2.1 SEMIC: PROMOTING SEMANTIC INTEROPERABILITY AMONGST THE EUROPEAN UNION MEMBER STATES (2016.07)

2.1.1 IDENTIFICATION OF THE ACTION

Type of Activity	Common Services
Service in charge	DIGIT.D2
Associated Services	European Commission (DIGIT, EAS, JRC, ESTAT, SG) EU Publications Office European Parliament, European Court of Auditors, Translation Centre for the bodies of the European Union, European Environment Agency (EEA), European Institute for Gender Equality (EIGE)

2.1.2 EXECUTIVE SUMMARY

Within the ISA Programme, the Action has produced specifications, created communities, and supported policy domains, DGs and MSs in improving semantic interoperability.
In the ISA² programme, the Action continues to develop specifications in the form of data standards and reference data in close collaboration with the domain owners and addressing needs from DGs and MSs.
The existing semantic interoperability solutions produced by the Action and other EC initiatives will be further promoted to be used in operational systems. Recent examples include the successful use of the Core Business vocabulary in the DG JUST Business Registers Interconnection System (BRIS), the DCAT-AP specification in the DG CNECT pan-European Open Data Portal (part of the CEF DSIs), the Core Vocabularies for the DG COMP State Aid Notification system, the Core Public Service Vocabulary in Estonia and the Maltese National Data Strategy where the Core Vocabularies are referred as the starting point for developing national core data standards. Moreover, resources will be invested to exploit these core data model in base registries, with the aim to develop the missing common semantic layer to allow exchange of basic data amongst the MSs, contributing to the "once-only" vision.

Furthermore, aiming at addressing the lack of a mature and interoperable tool for the management and publication of controlled vocabularies, the Action, in close collaboration with the Publications Office of the EU, will continue the development around the open-source thesaurus management tool VocBench.

Having experienced this first hand and unsatisfied with the existing tools on the market, the Food and Agriculture Organization of the United Nations (FAO) and the Publications Office of the EU together with the University of Rome (UNITOV) have decided to use their expertise for the further development of the open-source thesaurus management tool VocBench.

VocBench 3.0 offers a ready solution for the management, alignment and publication of controlled vocabularies as Linked Open Data, thus taking away the burdens for public administrations to maintain and publish their controlled vocabularies in an open and interoperable way.

VocBench 3.2 will evolve beyond the core functionalities which have been developed in the context of the first 2016 funding and of the extensions developed in the 2017 action. In particular, an extension of the visualization possibilities to include graph exploration and other advanced navigation modes still in interactive modality (i.e. the possibility to edit inline the selected elements or to take other actions connected to their type) is the core contribution of this action. Other refinements include advanced search functionalities and search customization (providing users with the possibility to define their own customized searches, defining the input elements and benefiting from automatically drawn user interfaces based on the above declared elements). Advanced SPARQL editing support will feature an improved interaction with the resource stored in the managed dataset, providing suggestions which are highly contextualized depending also on the portion of the query being edited. Finally, drag and drop will be implemented, with different semantics, for all different kind of tree visualizations (classes, properties, concepts, collections).

Since version 3.0, VocBench is able to deal with ontologies and RDF data models. In version 3.2, VocBench will be able to search for data models in public dataset repositories and registries and to retrieve and reuse them completely or partially for the creation of own data models

2.1.3 OBJECTIVES

The objective of this action is to promote semantic interoperability amongst the EU Member States (MSs) and the EU Institutions. This is achieved by:

- Supporting alignments and agreements on common definitions and specifications at the semantic layer for the MSs and the EU Institutions.
- Promoting the use of the ISA Core Vocabularies, the ADMS and DCAT-AP specifications.
- Identification of areas for new Core Vocabularies and/or common data models.
- Supporting relevant policy and projects executed by DGs and MSs.
- Using core data models in base registries.
- Supporting DGs and MSs for using the ISA data models in operational systems.
- Supporting and promoting metadata management policies to MSs and DGs.
- Promoting open and linked data approaches and technologies.
- Promoting best practices, and lessons-learnt in the area of semantic interoperability.
- Supporting the development of a new version of VocBench (version 3.2) as a toolbox for controlled vocabularies, metadata, glossaries, ontologies and data models.
- Continuation of the promotion and dissemination work including the SEMIC conference.

2.1.4 SCOPE

Development of new specifications, policy and technical support, pilots, promotion activities are within the scope of the Action. Emphasis will be given on promoting cross-domain interoperability but single-domain, cross-border and/or cross-EU institutions initiatives remain with the scope of the Action.

Moreover, the Action will develop a new version (version 3.2) of the collaborative thesaurus-management platform VocBench

The following activities are not considered within the scope of this Action:

- All technical development related to the Joinup platform including the implementation of the federation of interoperability solutions.

2.1.5 ACTION PRIORITY

2.1.5.1 Contribution to the interoperability landscape

The contribution of the action to the interoperability landscape, measured by the importance and necessity of the action to complete the interoperability landscape across the Union

Question	Answer
How does the proposal contribute to improving interoperability among public administrations and with their citizens and businesses across borders or policy sectors in Europe? *In particular, how does it contribute to the implementation of:* • *the new European Interoperability Framework (EIF),* • *the Interoperability Action Plan and/or* • *the Connecting European Facility (CEF) Telecom guidelines* • *any other EU policy/initiative having interoperability requirements?*	The Action supports the implementation of the EIF and the EIS, the INSPIRE Directive, the PSI Directive, the Service Directive by promoting semantic interoperability, through the definition and use of common specifications. Vocbench is an open-source reference tool for the management of semantic interoperability assets, it promotes the opening of data and facilitates the exchange of data between national and EU administrations. It also opens up the possibility for potential reuse by businesses and citizens. The proposal covers the following underlying principles of the EIF: Reusability, Multilingualism, Openness, Semantic interoperability, Technical interoperability and Standardisation.
Does the proposal fulfil an interoperability need for which no other alternative action/solution is available?	The Action tries to develop and promote common specifications when such specifications are not available. It is the only horizontal action for promoting semantic interoperability for the European public administrations. Moreover, VocBench is the most mature OSS platform for advanced thesaurus

	management. In the paper: ”VocBench: A Web Application for Collaborative Development of Multilingual Thesauri”; A. Stellato , S. Rajbhandari, A. Turbati, M. Fiorelli, C. Caracciolo, T. Lorenzetti, J. Keizer, M. T. Pazienza[23]; the most important systems (i.e. renowned, and satisfying more than half of the above requirements) among those listed above have been compared to VocBench 2.

2.1.5.2 Cross-sector

The scope of the action, measured by its horizontal impact, once completed, across the sectors concerned

Question	Answer
*Will the proposal, **once completed** be useful, from the interoperability point of view and utilised in two (2) or more EU policy sectors? Detail your answer for each of the concerned sectors.*	The specifications developed by the Action are generic and reusable in any policy domain. VocBench can be used for hosting any type of thesauri and vocabulary, allowing also mappings and alignment amongst assets coming from different policy areas. Vocbench will facilitate the cross-sector interoperability by use of semantic technologies/formats and by facilitating alignments between controlled vocabularies from different sectors/policy domains. E.g. Alignment of specialised thesauri from the “Agriculture, fisheries and foods” policy domain with specialised thesauri from the “Environment, consumers and health domain with generic thesauri that cover all EU policy domains such as EuroVoc
*For proposals completely or largely **already in***	The DCAT-AP spec is used in the PSI

[23] The paper is available here: http://link.springer.com/chapter/10.1007%2F978-3-319-18818-8_3 or in its self-archived camera-ready version here: http://art.uniroma2.it/publications/docs/2015_ESWC15_VocBench.pdf.

operational phase, *indicate whether and how they have been utilised in two (2) or more EU policy sectors.*	(open data) and in environmental (INSPIRE) policy areas. The Core Public Service Vocabulary has been used in the Service Directive (points of single contacts), eGovernment (national portals) and local government (OSLO initiative). VocBench has been used to create semantic interoperability between datasets in the fields of agriculture, environment, gender equality, education.

2.1.5.3 Cross-border

The geographical reach of the action, measured by the number of Member States and of European public administrations involved.

Question	Answer
Will the proposal, ***once completed,*** *be useful from the interoperability point of view and used by public administrations of three (3) or more EU Members States? Detail your answer for each of the concerned Member State.*	The proposal will facilitate the cross-border interoperability thanks to its inherent support for multilinguality. Further cross-border interoperability improvements can be expected through the alignment of the generic EuroVoc thesaurus that covers the EU policy domains with specialised EU and national controlled vocabularies through VocBench. Potentially all stakeholders listed under 1.1.9.1 and 1.1.9.2 can in such a way be linked using VocBench. VocBench is already used by public administrations in France, Italy, the Netherlands and Scotland and interest has been expressed from public administrations in Belgium and Slovenia.
For proposals completely or largely ***already in operational phase****, indicate whether and how they have been utilised by public administrations of three (3) or more EU Members States.*	VocBench is already used by public administrations in France, Italy, the Netherlands and Scotland. See 1.1.9.2 for more details.

2.1.5.4 Urgency

The urgency of the action, measured by its potential impact, taking into account the lack of other funding sources

Question	Answer
Is your action urgent? Is its implementation foreseen in an EU policy as priority, or in EU legislation?	Semantic interoperability has always been identified as an important barrier in digital EU policies (e.g. Digital Agenda for Europe, Digital Single Market) Implementation of this version of VocBench is not urgent, but as the major revision of VocBench 3.0 concerned in particular under the hood technological changes, the multilingual controlled vocabulary management community and the key stakeholders are eagerly waiting for the inclusion of the requirements identified in the context of the ISA 2015 action.
How does the ISA² scope and financial capacity better fit for the implementation of the proposal as opposed to other identified and currently available sources?	As semantic interoperability should be promoted at a horizontal and cross-domain level, the ISA² is a perfect match for hosting such activities.

2.1.5.5 Reusability of action outputs

Can the results of the action (following this proposal) be re-used by a critical part of their target user base, as identified by the proposal maker? For proposals or their parts already in operational phase: have they been re-used by a critical part of their target user base?

Name of reusable solution	Data models
Description	A set of 6 Core Vocabularies, three DCAT-related specifications, ADMS have been developed by the Action
Reference	https://joinup.ec.europa.eu/community/semic/description
Target release date / Status	2012 – today: All the specifications are released on Joinup.
Critical part of target user base	The specifications can be used by all system architects, designers and national authorities as horizontal data standards.

For solutions already in operational phase - actual reuse level (as compared to the defined critical part)	34 known implementations in 8 countries are presented on https://joinup.ec.europa.eu/community/semic/description

Name of reusable solution	VocBench 3 (revision 2)
Description	The main outcome of the work. A fully fledged collaborative platform for management of SKOS thesauri OWL ontologies and RDF datasets in general, with particular emphasis on Controlled Collaboration, Extensibility and Scalability
Reference	http://vocbench.uniroma2.it/
Target release date / Status	3.0 released to Publications Office on July 2017. Currently under evaluation. Planned 3.1 by July 2018. Planned 3.2 by the end of this project (at most, or intermediate release and further features added by the end).
Critical part of target user base	VocBench 3 is an open source system for collaborative editing of RDF data, OWL ontologies and SKOS and SKOS-XL thesauri. VocBench's adoption is not meant to be in any way restricted to the registered stakeholders, and its full compliance with W3C standards make it a perfect platform for the evolution of many organizations and authorities towards production and publication of Linked Open Data. At the time of writing there are already several users adopting VocBench 2, the previous version of VocBench developed before the ISA^2 action that funded development of VocBench 3. Major users and the thesauri managed through VB2 are described here: http://vocbench.uniroma2.it/support/community.jsf. There is also a mailing list (vocbench-user@googlegroups.com) with more than 100 registered users.
For solutions already in operational phase - actual reuse level (as compared to the defined critical part)	N.A.

2.1.5.6 Level of reuse by the proposal

The re-use by the action (following this proposal) of existing common frameworks and interoperability solutions.

Question	Answer
Does the proposal intend to make use of any ISA[2], ISA or other relevant interoperability solution(s)? Which ones?	Joinup is used as a publishing and community building tool for all the specifications developed under the Action. The specifications produced by the Action have the potential to be used by several other Actions in all data modelling tasks. For VocBench, the current platform already generates metadata descriptions of the controlled vocabulary in DCAT, DCAT-AP and ADMS. The module can be extended to cover other ISA interoperability solutions such as StatDCAT-AP and GeoDCAT-AP and the various DCAT-AP-xx developed by the single member states of the EU. This version will reuse information from the EU Data Visualization Catalogue of the EU Open Data Portal developed thanks to ISA2 funding:(http://52.50.205.146:8890/data_visualisation_catalogue/). A catalogue pointing to different visualization technologies..
For proposals or their parts already in operational phase: has the action reused existing interoperability solutions? If yes, which ones?	Joinup is the main publishing and community platform for the Action.

2.1.5.7 Interlinked

The extent to which the action (following this proposal) contributes to Union's initiatives such as the DSM.

Question	Answer
Does the proposal directly contribute to at least one of the Union's high political priorities such as the DSM? If yes, which ones?	Interoperability is considered critical for the progress of the Digital Single Market, and special references are included in the relevant strategy. Admittedly, semantic interoperability together with organisational are the two layers where most MSs experience the most severe problems. o The Action contributes to "better public services" in the Digital Single Market strategy.

What is the level of contribution?	o Copying from the DSM strategy "*The lack of open and interoperable systems and services and of data portability between services represents another barrier for the cross-border flow of data and the development of new services."* The Action contributes to better interoperability and data portability. o The DSM promotes the "Free flow of data" and identifies interoperability as an important issue for achieving this. o The DSM strategy asks for a revision of the European Interoperability Framework. One of the interoperability levels there refers to "semantic interoperability". o The DSM discusses about the interconnection of base registries and the Once-Only principle. Semantic interoperability is a prerequisite for achieving these targets. PSI and Open Data In line with the EC objectives in the area of the PSI Directive, the Action promotes policies towards both open data and open metadata by the MSs and the EU services as a support measure for the more general Open Data policy. One of the concrete measures developed to implement this policy entails the setting up of the EU Open data Portal which is operated by the Publications Office of the EU. Close collaboration and common activities with the Publications Office of the EU have already started in 2013 and continues to secure alignment of efforts and concrete results. The creation of the pan-European Open Data portal is also part European open data policy. The portal has started its operation in 2015, and the DCAT-AP specification is used as the metadata standard for the description of the datasets coming from over 70 Open Data portals from 34 countries. The Action closely collaborates with the relevant EC service (DG CONNECT/G3). In 2014, a revision of the DCAT-AP has been concluded. Moreover, two extensions of the specifications are being developed to cover description of geospatial (GEO/DCAT-AP) and statistical (STAT/DCAT-AP) data.

2.1.6 PROBLEM STATEMENT

The environment in which data exchange and service execution takes place amongst MSs is very complex contributing many barriers and challenges to the exchange of data during the execution of European Public Services. These barriers include divergent interpretations of the data, lack of commonly agreed and widely used data models, absence of universal reference data (e.g. code lists, taxonomies), from which only few are publicly available and even less as Linked Open Data, lack of interoperable tools/formats for the management and publication of reference data, the multilingual challenge. Last, in the sensitive area of labour markets, we currently experience fragmented European

labour markets which hinders European mobility as the data of jobseekers and the matching job vacancies are kept in different systems due to the existing digital segmentation etc.

Due to these pressures, semantic interoperability becomes an important element in many eGovernment and interoperability national agendas and interesting experience and lessons-learnt can already be shared at a European level.

The problem of	divergent interpretations of the data, lack of commonly agreed and widely used data models, absence of universal reference data (e.g. code lists, taxonomies), from which only few are publicly available and even less as Linked Open Data, lack of interoperable tools/formats for the management and publication of reference data, the multilingual challenge
affects	the exchange of data during the execution of European Public Services
the impact of which is	fragmented European labour markets which hinders European mobility.
a successful solution would be	having semantic interoperability as an essential elements in eGovernments and intareoperability national agendas and sgaring of leassons learnt and best practices

The problem of	Bringing existing vocabularies, thesauri, ontologies and datasets to Linked Open Data (LOD) standards
affects	EU institutions, national administrations and businesses
the impact of which is	Until this target is achieved, data interoperability is hampered
a successful solution would be	Provide system and tools for porting local data to LOD standards, facilitating (possibly collaborative) maintenance and evolution

The problem of	Exploring and visualizing data in editing tools, which are usually focused on optimizing editing
affects	Especially domain experts and naive users, less specialized in data modelling/editing and more focused on domain exploration and analysis
the impact of which is	The “understandability” of the content is affected.
a successful solution would be	Provide additional, alternative, visualization and exploration solutions

2.1.7 IMPACT OF THE ACTION

2.1.7.1 Main impact list

Impact	Why will this impact occur?	By when?	Beneficiaries
(+) Savings in money	Users adopt more tools for maintaining/browsing their data, as none of them satisfy all of user needs. The need to efficiently browse arbitrary data structures in a convenient way often asks for customized solutions that are developed ad-hoc. Adding more customization possibilities, especially for exploration navigation and search make an editing tool closer to suit different, specific needs of each user for exploring their own data.	As soon as the new version is adopted and the users abandon multiple solutions in favour of a centralized management in VocBench	All users
(+) Savings in time	Better, customizable interaction possibilities (both for data exploration, search and navigation) improve productivity and thus reduce working time	As soon as the new VocBench version is adopted	All users
(+) Better interoperability and quality of digital public service	Better data quality thanks to visual exploration of data maintained in VocBench facilitating identification of non-standard patterns	As soon as the new VocBench version is adopted	All users

(-) Integration or usage cost	The usage cost is reduced especially for those users more interested in exploring, analysing, navigating and searching the data	As soon as the new VocBench version is adopted	Final users, non-editors

Beneficiaries	Anticipated benefits
Member States' Public Administrations	o Forum to identify opportunities for harmonization at European level o A virtual place to share best practices and experiences o Core Vocabularies, reference data, XML schemata and thesauri ready to be used in public administration information systems development. o Linked data best practices. o Reduced development costs • due to reuse during the initial development phase • due to less interoperability conflicts while integrating systems or providing cross-agency/domain/country services
IT Services Industry	o Reduced development costs, as explained above. o Reuse and integrate controlled vocabularies and metadata produced by Member States and Public Administrations or EU institutions and bodies.
European Commission Services	o A set of specifications in the form of Core Vocabularies and other relevant tools, such as reference data (e.g. authority tables), XML schemata and thesauri that can be reused by the EC services to facilitate interoperability among different applications and systems o Reduced development costs, as explained above.
For VocBench	
European Commission Library	Reconciliation and mapping of the European Commission thesaurus (ECLAS) with EuroVoc in the purpose of re-indexing the European Commission Library collection (ECLAS) with EuroVoc.
European Commission, Informatics Directorate General (DIGIT)	Use VocBench for the maintenance and dissemination of the glossaries and ontologies produced by DIGIT.
EU and national Open Data Portals	Use VocBench for the maintenance and dissemination of the DCAT-AP themes used as a standard vocabulary to annotate the Open Data datasets in European, trans-European and national Open Data Portals.

FAO (Food and Agricultural Organisation)	Use VocBench for the maintenance and dissemination of the Agrovoc thesaurus, the Glossary of Biotechnology for Food and Agriculture, the Land and Water controlled vocabulary and the FAO main topics. A list of the partners acting in the VocBench Community is available at http://VocBench.uniroma2.it/support/community.jsf
JRC Inspire Registry	Use VocBench for the maintenance and dissemination of the JRC Inspire themes and features concept dictionary.
Supranational, national and regional authorities	Use VocBench for the maintenance of their controlled vocabularies and metadata, for the alignment of their supranational, national and regional controlled vocabularies with EuroVoc. Enrich the EU-level semantics of EuroVoc with in-depth national and regional coverage responding to the need of the Member States.
Academy and Education	Given the open and standard compliant nature of VocBench, it is being adopted in various areas of education (ranging from digital libraries, archiving and purely IT areas of learning) for teaching how to develop thesauri compliant with W3C standards for the web of data.
European Commission Library	Reconciliation and mapping of the European Commission thesaurus (ECLAS) with EuroVoc in the purpose of re-indexing the European Commission Library collection (ECLAS) with EuroVoc.

2.1.7.2 User-centricity

The specific actions listed in this proposal follow (and mostly complete) requirements gathered from:

- a synthesis of requirements and a development plan that followed a stakeholders meeting held in the Publication Office on February 2016, before the start of the first action on VocBench. The outcomes of this synthesis have been collected in deliverable: *D03.03 Report on the provided support for the development of VocBench 3.0*
- desiderata expressed by stakeholders along the first year of the project and during the VocBench day organised end 2016
- desiderata expressed by members of the EuroVoc maintenance committee making use of VocBench (e.g. for the Graph Visualization)
- feedback provided by the Publication Office following the release of the system at the end of the first action 2016/2017 (the system is under evaluation by the OP at the time of writing)

2.1.8 EXPECTED MAJOR OUTPUTS

Output name	Component for Graph Representation of RDF Data and Additional Alternative Data Visualizations
Description	The component will satisfy the following requirements: Lay the basis for additional visualizations (graph and others) taken from existing libraries to RDF management, as a dedicated library The module will enable the graphical visualisation of the

	ontology/vocabulary. The graphical visualisation will be possible for single concepts, concept collections or schemes and, concept search results and SPARQL query results. The ontology/vocabulary graphical visualisation will provide an extension point for exporting the graph into arbitrary graphics formats. An extension for the ontology/vocabulary graphical visualisation, will be implemented for exporting the graph to the SVG graphics format. Connect the additional visualizations with the pervasive addition of Extension Points: role/context-based extension points for pluggable actions on resources (developed in phase 2) Provide additional visualizations besides Graph-View, by checking available graphical frameworks in Javascript
Reference	The component will be developed by reusing existing data visualization technologies for Javascript and developing a dedicated framework for RDF. A state-of-the-art investigation will be conducted, though the current technologies have already been identified as potential candidates: JavaScript InfoVis Toolkit (http://philogb.github.io/jit/) Gephi: The Open Graph Viz Platform (https://gephi.org/) D3: Data-driven documents (https://d3js.org/) The investigation will also reuse information from the EU Data Visualization Catalogue of the EU Open Data Portal developed thanks to ISA^2 funding:(http://52.50.205.146:8890/data_visualisation_catalogue/) a catalogue pointing to different visualization technologies
Target release date / Status	Q2/2019

2.1.9 ORGANISATIONAL APPROACH

2.1.9.1 Expected stakeholders and their representatives

Stakeholders	Representatives
Member States	o ISA^2 Committee o ISA^2 Coordination Group or ISA CG equivalent o Persons/units in charge of National and/or regional

	Interoperability Frameworks and/or metadata standards catalogues/repositories, etc. o Persons/units in charge of governmental portals
Standardization bodies	o W3C o OASIS o GS1 o CEN o UN/CEFACT o Dublin Core Metadata Initiative o EU Interinstitutional Metadata Maintenance Committee (IMMC)
DIGIT	• DIGIT/B2, DIGIT/01 • Unit (s) responsible for promoting common specifications in new and existing systems
EU Publications Office	The unit responsible for metadata management The unit responsible for the EU Open Data portal The unit responsible for the Common Vocabulary for legal documents
Other EC DGs and EU Institutions	Representatives from the DGs who work as IT coordinators (IRMs) Representatives from DGs and EU Institutions (e.g. JRC, EEA etc.) who work in the area of metadata management. DG CNECT units for public services and G.3 PSI and open data The EC Secretariat General The EU Council, the European Parliament.
FAO	The unit responsible for metadata management using VocBench.

2.1.9.2 Identified user groups

In principle any Public Administration from any Member statesas well as EU Institutions are potential user groups.
For VocBench, other than the foreseen stakeholders, the project is already adopted by different entities. UNITOV offers open support through two mailing lists for end-users (vocbench-user@googlegroups.com) and developers (vocbench-developer@googlegroups.com) interested in the developing aspects of the system.
The lists are open and do not foresee any pre-identification of users. However, in two years of service, the following important actors have been identified are regular adopters of VocBench.
For each of them, we report the known managed resources:

- Food and Agriculture Organization (FAO) > AGROVOC, Biotechnology, Land and Water, FAO Topics
- FAO, CABI and NAL (joint effort) → GACS (Global Agriculture Concept Scheme)

- Institut National de la Recherche Agronomique (INRA) > in the context of AnaEE France project (ecosystems and biodiversity thesaurus) and Délégation à l'information scientifique et technique (DIST)
- CABI (Centre for Agriculture and Biosciences International), head office UK > CAB thesaurus
- European Commission Central Library > ECLAS (library catalogue)
- European Environment Agency (EEA) > GEMET
- EU Publications Office > EUROVOC
- European Holocaust Research Infrastructure (EHRI) project: EHRI holocaust thesaurus, supporting the work of the EHRI Editorial Board (deployed by Ontotext)
- Columbia University > IEDA Thesaurus
- Italian Senate > TESEO
- United Nations Environment Programme (UNEP) > InforMEA, in the context of MEA, the Information and Knowledge Management Initiative
- Scottish Government > Gov metadata
- United Nations Convention to Combat Desertification (UNCCD)
- UNESCO > UNESCO Thesaurus (deployed by Sparna)
- Harvard University > Unified Astronomy Thesaurus (UAT)
- Dutch Firebrigade

Another important group is represented by users working in education. We have known reports of professors adopting VocBench in order to:
1) Show how to develop a SKOSXL thesaurus (VocBench is in fact one of the very few systems which are at the same time: completely free, open source, and totally compliant and based on languages of the RDF family of standards).
2) Have students experience a collaborative experience in the development.

2.1.9.3 Communication plan

The Action will create the necessary links with the ISA^2 Communication Action to coordinate the promotion and dissemination of results to several types of recipients. The Action includes within its plan communication activities which are of technical nature, while Action 4.1.1 will promote the results to a broader audience, using appropriate non-technical language, and different channels.

Event	Representatives	Frequency of meetings / absolute dates of meetings?
The Action tries to disseminate its work with presentations, posters and panels in the major eGovernment and semantic technologies conferences and events		

Event	Representatives	Frequency of meetings / absolute dates of meetings?
The Action will try to raise awareness for the importance of metadata management, data standards and alignment in MSs and internally in the EC		
The Action will try to disseminate and promote the use of the Core Vocabularies and linked data best practices in the EU MSs		2 MSs visits per year
The Action will continue organising its annual conference which has been very successful so far		1 conference per year

VocBench is already a mature open-source project. Communication about its further releases, support and discussion are given and held through different channels, covering different needs with the most appropriate solution, and exhibiting redundancy to some extent.
By first, the official site of VocBench: http://vocBench.uniroma2.it represents an entry point for people/organizations willing to approach the system. Secondly, two mailing lists/discussion groups: Users[24] and Developers[25], provide assistance to, respectively, the community of users and of those willing to contribute the project or develop independent extensions for it. An Atlassian Suite instance provides project management, bug report and wiki services. Finally, the Agricultural Information Management Standards (AIMS) portal, managed by FAO, reports on news about VocBench and associated systems and tools.
The Publications Office of the EU will reserve a space on its future corporate EU Vocabularies website for information on VocBench. The Publications Office offers access to an instance of VocBench to other EU institutions and bodies through the Testa network. It will organise information and training sessions for services interested in managing their controlled vocabularies in VocBench.
Promotion beyond the EU institutions of the new version of VocBench is foreseen by making use of social media (Twitter), webinars and participation in conferences (posters). The project stakeholders will promote VocBench in their respective communities.

2.1.9.4 Key Performance indicators

Numbers before the KPI refer to section numbers in the deliverable D3.3 of ISA action 2015 identifying VB requirements

[24] http://groups.google.com/group/VocBench-user
[25] http://groups.google.com/group/VocBench-developer

Description of the KPI	Target to achieve	Expected time for target
Access to remote registries and repositories (not in the D3.3 deliverable, added following an action of the Denmark member state)	Access to remote registries and repositories for the search, retrieval and possibly provisioning of dataset (ontologies, thesauri, datasets in general)	Q1/2019
5.1.5 Mass Edit Actions	Mass Edit Actions for quick editing of large quantities of data with homogeneous characteristics	Q2/2019
5.3.2.1. Graph and Additional Visualizations for data		
Graph and Additional Visualizations part 1	Realization of a visualization framework for RDF with support for graphs and interactive graphical objects	Q1/2019
Graph and Additional Visualizations part 2	Realization of dedicated views for different knowledge structures, divided by resource type (e.g. RDF concepts, classes, properties, collections, schemes…)	Q2/2019
5.1.5.2. Drag and Drop Support	Support for Drag and Drop in all type of trees: rdfs/owl classes, skos concepts, skos collections, rdf properties and their specific owl subproperties with dedicated actions and outcomes depending on the resources and area of the user interface involved in the Drag and Drop	Q1/2019
Support for UNDO		Q1/2019
Action Level History Management	An History Based Time Machine allowing users to browse the evolution of resources with the granularity of single actions (previously it was available at the level of stored dataset versions)	Q2/2019

2.1.9.5 Governance approach

The Action is run by the European Commission, DG DIGIT, D2 (Interoperability Unit).
For the work expected in this Action, collaboration with MSs and other stakeholders (e.g. DG EMPL, the EU Publications Office, other DGs, EU institutions, agencies and bodies, standardisation bodies) is considered of critical importance. For this reason, communication with the MSs and third parties will be frequent and their active involvement will be encouraged via invitations for participation in ad hoc groups discussing specialised issues related to the content of the Action.
For Vocbench a Project Steering Committee will be established to provide overall guidance and to ensure the consistency of the project regarding the roadmap, the technical specifications and the programme actions. The Committee will consist of representatives from the Publications Office, UNITOV, DIGIT (European Commission), Joint Research Centre (Inspire Registry), Food and Agricultural Organisation (FAO), European Environment Agency (EEA) and representatives from the Member States' national or regional authorities.

2.1.10 TECHNICAL APPROACH AND CURRENT STATUS

Overview for work so far

The Action tries to overcome semantic interoperability problems by:

1. documenting what is available in each MS with regards to metadata policies and the management of semantic interoperability assets;
2. identifying opportunities for alignment;
3. promoting share and reuse of semantic assets, experiences and tools and facilitating agreements in key areas;
4. raising awareness on the importance of metadata management;
5. identifying and promoting successful practices and tools for linking governmental data;
6. developing and promoting common specifications.

Through this action, the ISA programme sets to achieve a better collaboration between European public administrations towards creating new and promoting existing agreements on the meaning of the exchanged information and on the common use of metadata. Joinup provides the necessary infrastructure for this collaboration while a broad network of stakeholders including several EU bodies like JRC, DG SANTE, DG EMPL, DG FISMA, DG GROW, the PO, DG TAXUD, DG MARE, DG CONNECT, DG COMP, DG COMM and DG JUST actively participate and/or follow the work.

The Action maintains communication links with relevant initiatives to promote semantic interoperability around the world (e.g. USA, Japan) to exchange views and share experiences.

Below an overview of the results of the Action during the last 5 years is provided.

In 2011, an assessment of the Action results for the period 2008-1010 took place. Based on this evaluation, recommendations were provided to identify areas where the Action could contribute. Following these recommendations, the Action produced the following results:

a) In the period 2011-2012, the Asset Description Metadata Schema (ADMS) was developed as a common vocabulary to describe semantic assets. ADMS has been extended to describe any type of interoperability solution. Based on this specification, a federation of semantic assets repositories has been developed and is available via the Joinup platform. This federation went live in January 2013. Currently, more than 20 repositories participate in the federation making available over 2.000 specifications. The Action will continue providing support to MSs to participate in the federation and use ADMS for describing their own assets. Nevertheless, the EFIR Action (action 4.2.4) remains the focal place for the federation maintenance and support in Joinup.

b) In the period 2011-2013, the Action has developed four Core Vocabularies, namely Core Business, Core Person, Core Location, and Core Public Service in close collaboration with other EU bodies including DG MARKT, Eurojust, and the JRC. These specifications (together with ADMS) have been endorsed by the ISA Coordination Group. In 2013-2014, implementations of the vocabularies have been implemented in 5 pilots with data coming from MSs and different EC/DGs.

c) In 2013, the DCAT Application Profile (DCAT-AP) as a specification to describe open data catalogues and datasets was developed in close collaboration with DG CONNECT and the EU Publications Office.

d) During the period 2011-2014 several case studies, surveys and recommendations have been published trying to raise awareness on issues related to semantic interoperability, open and linked data, metadata management and persistent Unique Resource Identifiers (URIs).

e) In 2014, the Action has tried to promote the use of core vocabularies both at national and at European level. Issues related to sustainability, governance, publication and persistence for these specifications were also identified and work has started to set up relevant policies. During the same year, the Action produced guidelines and support for the publication of high-value datasets, and initiated cross-institutional work to draft metadata and URI policies for the EU institutions which are reusable by the MSs. Work to create a stack of tools for metadata management with the participation of DG MARE, JRC, DG COMP and the PO has been also initiated.

f) In 2015/16, the Action delivered: a revision of the DCAT-AP and two extensions to describe geospatial and statistical data; supported the DCAT-AP implementation at Eu and MSs level; drafted a specification for publishing State Aid Notifications as open data working closely with DG COMP; supported the work to develop the Common Vocabulary for describing legal documents in the context of the IFC; drafted a specification for publishing the EU budget as open data with the PO; run a pilot with ESCO and DG Employment; run a pilot with the Core Public Service Vocabulary in Estonia; supported the creation of the data model for the Business Registers Interconnection System (BRIS) for DG Justice; provided the semantic layer for EIRA; supported the Secretariat General of the EC and EFSA to establish an information management strategy; supported EU inter-institutional work to establish persistent URIs management; supported local pilots to use the Core Vocabularies in Italy and Belgium.

Activities in the new WP

Activities promoting semantic interoperability at horizontal level

In the new ISA^2 Programme WP2017, the Action will continue to maintain, update whenever necessary, and promote the above-mentioned specifications. It will also try to identify and develop more Core Vocabularies (e.g. Core Budget), promote the real usage of the existing specifications, make available a robust toolset for metadata management, further elaborate on policies for metadata and URI management. More specifically the Action will deliver the following results:

- Core Vocabularies
    - Identification of areas for new Core Vocabularies and/or common data models.
    - Checking maturity and opportunities for developing specifications with communities.
    - New vocabularies development.
    - Sustainability of the existing specs.
- Policy and project support to DGs and MSs for promoting semantic interoperability with pilots and policy advice including the areas of open and linked data.
- Support for using the ISA data models in operational systems in DGs and MSs
- Use of core data models in base registries.
- Work for metadata management, including documenting existing policies and tools, drafting guidelines, architectures, providing direct support to MSs and DGs, aligning and coordinating relevant initiatives.
- Information management strategy: good practices, guidelines, recommendations, tools.
- Support work for open data and open data portals interoperability with further animation of the community of practice for DCAT-AP implementation.
- Continuation of the promotion and dissemination work including the SEMIC conference, identification of good practices, case studies, recommendations.

VocBench 3.2

Development will follow the modalities established in the previous ISA^2 actions, based on planning through JIRA, bug reporting, unit testing and possibly adding continuous development and deployment through a dedicated server.

The planned features (with section references – where applicable – to deliverable ISA 2015 SEMIC 7, D03.03: "Report on the provided support for the development of VocBench 3.0" and the set of user requirements expressed there) are:

5.3.2.1. Graph and Additional Visualisations for data

NF-O-GV-1. Lay the basis for additional visualizations

FR-O-63. Enabling the graphical visualisation of the dataset.

FR-O-64. The graphical visualisation will be possible for different resources and over search results and SPARQL query results.

FR-O-65. The dataset graphical visualisation will provide an extension point for exporting the graph into arbitrary formats.

FR-O-66. One of the formats above will be SVG

FR-New-GV-1. role/context-based extension points for pluggable actions on resources (developed in phase 2)

FR-New-GV-2. additional visualizations besides Graph-View

Advanced Search

FR-New-AS Restore VB2-like search with the following requirements:
FR-New-AS.1 - search by status
FR-New-AS.2 - search by scheme filtering
FR-New-AS.3 - search by outgoing relationship to resource/term
FR-New-AS.4 - search by ingoing relationship
FR-New-AS.5 - search by type

5.1.6. User Customizable Search Support (SPARQL based)

FR-New-SS.1 Customizable searches based on user-defined and storable SPARQL queries
FR-O-39. save and load specific search requests per user and project.
FR-O-40. Possibility to define private search requests (visible only to the specific user) and public search requests (accessible to the all the users in the same project).
FR-New-SS.2 privacy/visibility rules for storing SPARQL queries for custom search

5.1.7. Extended SPARQL Support

FR-O-42. The SPARQL Panel will provide autocompletion on resource names
FR-O-43./alt Ability to save and load specific SPARQL queries per user and projects
FR-O-44. Ability to store public/private queries
FR-O-45. Possibility to generate Query Templates
FR-O-46. Export datasets resulting from the execution of SPARQL queries

5.1.5.2. Drag and Drop Support

FR-O-30. Ability to drag and drop concepts in the hierarchical tree structure. The system will validate the proposed drag & drop gesture
FR-O-31. After a successful drag & drop gesture the system will present a list of the actions to be performed on the dataset, requiring confirmation by the user to proceed.
FR-O-32. On a failed of the drag & drop action the system will present to the user the reason why the action is not possible.
FR-O-33. The user will be able to select more than one concept at once for performing a drag and drop action.
FR-O-50. The Administrator or the Project Manager will be able to disable drag and drop functionality between concepts for a specific project.
FR-New-DDS-1. The drag and drop will be extended not only to concepts, but to all major kind of resources shown in hierarchical views. Different management possibilities will be offered depending on the resource

Support for UNDO

Enabled when history is active (or, at least, with the presence of the sail tracking component, developed in the 2016 VB3 action). Specific behaviour and dependency analysis when multiple users are editing mus be considered

2.1.11 COSTS AND MILESTONES

2.1.11.1 Breakdown of anticipated costs and related milestones

Main SEMIC action

Phase: Initiation Planning Execution Closing/Final evaluation	Description of milestones reached or to be reached	Anticipated Allocations (KEUR)	Budget line ISA/ others (specify)	Start date (QX/YYYY)	End date (QX/YYYY)
Operational	Identification of areas for new Core Vocabularies and/or common data models. Checking maturity and opportunities for developing specifications with communities. New vocabularies development. Update and sustainability of the existing specs.	551	ISA[2]	Q2/2016	Q4/2018
	Policy and project support to DGs and MSs for promoting semantic interoperability, linked and open data with pilots and policy advice.	550	ISA[2]	Q2/2016	Q4/2018
	Use of core vocabularies in base registries.	200	ISA[2]	Q2/2016	Q4/208
	Support for using the ISA data models in operational systems in DGs and MSs	300	ISA[2]	Q2/2016	Q4/2018
	Work for metadata management, including	200	ISA[2]	Q2/2016	Q4/2018

	documenting existing policies and tools, drafting guidelines, architectures, providing direct support to MSs and DGs, aligning and coordinating relevant initiatives.				
	Support to define information management strategies: good practices, guidelines, recommendations, tools.	100	ISA[2]	Q2/2016	Q4/2018
	Support work for open data interoperability with animation of a community of practice for DCAT-AP, maintenance and support for its real life-implementations.	150	ISA[2]	Q2/2016	Q4/2017
	Continuation of the promotion and dissemination work including country visits, the annual SEMIC conference, international workshops and webinars.	300	ISA[2]	Q2/2016	Q4/2018
	Total	**2351**			

VocBench subactivity

Phase: Inception	Description of milestones reached or to be reached	Anticipated Allocations (in KEUR)	Budget line ISA/ others (specify)	Start date (Q2/2016)	End date (Q2/2017)
Inception	Study, testing and comparison of state of the art technologies that are suitable for the design of VB3.0 and its technological innovation.	10	ISA[2]	Q2/2016	Q2/2016
Inception	Definition of the new	5	ISA[2]	Q2/2016	Q2/2016

	technological stack				
Execution	Designing and Developing the Semantic Turkey architecture as the core system of VB3.0.	80	ISA^2	Q3/2016	Q3/2017
Execution	Enabling all user aspects (Authentication, permissions etc.) in ST	25	ISA^2	Q3/2016	Q3/2017
Execution	Action History: completely rewritten wrt old module in VB2.0.	15	ISA^2	Q3/2016	Q3/2017
Execution	Controlled Collaboration & Validation: completely rewritten wrt old module in VB2.0.	25	ISA^2	Q3/2016	Q3/2017
Execution	Versioning	25	ISA^2	Q3/2016	Q3/2017
Execution	Provenance Management	15	ISA^2	Q3/2016	Q3/2017
Execution	Metadata Registry and Retriever and integration in VB	20	ISA^2	Q3/2016	Q3/2017
Execution	MAPLE Finalization and integration in VB	20	ISA^2	Q3/2016	Q3/2017
Execution	Lexical Linked Data Registry, starting and managing the Service and making VB compliant with it.	25	ISA^2	Q3/2016	Q3/2017
Execution	Improved Alignment: access and interaction with LOD data.	15	ISA^2	Q3/2016	Q3/2017
Execution	Improved Alignment : more automatism/support for validation of externally loaded mappings	15	ISA^2	Q3/2016	Q3/2017
Execution	Metadata producer: generators for ADMS/DCAT-AP/VoID/LIME Metadata.	10	ISA^2	Q3/2016	Q3/2017
Execution	Development of extensions points for : - customized export - connectable collaboration Frameworks	15	ISA^2	Q3/2016	Q3/2017
Execution	JIRA connector as an instance of a collaboration framework connector. Execution reported in 2017-2018 (Vocbench 3.1)	8	ISA^2	Q3/2016	Q3/2017
Execution	Dedicated SKOS-AP-EU exporter as an instance of the export extension point.	5	ISA^2	Q1/2017	Q3/2017
Execution	Property Language Tags Support (sections 5.1.5.3, 5.10 and 5.10.1 of D03.03)	9	ISA^2	Q3/2017	Q4/2017
Execution	Data Export Management: Excel Export (section 5.3.2 of D03.03)	1	ISA^2	Q3/2017	Q3/2017
Execution	Extensible URI Generation System	12	ISA^2	Q3/2017	Q4/2017
Execution	Alignment covering datasets exposed through HTTP dereferenciation and datasets	10	ISA^2	Q3/2017	Q4/2017

	exposed on a SPARQL endpoint (completion of requirement FR-I-35 started in first action on VocBench 3)				
Execution	Integrated Constrains Validation Service (sect. 5.9 of D03.03)	6	ISA²	Q3/2017	Q4/2017
Execution	User Notification Support (sect. 5.10.1.1 of D03.03)	21	ISA²	Q3/2017	Q2/2018
Execution	Sheet2RDF: Spreadsheet Import (part of it is described in sect. 5.3.1. Data Import Management of D03.03)	37	ISA²	Q3/2017	Q2/2018
Execution	Extension Points: role/context-based extension points for pluggable actions on resources	19	ISA²	Q3/2017	Q2/2018
Execution	MAPLE (was left out of first ISA action on VocBench 3)	25	ISA²	Q3/2017	Q2/2018
Execution	Support for OWL 2 (requirement NF-O-1 of D03.03)	28	ISA²	Q3/2017	Q2/2018
Execution	Rendering Engine Extension Point and three implementations for RDFS, SKOS and SKOSXL lexical models	8	ISA²	Q3/2017	Q3/2017
Execution	Search Support (indexing and autocompletion, sect. 5.1.6. of D03.03)	8	ISA²	Q3/2017	Q4/2017
Execution	Data Export Management (section 5.3.2 of D03.03)	12	ISA²	Q3/2017	Q4/2017
Execution	Overall Testing	14	ISA²	Q3/2017	Q2/2018
Execution	Metadata Registry and Retriever (adding retriever part and the LOD registry part, while metadata production has been implemented in the first action on VocBench 3)	26	ISA²	Q4/2017	Q2/2018
Execution	Alignment Management Service (sect. 5.5 of D03.03)	21	ISA²	Q4/2017	Q2/2018
Execution	Action Level History Management (sect. 5.7.3 of D03.03)	12	ISA²	Q4/2017	Q2/2018
Execution	Resource Level Metadata (sect 5.6.2 of D03.03)	21	ISA²	Q4/2017	Q2/2018
Execution	Provenance Management	12	ISA²	Q4/2017	Q2/2018

	(parts of 5.6.4 not dealt with in first action on VocBench 3)				
Execution	SKOS Vocabulary Constraint Management (sect. 5.9.1.1. of D03.03)	15	ISA^2	Q4/2017	Q1/2018
Execution	Advanced Concept Management (parts of 5.1.5 not dealt with in round 1)	12	ISA^2	Q1/2018	Q2/2018
Execution	Multiple Properties Editing Support (section 5.1.5.1 of D03.03)	4	ISA^2	Q1/2018	Q1/2018
Execution	Support for UNDO (sect. 5.10. of D03.03):	6	ISA^2	Q1/2018	Q1/2018
Execution	Other Collaboration Services (sect. 5.8.3. of D03.03): Development of extensions points for: connectable collaboration frameworks (ext.point.2 was removed in round 1)	14	ISA^2	Q1/2018	Q2/2018
Execution	Specific extension point for connectable collaboration frameworks implementation for JIRA Support (sect. 5.8.3.1. of D03.03)	9	ISA^2	Q1/2018	Q2/2018
Execution	Extension point for accessing dataset repositories and registries, searching for datasets of interest, importing them	17	ISA2	Q4/2018	Q4/2018
Execution	Development of Specific Connectors to repositories and registries for the provisioning of datasets	15	ISA2	Q4/2018	Q1/2019
Execution	Advanced Concept Management (parts of 5.1.5 not dealt with in round 1)	13	ISA2	Q1/2019	Q1/2019
Execution	Multiple Properties Editing Support (section 5.1.5.1 of D03.03)	5	ISA2	Q1/2019	Q2/2019
Execution	5.3.2.1. Graph and Additional Visualisations for data	38	ISA2	Q1/2019	Q2/2019
Execution	5.1.5.2. Drag and Drop Support	8	ISA2	Q4/2018	Q1/2019
Execution	Support for UNDO (sect. 5.10. of D03.03):	7	ISA2	Q4/2018	Q1/2019
Execution	Action Level History Management: History Based Time Machine (sect. 5.7.3 of D03.03)	14	ISA2	Q3/2018	Q2/2019
Execution	Extension Points: role/context-based extension points for pluggable actions on resources	21	ISA2	Q3/2018	Q2/2019

	Total	**833**			

2.1.11.2 Breakdown of ISA funding per budget year

Budget Year	Phase	Anticipated allocations (in KEUR)	Executed budget (in KEUR)
2016	Operational	950	950
2016	Execution (VocBench)	348	348
2017	Operational	746	746
2017	Execution (VocBenck 3.1)	347	347
2018	Operational	800	
2018	Execution (VocBenck 3.2)	138	

2.1.12 ANNEX AND REFERENCES

Description	Reference link	Attached document
Federation of Semantic Assets Repositories	https://joinup.ec.europa.eu/catalogue/all?filters=bs_current_version:true	
Core Vocabularies (Core Person, Core Business, Core Location, Core Public Service)	https://joinup.ec.europa.eu/community/core_vocabularies/description	
The DCAT Application Profile	https://joinup.ec.europa.eu/asset/dcat_application_profile/description	
Towards Open Government Metadata	http://www.semic.eu/semic/view/documents/towards_open_government_metadata.pdf	
Similar project in another, related domain: Specifications for the development of an open metadata schema for qualifications		20150717 Specifications.pdf 20150717 Annex to the Specifications.pdf