

DEVELOPMENT OF AN OPEN DATA SERVICE, SUPPORT AND TRAINING PACKAGE IN THE AREA OF LINKED OPEN DATA, DATA VISUALISATION AND PERSISTENT IDENTIFICATION (2016.18)

IDENTIFICATION OF THE ACTION

Type of Activity	Common Services and Reusable Generic Tools
Service in charge	Publications Office of the European Union
Associated Services	EC: JRC, CONNECT, DIGIT, COMM, ESTAT. Agencies: EMA, ERA, INEA.

EXECUTIVE SUMMARY

The action aims at providing data providers (institutions, agencies and other bodies) of the European Union Open Data Portal (EU ODP) and other stakeholders (Pan-European Data Portal including affiliated Member State portals) with an open data service package in order to enable them to further open up their data and increase data interoperability in view of better data reusability and data visualisation.

It has been triggered by the needs expressed by European Commission DGs, institutions and agencies during workshops and individual meetings with EU ODP teams and additionally brought to evidence by the interinstitutional survey on data visualisation needs (carried out between May and July 2015), namely

- for a **structural approach for sharing knowledge about open data management** (Guidelines, consultancy, assistance, training, awareness raising, etc.), reusable data structures, data visualisation and the development of shareable tools;
- for significantly **mutualising and rationalising efforts** towards the identification, development and sharing of modern **data processing and visualisation tools** (i.e. a corporate visualisation solutions pool);
- to **raise awareness** about the dependency of easily reusable visualisation tools on interoperable data structures, and to put in place structural measures in order to achieve this alignment;
- to support further opening of data namely by assistance in transforming data into Linked open data (LOD) and persistent identification.

Through its different action strands (implementing guidelines, consultancy and hands-on assistance, workshops and training sessions, development of toolsets, etc.), the project will deliver an

- interinstitutional **knowledge base** on open data publishing, open data treatment and processing, data visualisation;
- interinstitutional **pool of reusable** data processing and data visualisation **tools**;
- governance structure, **implementation guidelines** and development of a toolset for unique and **persistent identification based** upon the sub-domain data.europa.eu and accompanied, where suitable, by the Digital Object Identifier (DOI).

Moreover, it will contribute to building up an intra-institutional community of practice in the domains of open data and data visualisation (a network of open data and data visualization stewards) and an increased data publishing maturity of participating institutions, agencies and bodies.

The project will be implemented with existing key partners of the Open Data Portal (data providers, stakeholders) as well as with new partners to be identified during the follow-up phase of the interinstitutional survey on visualisation needs (ISA 1.1 – January-June 2016)

OBJECTIVES

The project aims at providing data providers (institutions, agencies and other bodies) of the Open Data Portal (EU ODP) and other stakeholders of the Open Data Portal (Pan-European Data Portal including affiliated MS portals) with an open data service package in order to enable them to further open up their data and increase data interoperability in view of better data reusability and data visualisation.

SCOPE

The project covers

- open data publishing, open data treatment and processing, data modelling in view of linked open data, data visualisation (setup of an interinstitutional knowledge base)
- the set-up of an interinstitutional pool of reusable data processing and data visualisation tools
- the implementation of a toolset for unique and persistent identification based upon the sub-domain data.europa.eu and accompanied, where suitable, by the Digital Object Identifier
- community building in all abovementioned areas

It entails activities in the areas of:

- consultancy and conceptual assistance,
- development of training material, delivery of training sessions/workshops,
- development of toolsets (including catalogues of tools)
- development of interoperability guidelines and documentation

ACTION PRIORITY

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
<p><i>Does the proposal directly contribute to implementing the European Interoperability Strategy, the European Interoperability Framework, or other EU policies with interoperability requirements, or needed cross-border or cross-sector interoperability initiatives? If yes, please indicate the EU initiative / policy and the nature of contribution.</i></p>	<p>The proposal contributes to the implementation of the European Interoperability Framework in particular to the underlying principles of openness, reusability, transparency, effectiveness and efficiency.</p> <p>It also corresponds to the European Interoperability Strategy, namely the strategic approach 14.7 for the accompanying measure 'Sharing Best Practices'</p>

<i>Does the proposal fulfil an interoperability need for which no other alternative solution is available?</i>	Yes, to our knowledge no other initiative among the EU institutions addresses the topic of interoperability for data visualisation purposes.

Cross-sector

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

Question	Answer
<i>Will the proposal, once completed be useful, from the interoperability point of view, and utilised in two (2) or more EU policy areas? If yes, which are those?</i>	Data visualisation is per se of cross-sector relevance. The participants in the working group come from across all institutions, agencies and other bodies. The pilots foreseen envisage the visualisation of research results (CORDIS), EU budget, etc.
<i>For proposals or their parts already in operational phase: have they been utilised in two (2) or more EU policy areas? Which are they?</i>	Not applicable

Cross-border

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

Question	Answer
<i>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?</i>	The elaboration of the catalogue and of the pilots considers also experiences made in Member States. Member States administrations could consult the catalogue, after ECAS authentication. Member States could also reuse the solution.
<i>For proposals or their parts already in operational phase: have they been utilised by public administrations of three (3) or more EU Members States?</i>	Not applicable

Urgency

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

Question	Answer
<i>Is your action urgent? Is its implementation foreseen in an EU policy as priority, or in EU legislation?</i>	Data visualisation and synergies in data management are key topics in the Communication on Data, Information and Knowledge Management at the European Commission. Furthermore, we embed and feed-back all outputs of this ISA ² project into the Commission internal initiative Data4Policy.
<i>Does the ISA² scope and financial capacity better fit for the implementation of the proposal as opposed to other identified and currently available sources?</i>	Yes, the strong focus of ISA ² on interoperability, standards, reuse and cross-service cooperation is especially propitious for the topics data visualisation, data management, unique and persistent identification and linked open data.

Reusability of action outputs

The re-usability of the action, measured by the extent to which its results can be re-used

Can the results of the 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	Common toolset in the domains data visualization and the reuse-oriented data management and data treatment (in particular LOD)
Description	Toolset (presented most probably as a catalogue of applications, libraries, etc.) consisting of reusable tools for data visualization and data treatment and conversion (methodologies, develop common libraries for data conversion and for visualisation, enabling the presentation of data with different granularity and modality (spatial, temporal, organizational, etc.).
Reference	
Target release date / Status	Q4 2017
Critical part of target user base	The output can be re-used by a critical part of their target user

	base.
For solutions already in operational phase - actual reuse level (as compared to the defined critical part)	Not applicable

Name of reusable solution	Governance structure, implementation guidelines and development of a toolset for unique and persistent identification based upon the sub-domain data.europa.eu
Description	<ul style="list-style-type: none"> • Setup of URI governance model • Development of a toolset for decentralised URI generation and management • Development of a central URI Register
Reference	
Target release date / Status	Q4 2017
Critical part of target user base	The output can be re-used by a critical part of their target user base.
For solutions already in operational phase - actual reuse level (as compared to the defined critical part)	Not applicable

Output name	Data models, ontologies and content of pilots treated within the project
Description	<ul style="list-style-type: none"> • Transformation of some data into LOD (CORDIS, TED, ESTAT, BUDGET, other interested DGs) and publishing them on EU ODP • Visualizations produced and listed on EU ODP
Reference	
Target release date / Status	Q4 2017
Critical part of target user base	The output can be re-used by some users.
For solutions already in operational phase - actual reuse level (as compared to the defined critical part)	Not applicable

Level of reuse by the proposal

The re-use by the action of existing common frameworks and elements of interoperability solutions.

Question	Answer
<p>Does the proposal intend to make use of any ISA², ISA or other relevant interoperability solution(s)? Which ones?</p>	<ul style="list-style-type: none"> • The proposal addresses the need of a common pool of reusable visualisation tools. Those tools are being described using ADMS¹. • It recommends that metadata describing those dataset be expressed in DCAT AP². • Recommendations of the interinstitutional URI Committee based upon the domain data.europa.eu³. • Core vocabularies
<p>For proposals or their parts already in operational phase: has the action reused existing interoperability solutions? If yes, which ones?</p>	<p>Not applicable</p>

Interlinked

The link of the action with Union initiatives to be measured by the collaboration and contribution level of the action to Union initiatives such as the DSM.

Question	Answer
<p>Does the proposal directly contribute to at least one of the Union's high political priorities such as the DSM? If yes, which ones? What is the level of contribution?</p>	<p>(Open) government data is a core asset for the knowledge-based economy, since its reuse is a basis for innovative information products and services. It makes a significant contribution to the Digital Single Market and is a key enabler for transparency, evidence-based decision-making and a broader participation in the political discourse. As a by-product, it enhances administrative efficiency through streamlined data management. The project aims at providing data providers (institutions, agencies and other bodies) of the EU Open Data Portal and other stakeholders of the EU Open Data Portal (Pan-European Data Portal including affiliated MS</p>

¹ <https://joinup.ec.europa.eu/asset/adms/home>

² https://joinup.ec.europa.eu/asset/dcat_application_profile/description

³ <https://joinup.ec.europa.eu/community/semic/news/persistent-uri-taskforce-experiments-central-uri-redirection-service>

	portals) with an open data service package in order to enable them to further open up their data and increase data interoperability in view a better data reusability and data visualisation.
--	---

PROBLEM STATEMENT

A) EU institutions, agencies and other bodies need a structural approach for sharing knowledge about open data visualisation and a pool of reusable data visualisation tools

Given the growing amount of data produced, gathered, processed and published by the EU institutions, there is a need to develop effective tools for visualising data. At the moment Commission DGs, institutions and agencies have common visualisation needs. Sharing of knowledge and tools would strengthen and facilitate their work and allow for cost reduction.

B) EU institutions, agencies and other bodies need a structural and reuse-oriented approach for sharing knowledge about open data management and publishing, data models, data identification, and the development of shareable tools.

EXPECTED BENEFICIARIES AND ANTICIPATED BENEFITS

Beneficiaries	Anticipated benefits
European Commission, all other institutions, agencies and bodies interested in participating in the projects	<ul style="list-style-type: none"> ○ Single point of access to a catalogue of shareable visualisation tools, as a first step towards a corporate pool of visualisation solutions. ○ Interoperability guidelines in the domains of data visualisation, data treatment for reuse and visualisation, open data publishing, in particular as Linked Open Data ○ Possibility to obtain training, share knowledge and conceptual support ○ Increased data publishing maturity of participating institutions, agencies and bodies ○ Building up an interinstitutional community of open data and data visualisation experts, evangelists and quasi-contact points
Users of the Open Data Portal and the Pan-European Portal (diverse reuse communities)	<ul style="list-style-type: none"> ○ Better data quality, including better documentation ○ More and better visualisations ○ More open datasets published on the Open Data Portal
Member States Portals and other portals	<ul style="list-style-type: none"> ○ Reuse of the developed guidelines and toolsets

EXPECTED MAJOR OUTPUTS

Output name	Knowledge base and the reuse-oriented data management and data treatment (in particular towards LOD) and data visualisation
Description	Knowledge base consisting of: <ul style="list-style-type: none"> ○ Implementation guidelines ○ Training material, footage of training sessions, ○ briefings on conceptual assistance, development of interoperability guidelines and documentation
Reference	
Target release date / Status	As of project inception, continuously enriched and improved

Output name	Delivery of training on open data, linked data visualisation and data visualisation
Description	The following training sessions are foreseen: <ul style="list-style-type: none"> ○ 3 introductory half-day training sessions, with 2 instructors, and ○ 6 two-day workshops, with minimum 2 instructors, on data publishing, visualisation and related tools.
Reference	
Target release date / Status	Q1-Q2 2017

ORGANISATIONAL APPROACH

Expected stakeholders and their representatives

Stakeholders	Representatives
Publications Office	Open Data Portal (Norbert Hohn, Agnieszka Zajac) Common repository (Cellar) : Peter Schmitz Metadata registry : Willem van Gemert EU Budget: Carmen Malagon
Eurostat	Marc Vanderperren (Renovation of the ESTAT dissemination chain)
CONNECT	Daniele Rizzi (Pan-European Data Portal)
JRC	Andrea Perego, Anders Friis Christensen
DIGIT	TBD

COMM	Jose Arcos
Others	Other key stakeholders identified during ISA 1.1 Action on data visualisation (Jan-June 2016)

Identified user groups

The following main user groups have been identified:

- The staff of the European Commission, institutions, agencies and other bodies working on data management, data visualisation, publishing and presentation.
- The staff of public administrations in Member States especially statistical institutes and national data portals.
- The users of data portals: researchers, journalists, data reusers.

Communication plan

The results of the visualisation survey and the actions to be derived from it will be ready to be discussed with ISPMB (The European Commission Information Systems Project Management Board) and the stakeholders' community not earlier than late September 2015. This process will require a structured approach over 2-3 months in order to properly collect input, and to define the follow-up strategy and the measures to be taken.

For this, it is proposed to use an action within ISA 1.1 as a communication platform for bringing together the stakeholders and define the requirements, priorities and work plan for a subsequent implementation within an ISA² action as of May 2016.

During the lifetime of the ISA² project, communication of the project will be ensured through the following channels:

- Open Data Portal network and the Publications Office's publishing officers network towards institutional data providers
- Pan-European Data portal towards affiliated Member State portals
- Open Data Portal social media and other channels towards reuser communities

Governance approach

The Publications Office will ensure the chairmanship and the project management.

All key stakeholders will be involved in the piloting of the data visualisation catalogue.

TECHNICAL APPROACH AND CURRENT STATUS

Following the preparatory work done under ABCIII SC 245 "D04.08 Reusable data visualisation tools to support data-driven policy making", the implementation of the catalogue of reusable data visualisation tools will involve taking decisions about its high-level architecture, starting the development of a pilot of the catalogue and building a knowledge base on data visualisation tools.

The catalogue will be created following a collaborative and iterative approach that will engage a number of key stakeholders from the EU Institutions, not limited to the OP.

The catalogue of reusable data visualisation tools will be a service package providing to EU staff support related to the use of data visualisation and data visualisation tools. It is composed of three elements:

- the catalogue itself, which represents a collection of data visualisation tools and visualisations made available via an online service. In addition to tools and visualisations, the knowledge base of the catalogue will also provide links to relevant documentation, studies and reports, learning resources, user reviews of data visualisation tools etc.;
- a community around the catalogue comprising EU staff who will exchange knowledge and ideas in an online environment to the EU Institutions; and
- a set of capacity building services for data visualisation, e.g. provision of training and creation of sample visualisations for selected datasets.

The evolution of the catalogue will be community-driven, built in a step-wise manner.

The core service will be developed first, by adding functionalities incrementally on the principle of iterative prototyping and based on user needs and demand. This approach allows for monitoring the catalogue's success over time and performing a cost-benefit analysis on the next steps.

The metadata associated with the description of data visualisation tools are described based on ADMS-AP⁴ (a profile of the DCAT-AP used to describe reusable solutions).

In regard to the implementation of the catalogue, four alternatives have been identified:

- the catalogue implemented using the Drupal 7 multisite platform of the European Commission;
- the catalogue as a part of Joinup;
- the catalogue as a part of a standalone service, or
- the catalogue as a part of EU Open Data Portal.

Current status: the Publications Office is preparing the specific contract within ABCIII Framework contract with PwC and Everis consortium for a period of 12 months.

COSTS AND MILESTONES

Breakdown of anticipated costs and related milestones

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)
	Project Management	50	ISA ²	Q2/2016	Q4/2017
	Development of catalogue of data visualisation tools and knowledge base	200	ISA ²	Q2/2016	Q4/2017
	Support services for developing open data capabilities in the EU	110	ISA ²	Q2/2016	Q4/2017

⁴ ADMS-AP: https://joinup.ec.europa.eu/asset/adms/asset_release/adms-ap-joinup-version-20

	Institutions				
	Publishing and visualising EU budget and related data	40	ISA ²	Q2/2016	Q4/2017
	Raising awareness and communication	18	ISA ²	Q2/2016	Q4/2017
	Total	418			

Breakdown of ISA funding per budget year

Budget Year	Phase	Anticipated allocations (in KEUR)	Executed budget (in KEUR)
2016		300	
2017		118	
2018			
2019			
2020			