



ODINE

H2020-ICT-2014-1

Project Number: 644683

D1.2 Data Management Plan

Coordinator: Luis-Daniel Ibáñez (SOTON)

Quality reviewer: Fabrizio Orlandi (IAIS)

Deliverable nature:	Report (R)
Dissemination level: (Confidentiality)	Public (PU)
Contractual delivery date:	31/07/2015
Actual delivery date:	31/07/2015
Keywords:	Data Management

Document Information

IST Project Number	644683	Acronym	ODINE
Full Title	Open Data Incubator for Europe		
Project URL	http://opendataincubator.eu		
Document URL			
EU Project Officer	Francesco Barbato		
Deliverable	D1.2 Data Management Plan		
Work Package	WP1 Project management		

1 Executive Summary

The Data Management Plan (DMP) outlines the datasets that will be generated or collected during the project's lifetime highlighting the following information:

- How datasets will be exploited/shared/licensed. For those that cannot be shared, the reasons why are explained.
- Which standards are followed for publishing datasets.
- Which strategies are used for curation and archiving of datasets.

2 List of Datasets

2.1 Management Deliverables

Dataset Reference	DS-1A
Belonging Work Package	WP1
Type	Generated
Origin	Project's partners
Scale (Approx End Volume)	Small (~50Mb)
Description	All deliverables of type report related to the project's management: Data Management Plan, financial reports, management reports and Project Management Board minutes.
Useful to	European Commission, project partners, future Horizon2020 projects and future initiatives concerning start-up incubation and Open Data.
Integration and Reuse possibilities	Linking to previous and same-date Horizon2020 projects. Re-use by future Horizon2020 projects.
Standards	.pdf files following guidelines and templates provided by the European Commission.
Data Sharing Policy	During their elaboration, documents will remain internal to

	partners. Once delivered, the European Commission will decide what is the appropriate sharing policy.
Archiving and Storage	During the project's lifetime, documents will be stored in the Google Drive space of the project. After the end of the project, the European Commission will decide the proper repository in which to store them.
Preservation Time	Decided by the European Commission for Horizon2020 projects.
Additional preservation cost	None.

2.2 General Statistics on Open Call

Dataset Reference	DS-1A
Belonging Work Package	WP2
Type	Collected
Origin	Submission Platform
Scale (Approx End Volume)	Small (~10Mb)
Description	<p>Statistics on the Open Call, per evaluation round and consolidated, including:</p> <ul style="list-style-type: none"> ● Number of applications ● Number of applicants ● Size in employees of applicants ● Number of shortlisted applications ● Number of successful applications ● Number of projects successfully completed. <p>For each of the above indicators, per country numbers will be available.</p>
Useful to	<p>Incubation initiatives and Decision Makers on start-up acceleration policies who can use these datasets to improve their work.</p> <p>Open Data publishers and users can get insight on the interest of Open Data for SMEs.</p>
Integration and Reuse	Integration with statistics from other incubation initiatives for

possibilities	Open Data can provide insight on the best funding/grant model. Comparison with incubation initiatives on other domains can provide insight on the net value of each domain and the dynamism of the business ecosystem for each domain.
Standards	Spreadsheets on .xls and .odt formats, plus .csv files
Data Sharing Policy	Creative Commons Attribution 4.0 International License .
Archiving and Storage	Stored in the Google Drive of the project and published on the project's website.
Preservation Time	To be defined in the project's sustainability strategy. Refer to task 6.3 (Sustainability).
Additional preservation cost	None.

2.3 Database of submissions

Dataset Reference	DS-2B
Belonging Work Package	WP2
Type	Collected
Origin	Submission Platform
Scale (Approx End Volume)	Small (~100Mb)
Description	The database of submissions contains the full information of all the applications submitted to ODINE.
Useful to	Evaluators and mentors within ODINE.
Integration and Reuse possibilities	None
Standards	Backend of the submission platform.
Data Sharing Policy	Not to be shared. The intellectual property of the ideas outlined in the submissions belong to the applicants. Only the statistics for dataset DS-2A will be collected.
Archiving and Storage	During the project's lifetime, it will be stored in the submission platform. Afterwards, it will be destroyed to guarantee

	confidentiality
Preservation Time	Project's lifetime.
Additional preservation cost	None.

2.4 Video Recordings of interviews

Dataset Reference	DS-2C
Belonging Work Package	WP2
Type	Generated
Origin	Submission Platform
Scale (Approx End Volume)	Medium (~100Gb)
Description	Submitted applications will be shortlisted for an interview with the evaluators to select the granted ones. These interviews will be recorded to help evaluators in their decisions, and mentors in their mentorship.
Useful to	Evaluators and mentors within ODINE.
Integration and Reuse possibilities	None
Standards	.mp4 format
Data Sharing Policy	Not to be shared. The intellectual property of the ideas outlined in the interview belong to the applicants.
Archiving and Storage	During the project's lifetime, it will be stored in the project's Google Drive folder. Afterwards, it will be destroyed to guarantee confidentiality
Preservation Time	Project's lifetime.
Additional preservation cost	In the event the current Google Drive space (15GB) is not enough to store the videos, extra space needs to be purchased. a 1TB quota in Drive costs 9.99\$ per month

2.5 Evaluation results

Dataset Reference	DS-2D
Belonging Work Package	WP2
Type	Generated
Origin	Submission Platform
Scale (Approx End Volume)	small (~50Mb)
Description	Unsuccessful applications that are not short-listed will receive an email with the comments of the evaluators. Unsuccessful applications that were shortlisted will receive more detailed comments agreed by both evaluators.
Useful to	Evaluators and applicants
Integration and Reuse possibilities	None
Standards	text and email.
Data Sharing Policy	Not to be shared. Comments belong to the applicants, as they reveal details about their ideas.
Archiving and Storage	During the project's lifetime, it will be stored in the project's Google Drive folder. Afterwards, it will be destroyed to guarantee confidentiality
Preservation Time	Project's lifetime.
Additional preservation cost	None.

2.6 Statistics on startup incubation (Open Data Acceleration)

Dataset Reference	DS-3A
-------------------	-------

Belonging Work Package	WP3
Type	Collected
Origin	Internal reports
Scale (Approx End Volume)	small (~10Mb)
Description	<ul style="list-style-type: none"> ● From the tracking activity, publish aggregate data about the level of satisfaction with the programme. ● Access statistics of online training materials: <ul style="list-style-type: none"> ○ number of users ○ the most and less consulted. ● Success rate of startups achieving milestones. (achieved milestones / total milestones). ● Percentage of successfully completed projects. ● Count of mentors split by sector and level of satisfaction of the mentorship.
Useful to	Other Open Data incubation initiatives. Institutions promoting Open Data could use this dataset to identify places where promotion has been done and its relative success. Reference point for learning about best practices.
Integration and Reuse possibilities	With similar data from other incubation initiatives.
Standards	.xls, .odt and .csv . This dataset will also be published in .pdf format as part of deliverable 3.3
Data Sharing Policy	Creative Commons Attribution 4.0 International License .
Archiving and Storage	During the project's lifetime, it will be stored in the project's Google Folder.
Preservation Time	To be defined in the project's sustainability strategy. Refer to task 6.3 (Sustainability)
Additional preservation cost	None.

2.7 Statistics on usage of Data and Computing Infrastructure.

Dataset Reference	DS-4A
-------------------	-------

Belonging Work Package	WP4
Type	Collected
Origin	Data and Computing Infrastructure
Scale (Approx End Volume)	small (~10Mb)
Description	<p>For the selected projects, statistics will be collected on:</p> <ul style="list-style-type: none"> • How many of them use the platform offered by ODINE. • Bugs and Issues with the platform. • Which of the tools provided by the consortium the project is using. • Which other (freely or payed) tools the project is using. • Which Open Datasets the project is using. • Which not open datasets the project is using.
Useful to	<ul style="list-style-type: none"> • Other Open Data incubation initiatives. • FI-Ware core developers and maintainers. • Developers of Open Data toolkits. • Dataset providers.
Integration and Reuse possibilities	Assigning each project an URI and linking it to the datasets it uses enables reuse through Linked Data technologies.
Standards	Bugs and issues will be published in the FI-Ware platform. .xls, .odt, .csv and RDF will be used for the rest.
Data Sharing Policy	Creative Commons Attribution 4.0 International License .
Archiving and Storage	Bugs and issues will be archived in the platform. The other stats will be collected by ODINE partners and stored in the project's Drive space.
Preservation Time	To be defined in the project's sustainability strategy. Refer to task 6.3 (Sustainability)
Additional preservation cost	None.

2.8 Data Value Chain.

Dataset Reference	DS-4B
-------------------	-------

Belonging Work Package	WP4
Type	Generated
Origin	Consortium research
Scale (Approx End Volume)	small (~10Mb)
Description	We will create a comprehensive database of potential data producers, business models, industries, and reuse prospects. For example, meteorological institutes produce weather data of varying granularity (micro and macro weather data, short, medium and long term prognosis). Insurance companies on the other hand are prospective users of such data (particularly of micro area and long term weather prognosis). Once such a database of data assets, requirements and business models is established, we can start to systematically engage and connect with stakeholders in the identified data value chains. Refer to deliverable 4.3 (Data Value Chain), for more details.
Useful to	<ul style="list-style-type: none"> • Entrepreneurs and SMEs. • Business analysts. • Organisers of networking events.
Integration and Reuse possibilities	Linking with datasets on companies and SMEs and with datasets on business models.
Standards	RDF
Data Sharing Policy	Creative Commons Attribution 4.0 International License .
Archiving and Storage	Special section on project's website.
Preservation Time	To be defined in the project's sustainability strategy. Refer to task 6.3 (Sustainability)
Additional preservation cost	To be defined.

2.9 Stakeholder Database.

Dataset Reference	DS-5A
Belonging Work Package	WP5

Type	Collected
Origin	<ul style="list-style-type: none"> ● Opencorporates ● various business directories ● Global Incubator ● FS6 ● ODI ● Wayra ● various smaller sources
Scale (Approx End Volume)	small (~100Mb)
Description	<p>As part of task 5.1, we will build and maintain a stakeholder database including entrepreneurs, SMEs and SME association, investors and data economy/community.</p> <p>Refer to deliverable 5.3 (Stakeholder database and network) for more details.</p>
Useful to	<ul style="list-style-type: none"> ● Business analysts of data economy ● Entrepreneurs and SMEs ● Venture capitalists and investors.
Integration and Reuse possibilities	Linking with existing datasets on SMEs, like Open Corporates, Open Check or government registries.
Standards	<p>During early stages of its collection, spreadsheets will be used. From 1000 entries, it will be migrated to a database and an API and visualization tool will be provided.</p> <p>An RDF version of the dataset will also be provided for linking with other datasets.</p>
Data Sharing Policy	Creative Commons Attribution 4.0 International License .
Archiving and Storage	A dedicated section on the website
Preservation Time	To be defined in the project's sustainability strategy. Refer to task 6.3 (Sustainability)
Additional preservation cost	To be defined.

