identifiers, like, e.g., urn:eu-de4a:CanonicalEvidenceType::BirthEvidence ▶ holder of the evidence: Person ▶ date of issue: the issuing date of the certificate. ▶ place of issue: the place of issue (location) of the certificate ▶ issuing authority: a Public Organisation with official authority in charge of issuing the certificate.	FUNC	M
MA-GE-02	The following information fields are needed for specifying the person entities: ▶ Identifier. ▶ given name. ▶ family name. ▶ gender ▶ place of birth ▶ date of birth ▶ citizenship	FUNC	M
MA-GE-03	A person is identified with a national ID. Not all Member States can provide this information. A person can be identified in various ways like NationalID, eIDAS IDs, etc. However, in some cases, such information is not provided.	NON-FUNC	S
MA-GE-04	Each location is represented as an address that includes: ▶ The region of the address ▶ Country. ▶ The Post Office Box number. ▶ Name of a passage or way through from one location to another. ▶ Building number	FUNC	M

Req#	Requirement Description	FUNC/ NON FUNC	Priority
	‣ Apartment number in the building ‣ City name or village. ‣ Post name. ‣ Post Code.		
MA-GE-05	Each country has the following administrative levels: National (NUTS0), the regional levels (NUTS1-3) that correspond to public administrations and the local (LAU).	NON-FUNC	M
MA-GE-06	The minimum administrative levels required in a canonical evidence are: Administrative level 1 (country), Administrative level 2 (usually a county, state).	NON-FUNC	S
MA-GE-07	The following fields assume values from controlled lists: ‣ country: according to ISO3166-1 a2. ‣ gender: according to the Human Sex vocabulary from the EU Vocabularies, the possible values are female, male, not applicable, not known, not stated. ‣ NUTS/LAU.	NON-FUNC	S
MA-GE-08	The date of birth can be expressed in days, months, years based on ISO 8601. Some Member States cannot provide all the information (e.g., only the year).	NON-FUNC	C
MA-GE-09	There are reusable initiatives for the development of moving abroad canonical evidence data models e.g., XSD files from ISA for Public Documents and ongoing efforts from SDG OOP.	NON-FUNC	S
MA-GE-10	There are cases where the data owner cannot provide both first name and given name.	NON-FUNC	C

4.1.3 Birth Canonical Evidence

This type of evidence proves the birth of a child. It is related to DE4A Moving Abroad Pilot (UC2) and can be used in many public services as shown in Table 5 below.

Table 5: Requirements for Birth Canonical Evidence

Req#	Requirement Description	FUNC/ NON FUNC	Priority
MA-BE-01	The following information fields are needed for specifying birth evidence: ‣ Birth event: The event indicating the moment a Child emerges from the body of another Person, i.e., start of life. ‣ Child: The Person who is born at the Birth and is of any age, who is a son or daughter. ‣ Parent: One of the two Persons who are jointly the cause of the Child's Birth, i.e., natural parent. ‣ Scope: proves that the person was born and date and place of birth of the person	FUNC	M
MA-BE-02	A child normally has two parents. There are cases with zero, one or "unknown" parent.	NON-FUNC	C

4.1.4   Marriage Canonical Evidence

This type of evidence proves the marriage of two persons. It is related to the DE4A Moving Abroad Pilot (MA) (UC2) and can be used in many public services as shown in Table 6 below.

Table 6: Requirements for Marriage Canonical Evidence

Req#	Requirement Description	FUNC/ NON FUNC	Priority
MA-ME-01	The following information fields are needed for specifying marriage evidence: ▶ Marriage: A legally accepted relationship between two Persons in which they live together. ▶ End of marriage: In case of divorce, this property includes information on the end of marriage	FUNC	M
MA-ME-02	The following information fields are needed for specifying a marriage event: ▶ Date of marriage: The date on which the Marriage took place. ▶ Place of marriage: The Location where the Marriage took place. ▶ Married person: A Person who has entered a Marriage	FUNC	M
MA-ME-03	The following information fields are needed for specifying the married person: ▶ Family Name after Marriage ▶ Family Name before Marriage ▶ Marital Status before Marriage	FUNC	M
MA-ME-04	The following information are needed for specifying the end of a marriage: ▶ date of end of marriage ▶ cause	FUNC	M
MA-ME-05	The following fields assume values from controlled lists: ▶ marital status: cohabitation, divorced person, married person, polygamy, separated person, single parent, unmarried person, widowed person (according to the Marital status vocabulary from the EU Vocabularies [31]) ▶ end of marriage: under discussion (possible values: divorce, void, death of a spouse, non-existence)	NON-FUNC	S

4.1.5   Domicile Registration Canonical Evidence

This type of evidence proves that an individual has successfully completed the procedural obligations to change address of domicile to another Member State of the EU. We describe the final set of requirements (functional and non-functional) for this type of evidence related to DE4A MA pilot UC1 as shown in Table 7 below:

Table 7: Requirements for Domicile Registration Evidence

Req#	Requirement Description	FUNC/ NON FUNC	Priority
MA-DRE-01	The following information about the relevant concepts which are required to describe domicile registration evidence: ▸ Domicile Registration event: It indicates that an individual has successfully completed the domicile registration, in terms of change of address. ▸ Inhabitant: Person, living in the residence or domicile ▸ Domicile: state/country that a person treats as their permanent home, or lives in and has a substantial connection with. It is independent of a person's nationality, and it may change from time to time. ▸ Scope: It proves that a person is living in a state/country at a given address and date.	FUNC	M
MA-DRE-02	Concentrate on re-using the naming convention of the existing semantic models/documents, like Domicile and/or Residence form from Annex X of EU Regulation for Public Documents [32], for the names of concepts for the domicile registration evidence.	NON-FUNC	S
MA-DRE-03	There are cases where the data owner cannot provide both first name and given name.	NON-FUNC	C

4.1.6 Pension & Deregistration Canonical Evidence

The requirement elicitation for both pension and deregistration canonical evidences are an ongoing process based on the second iteration and can be found in DE4A WIKI page [29]. The corresponding XSD files can be found on the GitHub Repository [30]. The related pages will be updated during the course of the project.

As of February 2022, in Table 8 below, we describe the requirements for the canonical evidence types for the use cases "UC3 - Request Pension Information - Claim Pension" and "UC1 - Request Address Change", respectively.

4.1.6.1 Means of Living Evidence

The means of living evidence consists of three different canonical evidence types regarding pension, unemployment and working life benefits. Table 8 presents the requirements of the three canonical evidence types combined.

Table 8: Requirements for Means of living Evidence

Req#	Requirement Description	FUNC/ NON FUNC	Priority
MA-ML-01	The following general concepts are needed for specifying the MeansOfLiving evidence: ▸ Data subject: The person who is subject to this information. ▸ Pension list: Contains information on the pensions of the person for whom the query was made. ▸ Unemployment data: Current unemployment data of the citizen ▸ Working life: current and/or future Working life of the citizen	FUNC	M

Req#	Requirement Description	FUNC/ NON FUNC	Priority
MA-ML-02	The following information fields are needed for specifying the pension: ▶ Pension category: A list of possible values for the pension category. ▶ Status: Status of the benefit (e.g., active, non-active, etc) ▶ Gross amount: It corresponds to the amount of the benefit, including the amounts of possible deductions plus the amount of Personal Income Tax (if applicable). ▶ Net amount: It corresponds to the amount of the benefit after the possible deductions have been applied, plus the amount of Personal Income Tax (if applicable). ▶ Number payments year: It refers to the number of payments corresponding to that benefit in the period of one year. Unless there is an error, this field will always be returned within the benefits of the Registry of Public Social Benefits. ▶ Period of time: It includes the effect date of the benefit and the date on which the benefit ends	FUNC	M
MA-ML-03	The pension category assumes a value from a controlled list	NON-FUNC	C
MA-ML-05	The type of currency is specified for all the fields of amounts.	FUNC	S
MA-ML-06	The type of currency assumes a value from a controlled list.	NON-FUNC	C
MA-ML-07	The following information fields are needed for specifying the unemployment data: ▶ Status: Status of the benefit (e.g., active, non-active, etc). ▶ Period of Time: Effective start date and expected end date or effective end date in case of suspension/cancellation/finish.	FUNC	M
MA-ML-09	The status assumes a value from a controlled list	NON-FUNC	C
MA-ML-10	The following information fields are needed for specifying the working life: ▶ Situation list: Contains the list of situations for the consulted citizen.	FUNC	M
MA-ML-11	The following information fields are needed for specifying the situation: ▶ Social security number: Social security number of the situation returned. ▶ Status: Status of the benefit (e.g., active, non-active, etc). ▶ Period of time: Period of work activity. It includes the date of start of work activity and the date from which the registration takes effect ▶ Contract type: Identifier of the employment contract.	FUNC	M
MA-ML-12	The contract type assumes a value from a controlled list, respectively.	NON-FUNC	C

4.1.6.2 Domicile Deregistration Canonical Evidence

This type of evidence proves that an individual has successfully completed the deregistration of his/her domicile, for the use case UC1 - Request change of address. We describe the final set of requirements (functional and non-functional) for this type of evidence related to DE4A MA pilot UC1 as shown in Table 9 below:

Table 9: Requirements for Deregistration Canonical Evidence

Req#	Requirement Description	FUNC/ NON FUNC	Priority
MA-DE-01	The following information fields are needed for specifying marriage evidence: ‣ Inhabitant: Person, living in a residence or domicile ‣ Domicile: Current domicile inhabited by person	FUNC	M
MA-DE-02	The current domicile is represented as an address.	NON-FUNC	C

4.1.7 Higher Education Diploma Evidence

Higher education diploma evidence along with the following types a) evidence for completion of secondary education and b) non-academic information evidence address the Studying Abroad Pilot (SA) use cases. Deliverable D4.1 *"Studying abroad - Use cases Definition and Requirements"* [33], as well as the results from the first pilot iteration and the associated discussions with the pilot partners served as input in this context.

Higher education Diploma Evidence proves that an individual has acquired a higher education diploma and shall encompass all pertinent aspects for submitting an initial application for admission to a public tertiary education institution. Below is the final set of requirements for this type of evidence.

Table 10: Requirements for Higher Education Diploma Evidence

Req#	Requirement Description	FUNC/ NON FUNC	Priority
SA-HE-01	The following information fields are needed for specifying a higher education diploma evidence: ‣ holder of the achievement: the person that has obtained the academic title or degree. ‣ country: the country where the study programme was completed by the student. ‣ institution name: the name of the higher education institution where the student obtained the degree. ‣ main field of study: the field of finished higher education. ‣ study program: the name of a study programme that the student finished at the higher education institution in order to obtain the degree. ‣ degree: academic title or degree obtained by the student and proven by this evidence. ‣ date of issue: the issuing date of the certificate or diploma. ‣ place of issue: the place of issue (location) of the certificate or diploma. ‣ duration of education: the official duration of education needed for obtaining the specific diploma.	FUNC	M

Req#	Requirement Description	FUNC/ NON FUNC	Priority
	▶ mode of study: specifies whether the specific diploma was acquired via full-time, part-time studying, etc. ▶ scope: refers to the official workload of the study programme in the ECTS (European Credit Transfer and Accumulation System) credit points. ▶ average grade: average grade awarded for the diploma.		
SA-HE-02	The following information fields are needed for specifying the holder of the achievement (in parenthesis: Mandatory/Optional): ▶ first name (M). ▶ family name (M). ▶ gender (M). ▶ date of birth (M). ▶ place of birth (O). ▶ country of birth (M). ▶ citizenship (M). ▶ address (M).	FUNC	M
SA-HE-03	The following fields assume values from controlled lists: country, main field of study, mode of study, scope.	NON-FUNC	C
SA-HE-04	The following fields support multilinguality: institution name, study program name, degree.	NON-FUNC	C
SA-HE-05	The duration of education can be expressed in days, months, and years.	NON-FUNC	C
SA-HE-06	The average grade is accompanied by optional information specifying the grading scheme.	NON-FUNC	C

4.1.8   Secondary Education Completion Evidence

This type of evidence proves that an individual has completed their secondary education. The following below contains the final set of requirements elicited by the date of this deliverable (February, 2022).

Table 11: Requirements for Secondary Education Evidence

Req#	Requirement Description	FUNC/ NON FUNC	Priority
SA-SE-01	The following information fields are needed for specifying a secondary education completion evidence: ▶ country: the country where the secondary education was completed by the student. ▶ degree: secondary education degree obtained by the student and proven by this evidence. ▶ name of school: the name of the secondary school that the person finished. ▶ name of program: the name of the secondary school programme that a person successfully finished (e.g., General secondary programme). ▶ grade: a mark indicating a degree of accomplishment. ▶ issuing date: the issuing date of the accomplishment.	FUNC	M
SA-SE-02	The country field assumes values from a controlled list.	NON-FUNC	C

Req#	Requirement Description	FUNC/ NON FUNC	Priority
SA-SE-03	The following fields support multilinguality: degree, name of school, name of program.	NON-FUNC	C
SA-SE-04	The grade is accompanied by optional information specifying the grading scheme.	NON-FUNC	C

4.1.9   Non-Academic Information Evidence

This type of evidence provides additional non-academic information about an individual, i.e., household composition and income (e.g., for the purposes of awarding a scholarship or grant). The model proposed here is adopted from SDG, but within DE4A there is still an ongoing discussion for revisions to the model, as data evidence providing partners (SGAD) cannot provide evidence on household income and household composition. The final model will be made available at the respective DE4A repository on GitHub [30].

This type of evidence provides additional non-academic information about an individual. Below is the final set of requirements for this type of evidence elicited by the date of this deliverable (February, 2022) as shown in Table 12.

Table 12: Requirements for Non-Academic Information Evidence

Req#	Requirement Description	FUNC/ NON FUNC	Priority
SA-NA-01	The following information fields are needed for specifying the non-academic information evidence: ▸ household income: financial information regarding the income of the household. ▸ household composition: information regarding the composition of the household.	FUNC	M
SA-NA-02	The following information fields are needed for specifying the household income: ▸ yearly gross income: amount of yearly gross income of a household for that tax year. ▸ taxed income: amount of income taxed in that tax year. ▸ yearly net income: amount of yearly net income of a household for that tax year.	FUNC	M
SA-NA-03	The type of currency is specified for all the fields of the household income.	FUNC	S
SA-NA-04	The type of currency assumes a value from a controlled list.	NON-FUNC	C
SA-NA-05	The following information fields are needed for specifying the household composition: ▸ permanent address: the address of permanent residence of the citizen. ▸ other residents: the (non-negative) number of people living together with the applicant.	FUNC	M

4.2   Requirements for Evidence Exchange

For the purposes of evidence exchange, DE4A introduces the **Information Desk (IDK)** as a collection of central components that provide the necessary information to Data Consumers (DC) and Data Providers (DP) for performing the evidence exchange (c.f. D3.3 *"Semantic Framework - Initial version" [*34]*,* D3.5 *"Semantic Toolkit – Initial Version"*[35]). The core requirements for the IDK (as specified in the deliverable D2.1 *"Architecture Framework"* [12]) are the following:

1. Informs the DC what evidence types can be obtained.
2. Informs the DC where the evidence can be obtained.
3. Informs the DC about the data needed for building the request message.
4. Informs the DP whether the request is allowed.
5. Provide information that allows locating the routing information of the participants.

Regarding the needs of the DE4A pilot use cases, the IDK should contain the following information:

‣ What evidence types are available.
‣ What cross-border evidence services are available per evidence type and Member State.
‣ What issuing authority is supplying the evidence data per evidence service.
‣ What routing information is needed for invoking a particular evidence service.

4.2.1   Information Desk Key Concepts

This subsection introduces the IDK key concepts: Canonical Evidence Type (Table 13), Evidence Provision (Table 14), Data Element (Table 15), Data Owner (Table 16), Canonical Event Catalogue Type (Table 17), Subscription Provision (Table 18).

4.2.1.1   Canonical Evidence Type

Table 13: Canonical Evidence Type, Definition & Specifications

Concept	Canonical Evidence Type
Definition	A canonical evidence type is an agreed dataset with a common data model that is an application profile of the corresponding domain ontology or vocabulary.
Specifications	According to the DE4A policy for identifiers, a canonical evidence type is identified with a URN, which includes the token name of such canonical evidence.
	A canonical evidence type is composed by a set of data elements according to its common data model.
	The information provided by a canonical evidence type offers the same proof regardless of the data owner that issues the data; otherwise, the dataset should be reorganised in different canonical evidence types (e.g., Diploma vs Course Results).
	The data provided according to the common data model defined for a canonical evidence type is called canonical evidence. The issuing authority must guarantee the equivalence of the proof provided by a canonical evidence and by the associated domestic evidence – evidence originally issued by competent authorities with legal value.

4.2.1.2   Evidence Provision

Table 14: Evidence Provision, Definition & Specifications

Concept	Evidence Provision
Definition	Availability of the provision of certain canonical evidence types by certain data owners.
Specifications	An evidence provision involves one and only one canonical evidence type.
	An evidence provision includes one and only one Data Owner (DO).
	An evidence provision is identified with a URN combining the identifier of the corresponding canonical evidence type and data owner, according to the DE4A policy for identifiers.
	If a canonical evidence type is provided by several data owners within a country, all of them belong to the same administrative territorial level but not to the same Administrative Territorial Unit (ATU).
	The evidence provision uses either the Intermediation Pattern (IP Provision) or the User Supported Intermediation Pattern (USIP Provision); in the first case, the evidence provision can be either a Direct Provision or a Proxy Provision.
	In the USIP provision, the URLs for redirecting the user forward and backward are required for exchanging a specific message. The request has the URL to send the user back to the DE, and the response from the DO has the URL to redirect the user to the DO.
	An IP provision may include additional information required by the data owner to properly locate the requested evidence, i.e., input parameters for record matching. This additional information has to be required to the user at the Data Evaluator (DE) portal.
	A Proxy Provision could involve more than an Administrative Territorial Unit (ATU) in the same level. The available ATUs to redirect the request to have to be defined in the data service as additional parameters.
	An IP Provision can specify alternative input parameter sets to properly locate the evidence (i.e., record matching), each of them identified by a sequential number and, alternatively, by a short descriptive title.
	An input parameter set is composed of a set of data elements.
	An evidence provision is implemented by at least one service endpoint created by a data transferor within the same country of the corresponding data owner. Service endpoints and data transferors are managed by SMP routing components, and they are out of the scope of this document.

4.2.1.3   Data Element

Table 15: Data Element, Definition & Specifications

Concept	Data Element
Definition	A data element is a piece of information involved in the evidence provision, either in the evidence request as additional information (input parameters) or in the evidence response as part of a canonical evidence

Concept	Data Element
Specifications	A data element is identified by a URI path that represents the hierarchical relationship of the corresponding term within the ontology or general-purpose vocabulary.
	A data element can be an entry of a list, a complex concept, a simple concept. These concepts determine the type of the data element. A complex concept is defined by a set of data elements, each of them either a complex or a simple concept. If the complex concept is a code list, it is defined by a set of list entry concepts.
	A data element is defined by its URI path, type and cardinality, along with a label and optionally a description and an example of the term in different languages, at least in English.
	The translation of the data element label/description to a specific language can be tagged as verified or not verified. The verification of such translation means that it can be considered a trusted translation. Automatic translations are tagged as not verified.

4.2.1.4   Data Owner

Table 16: Data Owner, Definition & Specifications

Concept	Data Owner
Definition	Public organisation, identified by an ISO6523 identifier, that is responsible for the provision of canonical evidences and events.
Specifications	The Data Owner is also the evidence issuing authority.
	A Proxy provision has a Data Owner (DO) that is an organisation running an intermediation platform that hides the organisational complexity behind it. This platform provides proxy services that are either redirect services or choreography services that connect with services provided by issuing authorities.
	If needed, to properly locate the issuing authority (e.g., a regional public administration), a Proxy Provision can require the ATU at a given administrative level as an input parameter.

4.2.1.5   Canonical Event Catalog Type

Table 17: Canonical Event Catalog Type, Definition & Specifications

Concept	Canonical Event Catalog Type
Definition	A canonical event catalog type is an agreed set of event types on the information managed by a Data Owner.
Specifications	According to the DE4A policy for identifiers, a canonical event catalog type is identified with a URN, which includes the token name of such canonical catalog type.
	Each event type in a canonical event catalog is uniquely identified by a token name.
	A DE can subscribe to a specific canonical event catalog by a subscription provision offered by the corresponding DO.

4.2.1.6   Subscription Provision

Table 18: Subscription Provision, Definition & Specifications

Concept	Subscription Provision
Definition	A subscription provision is offered to Data Evaluators by the Data Owners that are responsible for canonical event catalogues.
Specifications	According to the DE4A policy for identifiers, a canonical catalog type is identified with a URN, which includes the token name of such canonical event catalog type.
	The subscription provision is according to the Intermediation Pattern.
	A DE can be subscribed to several canonical event catalogues.
	A DO can provide subscription to a canonical event catalog when events of the types included in that catalog are to be notified by them to the subscriptors to the catalog.
	A DO can provide subscription to several canonical event catalogues under their responsibility.

4.2.2   IDK Semantic Model

Based on the above, this subsection presents the final set of functional requirements for the IDK, in the form of **Competency Questions** (**CQs**). CQs were initially introduced in D3.5 "Semantic Toolkit – Initial Version" and constitute natural language queries that express a pattern for a type of question a semantic model should be able to answer. In a MoSCoW context (see also Section 4.1), all CQs are classified as belonging to priority M/S. The Table 19 shows the final set of CQs that drive the design of the IDK. These CQs were extracted from the respective requirements presented in D3.3 *"Semantic Framework - Initial version" [*34]and D3.5 *"*Semantic Toolkit – Initial Version*"*[35], as well as on the architecture, pilots and common technical components needs.

Table 19: Requirements for IDK Semantic Model

CQ#	Competency Question	Sample Responses
CQ1	How can a data evaluator know how to obtain an evidence or to subscribe to an event catalog?	By the provisions specified by data owners in the IDK.
CQ2	How is a provision identified?	By the object type -canonical evidence or event catalog type- provided and the data owner who provides it.
CQ3	What is the object of a provision?	For evidence provisions a CanonicalEvidenceType, and for subscription provisions a CanonicalEventCatalogType.
CQ4	What are the types of canonical evidence?	BirthEvidence, CompanyRegistration, etc.
CQ5	What are the types of event catalogs?	Currently only one, BusinessEvents.
CQ6	What are the events in the event catalog?	- Company ended operations - Company changed legal form - Company merger or takeover - Company moved to another location - Company administration changed - Company registration evidence changed

CQ7	How is a canonical evidence type or a canonical catalog type identified?	With a URN, according to the DE4A policy for identifiers, like, e.g., urn:eu-de4a:CanonicalEvidenceType::HigherEdCertificate.
CQ8	How many canonical evidence types does evidence provision [X] provide?	Only one. E.g., CompanyRegistration.
CQ9	How many canonical event catalog types does subscription provision [X] provide?	Only one. E.g., BusinessEvents.
CQ10	How are Data Owners (DOs) identified?	With a URN, according to the DE4A policy for identifiers, like, e.g., iso6523-actorid-upis::9991:PT990000101
CQ11	How many DOs are responsible for a provision object of type [X]?	Only one at a specific administrative territorial unit (ATU). Thus, if there are two DOs responsible of [X] at ATU[Y] and ATU[Z], and the latter is in a lower administrative level than the former (e.g., a region within a country), only the DO at highest administrative level (ATU[Y]) can be included as responsible of [X]. In this way, if there are more than one DO responsible to provide [X], all of them are in the same administrative territorial level but not in the same administrative territorial unit
CQ12	Who is the DO that provides a provision object type [X]?	E.g., iso6523-actorid-upis::9991:PT990000101.
CQ13	How many types of evidence provisions are there?	Two types: USIP Provision & IP Provision. USIP provision corresponds to the User Supported Intermediation Pattern (USIP). IP provision corresponds to the Intermediation Pattern (IP) and has two sub-types: Proxy Provision and Direct Provision.
CQ14	What are the differences between the types of evidence provisions?	USIP provision involves a two-stage response message, first to get the URLs for redirecting the user to the DO portal and back to the DE portal, and second to send the evidence request; IP provision MAY include additional information required by the service to ask the user for them and include them in the evidence request.
CQ15	What is the administrative territorial unit that delimits the competence scope of a provision?	The ATU of the DO when the evidence provision is of type USIP provision or IP Direct Provision. When the type is Proxy Provision, more than one ATU can be involved and the final ATU source of the evidence will be set by an additional attribute provided by the user.
CQ16	How many administrative levels can have administrative territorial units that provide evidence type [X] in country [Y]?	Only one.
CQ17	How many administrative territorial units are associated with a DO?	Only one that corresponds to the DO competence territorial scope.

CQ18	What is the administrative level of the DOs that issue evidence type [X]?	E.g., NUTS2.
CQ19	How many countries are involved in the provision of evidence service [X]?	Only one.
CQ20	How does the evidence requester know what to do for requesting the user the additional information required by an IP provision service?	Data services use a common format to define additional information attributes, identifying each attribute by a URI. The URI, type, cardinality, and explanation of these attributes will be available in the Multilingual Ontology Repository (MOR), along with their label, description, and examples in several languages.
CQ21	How does the data owner portal know to show the evidence to transfer for the user preview in the corresponding language of the data evaluator portal?	Each evidence type and their elements are defined in the MOR. Each element is defined by a URI, type, cardinality, and explanation, along with a label, description, and examples in several languages. In this way, the user may preview the evidence in any of the available languages.
CQ22	How does MOR provide information about code lists?	An element whose value is restricted to a code list is defined in MOR of a type [X] corresponding to that code list. In this way, the code list is also an element defined in MOR of type "enumeration" and each value of the code list is also an element defined in MOR of type "token" and URI path [X/z], where [z] is the code and the label to show the user is defined in every available language.
CQ23	What are the types that a MOR element can be defined by?	The type of MOR element [X] can be another MOR element [Z] or a simple type defined according to the XML Schema specification (xsd:string; xsd:int, etc.). In the first case, the MOR element [X] may overload any property of the definition of the [Z] or extend it with more sub-elements.
CQ24	How do data owners know if a data evaluator is granted to use their provisions?	By the corresponding authorization registered in the Authorization Controller (AC), if any.
CQ25	What is an authorization registered in the AC?	It is the authorization to use a provision by defining the data owner and the provision object type (canonical event catalogue or evidence type) and the data evaluator and the category of the DE's procedure where the provision is to be used.
CQ26	How is AC authorization configured?	By specifying the provision -the data owner and the identification of the provision object type- along with the data evaluator and the category of procedures authorized to use that provision.
CQ27	What are the possible categories of a procedure?	The ones corresponding to the procedures listed in SDGR Annex II [36], the four types of Directives mentioned in SDGR Article 14(1) and a category for the rest of cases, according to the corresponding controlled code list.

CQ28	How are Data Owners (DOs) and Data Evaluators (DEs) identified in the AC?	With a URN, according to the DE4A policy for identifiers, like, e.g., iso6523-actorid-upis:9991:PT990000101
CQ29	How is a provision object type identified in the AC?	With a URN, according to the DE4A policy for identifiers for both canonical event catalog types and canonical evidence types -e.g., urn:eu-de4a:CanonicalEvidenceType::HigherEdCertificate- along with the object type, i.e., "EVENTCATALOG" or "EVIDENCE".
CQ30	Who can register an authorization in the AC?	Any DO may register an authorization in AC regarding a canonical evidence type provided by them.
CQ31	How can an authorization be registered for groups of procedure categories, data evaluators or provision object types?	If the character "*" is used as the value of any authorization attribute by the data owner, the authorization is granted to any possible value of such an attribute.

4.3   Requirements for the Provision of Public Procedures

The provision of public procedures is addressed via the **Information Exchange Model (IEM)**, the development of which is based on the requirements provided by D3.3 *"Semantic Framework - Initial version"*, D3.5 *"Semantic Toolkit – Initial Version"*, the DE4A architecture, pilots and technical design of common components.

4.3.1   IEM Basic Assumptions

The Table 20 below lists the basic assumptions for the IEM:

Table 20: IEM Basic Assumptions

IEM Basic Assumptions
The DE4A IEM allows the message exchange between data owners and data evaluators for the processing of administrative procedures.
A data owner, along with its data transferor, can automatically generate an IEM response message for the received request message.
The DE4A IEM models common information to include, such as payload of request and response messages, including error responses.
The IEM is abstract enough to allow the exchange of data or documents in any business domain.
The DE4A IEM is based on the agreed canonical forms of evidence types and existing European or international vocabularies or standards.
The IEM satisfies the specific needs of the DE4A architecture, pilots and technical common components.

4.3.2   IEM Basic Requirements

Table 21 below specifies the basic requirements for the IEM, followed by the basic requirements for evidence exchange (Table 22) and basic elements for event subscription (Table 23):

Table 21: Basic Requirements for the IEM

IEM Basic Requirements
The DE must be able to request from the DO information about the user of the procedure, to be used as evidence in the procedure processing.
The DO must be able to unambiguously understand and automatically process the request message.
The DE must be able to unambiguously understand and automatically process the response message.
A response should unambiguously refer to its corresponding request.
The DT and the DR must be able to transmit the messages to each other.
The exact time of the transmission should be specified in each message.
The DE must provide data that identifies the user, and the data subject if they are not the same one.
Competent authorities involved in the exchange must be identified.
The DE must be able to specify the purpose of the request by providing the name of the involved procedure.
If the DO cannot respond to a request message, the reasons must be given.

Table 22: Basic Requirements for Evidence Exchange

Evidence Exchange Basic Requirements
The DO must be able to provide evidence with legal value in a format lawfully issued.
The DE must specify that the user explicitly requested the use of the Once-Only technical system (OOTS) for the retrieval of the evidence or to state that a law prevents such an explicit request.
The DO must be able to provide evidence in the public documents domain according to its multilingual form in a format selected by the authorities.
Transmitted evidence shall be limited to what has been requested according to the agreed canonical evidence types.

Table 23: Basic Requirements for Event Subscription

Event Subscription Basic Requirements
The DE must be able to subscribe to an event catalog provided by a DO about a specific person, under the person's consent or an applicable law.
The DO must be able to notify events to data evaluators that have an active subscription to an event catalog where the notified event belongs to.
The notified event must be related to the person who is the subject of the subscription, but the notification must be able to include other persons affected by the event.
The DE must be able to request an evidence exchange to the DO that sent an event notification on behalf of this notification and according to the applicable rules.
Event subscriptions must be able to have a fixed starting point and ending point.
The DE must be able to change the time limits of an event subscription or cancel it.

4.3.3 IEM Semantic Model

For outlining the scope of information to include in the IEM and following a similar requirements elicitation procedure like the one described in the earlier subsection, a set of **Competency Questions (CQs)** is used, namely, natural language queries that express a pattern for a type of question the IEM should be able to answer.

4.3.3.1 IEM Request Requirements

Table 24 shows the CQs for the request messages in the IEM.

Table 24: Requirements for IEM request

CQ#	Competency Question	Sample Responses
CQ1	What is the format of the request message [R]?	An XML according to the specification whose identifier is specified in the message.
CQ2	How is a request message [R] identified?	By a UUID.
CQ3	When is request [X] sent?	At the specific timestamp.
CQ4	What is the purpose of the request message[R]?	The administrative procedure [X] whose processing requires the requested information.
CQ5	How is the administrative procedure [X] identified?	The request message identifies the related administrative procedure [X] by the classification provided by the Single Digital Gateway.
CQ6	What authority is sending the request message [R]?	The data evaluator [Y], who is responsible for processing the procedure thus also for the information contained in the response of the message.
CQ7	What authority is requested to send the response to the request message [R]?	The data owner [Z], who is responsible for lawfully issuing the requested evidence or for the event catalogue.
CQ8	How is a data evaluator [Y] or a data owner [Z] identified in the message?	With a URN according to the ISO 6523 specification, i.e., iso6523-actorid-upis::9991:PT990000101, and optionally with one or more preferred names.
CQ9	Who is the subject of the requested message [X] about?	The data subject.
CQ10	Is the data subject [X] a natural or a legal person?	Either type of person.
CQ11	How is the data subject [X] identified?	With the mandatory dataset established by the eIDAS regulation.
CQ12	Is the data subject the user of the procedure?	Yes, except if the user is a natural person representing the data subject (e.g., father requesting his child's birth certificate).
CQ13	How is the data subject's representative [X] identified?	With the mandatory dataset established for natural persons by the eIDAS regulation, along with any additional parameter required by the data service.
CQ14	What elements can be requested in a request message [Q]?	Either a canonical evidence type or a subscription to a canonical event catalog type or the URLs for the user redirection according to a USIP evidence provision.

CQ15	How many elements can be requested in a single request message?	One or more of any type.
CQ16	How many request element types are there?	Three types of elements, for requesting an evidence, the subscription to an event catalog or to get the DO's portal URL to redirect the user to when using an USIP provision.
CQ17	How does a request message differentiate the request elements included?	By a sequential number or a UUID.
CQ18	How does a request message differentiate the different types of elements required?	Each element type has a particular name and structure.
CQ19	What information is required to send for requesting a redirect URL for a USIP evidence provision?	The data evaluator's portal URL to redirect the user back.
CQ20	What information is required to send for requesting a subscription to an event catalog?	The data subject of the events, the canonical event catalog URN and, optionally, the start and/or end date and time of the subscription.
CQ21	How does the request message [R] identify a canonical event catalog type [C]?	By the corresponding to the canonical event catalog type ID, a URN according to the DE4A policy for identifiers.
CQ22	What information is required to send for requesting canonical evidence?	The data subject of the evidence, the canonical evidence type URN, the request grounds and, if required and when the request is using an IP provision, the values for the additional parameters set by the corresponding IP provision.
CQ23	How does the request message [R] identify a canonical evidence type [T]?	By the corresponding to the canonical evidence type ID, a URN according to the DE4A policy for identifiers, like, e.g., urn:eu-de4a:CanonicalEvidenceType::HigherEdCertificate:1.0.

4.3.3.2   IEM Response Requirements

Table 25 shows the CQs that drive the design of the model for representing the response messages in the IEM.

Table 25: Requirements for IEM Response

CQ#	Competency Question	Sample Responses
CQ1	What is the format of the response message [R]?	The corresponding to the specification identified in the message.
CQ2	What request message is the response message [R] responding to?	The request message identified by the Request ID.
CQ3	When is response [X] sent?	At the specific timestamp.

CQ4	What authority is the addressee of the response message [R]?	The data evaluator [Y] who sent the corresponding request message.
CQ5	What authority is sending the response message [R]?	The data owner [Z], who was addressed in the corresponding request message.
CQ6	How is a data evaluator [Y] or a data owner [Z] identified in the message?	With a URN according to the ISO 6523 specification, i.e., iso6523-actorid-upis::9991:PT990000101, and optionally with one or more preferred names.
CQ7	Who is the person subject of the response message [R]?	The data subject that was identified in the corresponding request message.
CQ8	What is the response provided in the response message [R]?	For each of the request elements included in the corresponding request message, the response message includes either an error or the corresponding response, i.e., the requested canonical evidence, the requested URL to redirect the user for the USIP provision or the confirmation of the requested subscription to a canonical event catalog.
CQ9	What error information is provided in the response message [R]?	The code, the message and, optionally, any other relevant additional information of the error (e.g., 1403, "Evidence not found", "No record found for the given and family name of the data subject").
CQ10	What information is provided in response message [R] for a request element of a redirect URL in an USIP provision?	The URL of the provider's portal for redirecting the user to and the canonical evidence type that was requested to be provided.
CQ11	What information is provided in response message [R] for a request element of a subscription to a canonical event catalog?	The parameters of the registered subscription, i.e., the corresponding data subject, canonical event catalog URI and subscription period.
CQ12	What information is provided in response message [R] for a request element of a canonical evidence?	The data subject of the evidence, the requested evidence in its canonical format and, optionally, if the canonical form has no legal value, in its domestic original format and/or its multilingual format with legal value.
CQ13	How is a canonical format of an evidence [X] represented?	With a canonical evidence type URN, according to the DE4A policy for identifiers, like, e.g., urn:eu-de4a:identifiers:CanonicalEvidenceType::HigherEdCertificate:1.0, along with the data evidence in according to the agreed XML Schema for such canonical evidence type.
CQ14	How is an original domestic format of an evidence [X] represented?	With any format represented in base64 encoding according to the specified mime-type (e.g., application/pdf), data languages (iso 639-2 code, e.g., "es"), the issuing type (either original or multilingual form) and, optionally, any other additional information (e.g., "scanned original page of the birth civil registry volume"). If the issuing type is multilingual, the data

		languages are two different languages. The base64 encoding is attached to the IEM Request message.

4.3.3.3 IEM Notification Requirements

Table 26 shows the CQs for the notification messages in the IEM.

Table 26: Requirements for IEM Notification

CQ#	Competency Question	Sample Responses
CQ1	What is the format of the notification message [M]?	The corresponding to the specification identified in the message.
CQ2	How is a notification message [M] identified?	By a UUID.
CQ3	When is notification message [M] sent?	At the specific timestamp.
CQ4	What is notified with a notification message [M]?	That some events have occurred.
CQ5	How many events can be notified in a notification message [N]?	One or more.
CQ6	Who is the sender of the notification message [M]?	The DO responsible for the canonical event catalogs where the types of the notified events belong to.
CQ7	Who is the addressee of the notification message [M]?	The DE subscribed to the canonical event catalogs where the types of the notified events belong to.
CQ8	How is each event included in a notification message [M] identified?	By a sequential number or a UUID.
CQ9	What is an event [E] included in the notification message about?	An event related to an active subscription of the addressee of the notification message.
CQ10	What kind of event is a notified event [E]?	The event type [T] specified in an event included in the notification message [M].
CQ11	How is an Event type [T] identified?	With the ID of the event type, which is a token name that uniquely identifies the event type within a canonical event catalog.
CQ12	Which is the canonical event catalog where the event type [T] belongs to?	The canonical event catalog specified in an event included in the notification message [M].
CQ13	How is a canonical event catalog [C] identified in the notification message?	With an URN, according to the DE4A policy for identifiers.
CQ14	When did a notified event [E] happen?	At the time and date specified in the event.
CQ15	Who is the subject of a notified event [E]?	The event subject specified in the event [E] included in the notification message [M], who is the subject of the DE's

		subscription to the canonical event catalog [C] where the type of event [T] belongs to.
CQ16	Can other subjects be affected by a notified event [E]?	Yes. There can be one or more subjects that are also affected by the notified event [E] included in the notification message [M]. For instance, in an event of type "Company Merge", at least there are type companies that are the subjects of the merge.
CQ17	How is an event subject [X] identified?	With the mandatory dataset established by the eIDAS regulation.

5 Conclusions

The results of this deliverable "D3.2 - *Final requirements for the semantic assets*" concentrate on a set of guidelines regarding semantic interoperability needs that provide foundations for the development of the DE4A semantic interoperability framework.

The initial version of this deliverable "D3.1 - *Initial requirements for the semantic assets*" [1] analysed results of existing literature and relevant projects, like TOOP [4], SCOOP4C [5],and initiatives, like SDG, eIDAS, BPMN. The deliverable (a) defined all the concepts for the whole WP3, (b) listed and explained semantic vocabularies, standards and technologies related to the DE4A use cases, (c) proposed an initial set of requirements and guidelines that are finalised in this deliverable, (d) highlighted common issues to achieve semantic interoperability at national or international level, and (e) provided an agile co-creation methodology that brought together users (DE4A pilots) and designers (WP3) to successfully define a DE4A semantic interoperability framework.

The implementation process in the present deliverable (D3.2) follows an agile methodology that started at a baseline level with D3.1 "*Initial requirements for the semantic assets*". Following the methodology, we iteratively improved the requirements by adding the tools resulting from pilots and other emerging assets. This document (D3.2) acts, therefore, as the final definition of all semantic assets used by WP3. The latter implemented the "reuse before adapt before develop" principle to all software development activities. The initial co-creation methodology is fully elaborated and aligned with the broader projection so it can be used by other similar initiatives.

The main aspect of the deliverable are the final requirements for the DE4A semantic assets, needed by the Information Desk (IDK), Information Exchange Model (IEM) and pilots. The deliverable provides uniquely identified requirement descriptions and competency questions with sample responses. These are provided for the pilot evidences: (a) Doing Business Abroad pilot: company information evidence; (b) Moving Abroad pilot: birth evidence, marriage evidence, domicile registration evidence; and (c) Studying Abroad pilot: higher education diploma evidence, secondary education completion evidence, and non-academic information evidence. The IDK as an information system similarly defines the requirements. Apart from IDK, the requirements are analogously prepared for the IEM that defines processes and other metadata for information exchange.

This deliverable also presents an overview of other common issues in semantic assets and findings on semantic data models for evidence as per the scope of DE4A. The WP3 team covered the aforesaid aspects through an agile methodology and iterative way. The results presented in this document (D3.2) provide a basis for selecting the semantic assets to be used in the DE4A Semantic Framework (D3.4 "Semantic Framework- Final Version") and the Semantic Toolkit (D3.6 "Semantic Toolkit- Final Version").

References

[1] DE4A Project Deliverable D3.1: Initial Requirements for Semantic Assets, [online], available at: https://www.de4a.eu/_files/ugd/f739c2_1d776fe0eb48414592594f0ba95e48fe.pdf

[2] Into a new level of cooperation: From ISA² to building Interoperable Europe - ISA² - European Commission. (2022). Retrieved 10 February 2022, from https://ec.europa.eu/isa2/news/new-level-cooperation-isa%C2%B2-building-interoperable-europe_en

[3] The European Interoperability Framework in detail. (2022). Retrieved 10 February 2022, from https://joinup.ec.europa.eu/collection/nifo-national-interoperability-framework-observatory/european-interoperability-framework-detail

[4] The Once-Only Principle Project | TOOP.EU. (2022). Retrieved 10 February 2022, from https://toop.eu/info

[5] Policy Recommendations | SCOOP4C. (2022). Retrieved 10 February 2022, from https://scoop4c.eu/node/729?field_recommendation_area_target_id=All&sort_by=field_recommendation_area_target_id&sort_order=ASC

[6] Regulation (EU) 2018/1724 of the European Parliament and of the Council of 2 October 2018 establishing a single digital gateway to provide access to information, to procedures and to assistance and problem-solving services and amending Regulation (EU) No 1024/2012, OJ L 295, 21.11.2018, p. 1–38

[7] DE4A Project Deliverable D3.5: DE4A Semantic Toolkit- Initial Version, [online], available at: https://www.de4a.eu/_files/ugd/2844e6_c268b6230223479eb5e088eed495ed1a.pdf

[8] Rodolphe Arthaud, an "Agile" lifecycle for requirements, Web link: https://re-magazine.ireb.org/articles/an-agile-lifecycle-for-requirements , Accessed on 25 Jan 2022

[9] Interoperable Europe, Web link: What is User Story? (visual-paradigm.com), Accessed on 8th Feb, 2022.

[10] EU Co-Inform Project , Co-creation Framework – building a sustainable ecosystem, Web link : https://coinform.eu/wp-content/uploads/2019/02/D1.1-H2020-Co-Inform-Co-creation-Framework.pdf, Accessed on 1st Feb, 2022

[11] DE4A Project Deliverable D7.1: Overview of legal and ethical requirements, [offline]. Distribution level Confidential.

[12] DE4A Project Deliverable D2.1: Architecture Framework, [online], available at: https://www.de4a.eu/_files/ugd/2844e6_ec9fb88703ef431c8a55c63b294fb2cd.pdf

[13] OP Core metadata - EU Vocabularies - Publications Office of the EU. (2022). Retrieved 25 February 2022, from https://op.europa.eu/el/web/eu-vocabularies/opcm

[14] Authority tables - EU Vocabularies - Publications Office of the EU. (2022). Retrieved 25 February 2022, from https://op.europa.eu/en/web/eu-vocabularies/authority-tables

[15] TOOP Exchange Data Model - TOOP Solution Architecture v2.1.0 - UPRC Wiki. (2022). Retrieved 25 February 2022, from http://wiki.ds.unipi.gr/display/TOOPSA210CEF/TOOP+Exchange+Data+Model

[16] XML Editor: XMLSpy. (2022). Retrieved 25 February 2022, from https://www.altova.com/xmlspy-xml-editor

[17] VocBench 3 - ISA² - European Commission. (2022). Retrieved 25 February 2022, from https://ec.europa.eu/isa2/isa2conf18/vocbench-3_en

[18] GBM Ontology - CCTF - UPRC Wiki - CCTF - UPRC Wiki. (2022). Retrieved 25 February 2022, from http://wiki.ds.unipi.gr/display/CCTF/GBM+Ontology

[19] Criterion And Evidence Type Rule Base - TOOP Solution Architecture v2.1.0 - UPRC Wiki. (2022). Retrieved 25 February 2022, from http://wiki.ds.unipi.gr/display/TOOPSA210CEF/Criterion+And+Evidence+Type+Rule+Base

[20] Data Services Directory - TOOP Solution Architecture v2.1.0 - UPRC Wiki. (2022). Retrieved 25 February 2022, from http://wiki.ds.unipi.gr/display/TOOPSA210CEF/Data+Services+Directory

[21] Registry of Authorities - TOOP Solution Architecture v2.1.0 - UPRC Wiki. (2022). Retrieved 25 February 2022, from http://wiki.ds.unipi.gr/display/TOOPSA210CEF/Registry+of+Authorities

[22] Semantic Repository - TOOP Solution Architecture v2.1.0 - UPRC Wiki. (2022). Retrieved 25 February 2022, from http://wiki.ds.unipi.gr/display/TOOPSA210CEF/Semantic+Repository

[23] BRegDCAT-AP. (2022). Retrieved 25 February 2022, from https://joinup.ec.europa.eu/collection/access-base-registries/solution/abr-bregdcat-ap/release/200

[24] GitHub - SEMICeu/SDG-sandbox: The SDG Sandbox creates a space for the review of data models produced by WP4 - Data semantics, formats and quality - in the context of the preparatory work for the Single Digital Gateway Regulation. (2022). Retrieved 25 February 2022, from https://github.com/SEMICeu/SDG-sandbox

[25] EUR-Lex - 32016R1191 - EN - EUR-Lex. (2022). Retrieved 25 February 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R1191

[26] MOSCOW Prioritization Technique. Retrieved 10 February 2022, from https://www.agilebusiness.org/page/ProjectFramework_10_MoSCoWPrioritisation

[27] NACE.Europa - RAMON - Classification Detail List. (2022). Retrieved 25 February 2022, from https://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=CL_NACE2&StrLanguageCode=EN&IntPcKey=&StrLayoutCode=HIERARCHIC

[28] BRIS. Retrieved 25 February 2022, from https://ec.europa.eu/isa2/sites/default/files/publications/eu-core-vocabularies-in-action-using-core-business-vocabulary-in-the-eu-v03.pdf

[29] Canonical Evidence - DE4A. (2022). Retrieved 24 February 2022, from https://wiki.de4a.eu/index.php/Canonical_Evidence

[30] GitHub - de4a-wp3/moving_abroad: Repository of canonical evidences of Moving Abroad Pilot. (2022). Retrieved 24 February 2022, from https://github.com/de4a-wp3/moving_abroad

[31] EU Vocabularies - Publications Office of the EU-marital status. Retrieved from: https://op.europa.eu/en/web/eu-vocabularies/concept/-/resource?uri=http://eurovoc.europa.eu/4184

[32] EU-The Regulation on Public Documents (Regulation 2016/1191), weblink: https://e-justice.europa.eu/content_public_documents-551-en.do, Accessed on 1st Feb, 2022

[33] DE4A Project Deliverable D4.1: Studying Abroad- Use Case Definition and Requirements, [online], available at: https://www.de4a.eu/_files/ugd/2844e6_cd8abc5ce50048fa8fc32bb283775bc9.pdf

[34] DE4A Project Deliverable D3.3: Semantic Framework- Initial version, [online], available at: https://www.de4a.eu/_files/ugd/f739c2_bbefa1524d5f4dc9917b049cee1c28f6.pdf

[35] DE4A Project Deliverable D3.5: Semantic Toolkit- Initial version, [online], available at: https://www.de4a.eu/_files/ugd/2844e6_c268b6230223479eb5e088eed495ed1a.pdf

[36] EU REGULATION 2018/1724 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 2 October 2018 about establishing a single digital gateway to provide access to information, to procedures and to assistance and problem-solving services and amending Regulation (EU) No 1024/2012 and SDGR, OJ L 295, 21.11.2018, p. 1–38, Article 3(1), available at : https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2018.295.01.0001.01.ENG&toc=OJ:L:2018:295:TOC#d1e1017-1-1

[37] DE4A Project Deliverable D2.4: Project Start Architecture, [online], available at: https://www.de4a.eu/_files/ugd/f739c2_fa17b6508b4744bf9872b59faa70cca3.pdf

[38] Code lists - Eurostat. (2022). Retrieved 6 February 2022, from https://ec.europa.eu/eurostat/data/metadata/code-lists

[39] Deloitte Study "Study on Data Mapping for the cross-border application of the Once-Only technical system SDG", March 9, 2020

[40] Controlled vocabularies - EU Vocabularies - Publications Office of the EU. (2022). Retrieved 6 February 2022, from https://op.europa.eu/en/web/eu-vocabularies/controlled-vocabularies

[41] Criteria. Merriam-Webster dictionary, available at: https://www.merriam-webster.com/words-at-play/criteria-vs-criterion-singular-plural-grammar#:~:text=Criterion%20is%20defined%20as%20%E2%80%9Ca,vegetarian%20entrees%20on%20the%20menu.%E2%80%9D

[42] Criteria. Cambridge dictionary, available at : https://dictionary.cambridge.org/dictionary/english/criterion

[43] Formal ontology - EU Vocabularies - Publications Office of the EU. (2022). Retrieved 9 February 2022, from https://op.europa.eu/en/web/eu-vocabularies/concept/-/resource?uri=http://publications.europa.eu/resource/authority/asset-classification/c_89b4bdb7

[44] Proof. Cambridge Dictionary, available at: https://dictionary.cambridge.org/dictionary/english/proof

[45] Proof. Merriam-Webster Dictionary, available at: https://www.merriam-webster.com/dictionary/proof

[46] Public Service. Available at: https://eur-lex.europa.eu/summary/glossary/public_service.html

[47] Taxonomy - EU Vocabularies - Publications Office of the EU. (2022). Retrieved 9 February 2022, from https://op.europa.eu/en/web/eu-vocabularies/concept/-/resource?uri=http://publications.europa.eu/resource/authority/dataset-type/TAXONOMY

[48] Ontologies - W3C. (2020). Retrieved 29 September 2020, from https://www.w3.org/standards/semanticweb/ontology.