



# D1.7 Data Management Plan adhering to the H2020 Open Research Data Pilot v2



This project has received funding from the European Union's Horizon 2020 Programme under Grant Agreement N°. 825014.

<b>Workpackage</b>	WP1 – Project Management
<b>Editor(s)</b>	KCL, ZABALA
<b>Responsible partner</b>	KCL
<b>Contributors</b>	KCL, ZABALA
<b>Status-Version</b>	v1
<b>Due to</b>	30/06/2020
<b>Submission date</b>	12/06/2020
<b>EC Distribution</b>	PU
<b>Abstract</b>	<p>This document is the Data Management Plan v2 of the Data Market Services Accelerator. (DMS Accelerator).</p> <p>DMS Accelerator is a €3M H2020 project that aims to support the emergence of data markets and the data economy in Europe.</p>

## Document History

Version	Date	Comment
v.01	18/05/2020	First draft prepared by KCL
v.02	21/05/2020	Second version of the deliverable including ZABALA's inputs
v.1	12/06/2020	Final version to submit

## Contents

1. Introduction	4
1.1. Project Participants	4
1.2. Purpose of the data collection and generation	5
1.3. Types of personal data generated/collected	6
2. Data description	6
2.1 Overview of datasets	6
Datasets details	7
EU DMS approached participant portfolio	8
DMS applicant SMEs	9
Applicant SME ratings (cohorts 1, and 2)	10
DMS participant portfolio (cohorts 1 and 2)	11
Webinar recordings	12
Webinar engagement reports	13
Course engagement reports	14
3. Data security	15
4. Ethical aspects	15

## Tables

Table 1: Project participants	5
Table 2: Overview of datasets	7
Table 3: EU DMS approached participant portfolio	8
Table 4: DMS applicant SMEs	9
Table 5: Applicant SME ratings (cohort 1)	10
Table 6: DMS participant portfolio (cohort 1)	11
Table 7: Webinar recordings	12
Table 8: Webinar engagement reports	13
Table 9: Course engagement reports	14

## 1. Introduction

This document is the updated Data Management Plan of the Data Markets Services Accelerator Project (DMS).

DMS Accelerator is a €3M H2020 project that aims to support the emergence of data markets and the data economy in Europe.

### 1.1. Project Participants

The participants of the project are listed in Table 1. There has been one addition, KCL, since the first version of this DMP was submitted. Their role, together with its relevance to the data management plan, is described in column 4.

Name of organisation	Short name	Country	Role
Zabala Innovation Consulting S.A.	ZABALA	ES	Coordination. Set up of the mechanism of monitoring. Relations with additional initiatives from Big Data Value PPP (European Data Incubator (EDI), Big Data Value eCosystem Project (BDVe)), FIWARE Foundation, Innovation Radar, Start-up Europe. Zabala handles the DMS email account, and as such is on hold of email details of participants.
Ogilvy worldwide S.A.	OG1	ES	Building of the contents for the promotion portfolio of initiatives. They provide training and may hold details of training participants for communications specific to the training sessions.
University of Southampton	SOTON	UK	Data science courses provider. They hold details of course participants.
King's College London	KCL	UK	Data science courses provider. They hold details of course participants.
IPtector Consulting	IPtector	DK	IPR training and GDPR. They provide training, and may hold details of training participants for communications specific to the training sessions.
GEIE ERCIM	ERCIM/W3C	FR	Standardisation. They will provide training, and may hold details of training participants for communications specific to the training sessions.

Bright Pixel	BRPX	PT	Company builder owned by a corporation. They provide training and mentoring and may hold details of participants for communications specific to these purposes.
Spherik Accelerator	SPHK	RO	Accelerator in Eastern Europe. They provide training and mentoring and may hold details of participants for communications specific to these purposes.
Spinlab	SPINLAB	DE	Accelerator in Germany. They provide training and mentoring and may hold details of participants for communications specific to these purposes.
The Next Web	TNW	NL	Incubator, Media publication, Data market itself. They provide promotion in the TNW conference, training, and mentoring, and may hold details of participants for communications specific to these purposes.

Table 1: Project participants

## 1.2. Purpose of the data collection and generation

DMS Accelerator is a Horizon 2020 project that aims to overcome the barriers of data-centric European SMEs and start-ups by providing free support services around data skills, entrepreneurial opportunities, legal issues, and standardisation. The expected project deliverables are a 150-data-based company portfolio, and the delivery of free services sorted out among 5 different categories: (1) Fundraising; (2) Acceleration; (3) Standards & Legal; (4) Data skills and (5) Promotion. To achieve this, data about SMEs will be collected for the following purposes:

- Contacting and engaging with the enterprises who will benefit from the services, in the form of marketing communications and individual communications when support is requested.
- Providing services such as fundraising, acceleration, standardisation & legal advice, data skills training, and promotion.
- Monitoring and analysing these services and their outcomes, in alignment with the project's Key Performance Indicators.
- Disseminating the results of such services, primarily in the form of white papers.

As the above list of purposes reflects, **no research data has been or will be handled in this project**. Therefore, this document will only include information about:

- what data will be collected, processed, and/or generated
- which methodology & format will be used
- how data will be curated & preserved (including after the end of the project).

### 1.3. Types of personal data generated/collected

Since no research data will be handled, personal data will be the main concern of DMS Accelerator relevant to this data management plan. There are three types of personal data collected by DMS:

- **Contact information** i.e. names, email addresses, and phone numbers – in order to communicate with applicants, both successful and unsuccessful. The legal basis for which are (unless otherwise specified\*) legitimate interests – to identify potential data providers.
- **Recordings of DMS Accelerator webinars.** These webinars will contain voice and written contributions from the participants. It will be necessary for DMS Accelerator legitimate interests to keep these recordings, and release them as services to participant SMEs.
- **Course participation data.** Participants' interactions in the online courses provided by DMS Accelerator will be recorded and analysed for monitoring and course evaluation purposes. The course participants will be identified through their names and email contact details.

## 2. Data description

### 2.1 Overview of datasets

The project collects and generates datasets with various data types and formats. Table 2 is a list of the datasets of the second cohort. These datasets are based on datasets of cohort 1.

Dataset reference	Relevant work package(s)	Dataset Name	First recorded in DMP
DMS-2A(y2)	2	EU DMS approached participant SMEs portfolio (new cohorts added to this dataset)	M16
DMS-2B(y2)	2	DMS applicant SMEs (new cohorts added to this dataset)	M16
DMS-2D(y2)	2	Applicant SME ratings (new cohorts added to this dataset)	M24
DMS-2E(y2)	2	Participant portfolio (new cohorts added to this dataset)	M24
DMS-3A(y2)	3	Webinar recordings (For all rounds, years 1,2 and 3)	M12

DMS-4A(y2)	4	Webinar engagement reports (For all rounds, years 1,2 and 3)	M12
DMS-4B(y2)	4	Course engagement reports (For all rounds, years 1,2 and 3)	M12

Table 2: Overview of datasets

### **Datasets details**

Each of the datasets used in this project is described below. Each description contains the following information:

- Dataset reference: A unique reference to each of the datasets, as listed in Table n.2
- Relevant work package/s
- Type
  - Collected
  - Generated
- Origin: Where does the data in the dataset come from.
  - Where does the data in the dataset come from?
  - From which sources it has been collected
- Scale: the size of the dataset in Mb/Gb
- Description: A description of the dataset, including fields when tabular.
- Useful to: the reason why it is worth generating, keeping and sharing when applicable
- Methodology & format:
  - How the data will be input
  - Which tool will be used
  - The format in which the data will be kept
- Data sharing policy: stakeholders with whom the data will be shared
- Archiving and storage: where the data will be stored.
- Preservation time
- Additional preservation cost



### EU DMS approached participant portfolio

Dataset Reference	DMS-2A(y2)
Relevant Work Package	WP2 SME engagement
Type	Collected
Origin	Project partners' networking connections
Scale (Approx End Volume)	Small (~50Mb)
Description	<p>All SMEs approached to apply for DMS services. The dataset contains the following fields:</p> <ul style="list-style-type: none"> <li>● SME name</li> <li>● SME URL</li> <li>● Partner-owner (who owns the contact)</li> <li>● Contact name</li> <li>● Contact email</li> <li>● Notes (description, comments)</li> </ul>
Useful to	The DMS consortium in approaching companies
Methodology & format	Manually created in Google Spreadsheets, manual input by participants.
Data Sharing Policy	Internal.
Archiving and Storage	This dataset will be stored in the project Google Drive and destroyed at the end of the preservation time. This includes the destruction of any copies on Google Drive and personal computers.
Preservation Time	Five years beyond the end of the project.
Additional preservation cost	None.

Table 3: EU DMS approached participant portfolio

**DMS applicant SMEs**

Dataset Reference	DMS-2B(y2)
Relevant Work Package	WP2 SME engagement
Type	Collected
Origin	Online form in DMS website
Scale	Small (~100Mb)
Description	<p>All SMEs which applied for DMS services by filling a form. The dataset contains the following fields:</p> <ul style="list-style-type: none"> <li>● Company name and email</li> <li>● Applicant name</li> <li>● Company financial details (yearly revenue, % of public funding)</li> <li>● Growth stage of the company</li> <li>● Relationship with the data market</li> <li>● Company category</li> <li>● Types of services applying for</li> <li>● Description of equality policy</li> <li>● Description of needs relevant to DMS services.</li> </ul>
Useful to	The DMS consortium in selecting SMEs for providing them with services.
Methodology & format	<p>Automatically collected from the online form linked to the DMS website, over three iterations, one per year.</p> <p>The online form lies in INDEX, a daughter company of TNW. INDEX is built with <a href="#">Datalicious</a>, and one of their tools are online forms.</p> <p>The online form generates a CSV, only accessible to TNW.</p> <p>Data from new cohorts (2 and 3) will be added to this dataset.</p>
Data Sharing Policy	Internal.
Archiving and Storage	This dataset will be stored in the servers of INDEX TNW, the partner in charge of WP2.
Preservation Time	Five years beyond the end of the project.
Additional preservation cost	None.

Table 4: DMS applicant SMEs

### Applicant SME ratings (cohorts 1, and 2)

Dataset Reference	DMS-2D(y2)
Relevant Package Work	WP2 SME engagement
Type	Generated
Origin	TNW's DMS Applicant SMEs dataset (DMS-1B)
Scale	Small (~100Mb)
Description	<p>All data contained in DMS-2B plus numerical and categorical ratings to each of the SMEs from four of the accelerator partners on the following categories:</p> <ul style="list-style-type: none"> <li>● Motivation (0-10)</li> <li>● Services required (0-5)</li> <li>● Estimated annual revenue (1-3)</li> <li>● Public funding &lt;40% (Yes/No)</li> <li>● Equality (0-1)</li> </ul>
Useful to	The DMS consortium in selecting SMEs for service provision
Methodology & format:	<p>CSV from DMS-2B is loaded in Microsoft Excel, and copies are distributed by Zabala to each of the rating partners. Partners rate and return to Zabala. Zabala merges the ratings.</p>
Data Sharing Policy	Internal.
Archiving and Storage	This dataset will be stored in Zabala servers and destroyed at the end of the preservation time. This includes the destruction of any copies on Excel and personal computers.
Preservation Time	Five years beyond the end of the project.
Additional preservation cost	None.

Table 5: Applicant SME ratings (cohort 1)

***DMS participant portfolio (cohorts 1 and 2)***

Dataset Reference	DMS-2E(y2)
Relevant Work Package	WP2 SME engagement
Type	Collected
Origin	TNW's DMS Applicant SMEs dataset (DMS-1B)
Scale (Approx End Volume)	Small (~100Mb)
Description	Services to which each of the 50 selected SMEs have signed up to.
Useful to	The DMS consortium in allocating services to selected SMEs
Methodology & format:	CSV from the survey collected in Typeform.
Data Sharing Policy	Internal.
Archiving and Storage	This dataset will be stored in Zabala's Typeform account and destroyed at the end of the preservation time.
Preservation Time	Five years beyond the end of the project.
Additional preservation cost	None.

Table 6: DMS Accelerator participant portfolio (cohort 1)

### Webinar recordings

Dataset Reference	DMS-3A(y2)
Relevant Work Package	WP3 Services
Type	Generated
Origin	DMS account in LogMeIn
Scale (Approx. End Volume)	Medium (~100Gb)
Description	Recordings of all video teleconference webinar sessions run by all consortium members, participated by selected SMEs members of staff, over the three iterations of the project.
Useful to	Participant SMEs as beneficiaries of the DMS Accelerator services.
Integration and Reuse possibilities	None
Methodology & format	Webinars are recorded in GotoWebinar and shared privately (only those with the link can see) in YouTube, for streaming only. The Youtube video link is shared on the DMS Accelerator website, only accessible through a password.
Data Sharing Policy	Private, accessible to participant SMEs through password.
Archiving and Storage	These files will be stored in Zabala's LogMeIn account, accessible for streaming through GoToStage (a tool within LogMeIn) during the preservation time period.
Preservation Time	Five years beyond the end of the project.
Additional preservation cost	None.

Table 7: Webinar recordings

**Webinar engagement reports**

Dataset Reference	DMS-4A(y2)
Relevant Work Package	WP4 Monitoring and analysis
Type	Collected
Origin	DMS account in LogMeIn
Scale (Approx. End Volume)	Small (~100Mb)
Description	Data about the interactions of participants in all webinars. Relevant fields are: <ul style="list-style-type: none"> <li>• Name</li> <li>• Email</li> <li>• Engagement measurements</li> <li>• Satisfaction survey</li> </ul>
Useful to	The WP4 leader for monitoring and analysing the performance of the services, and producing deliverables
Methodology & format	Automatically generated in LogMeIn. Downloaded by WP4 leader for aggregate analysis. Name and email fields deleted immediately after download.
Data Sharing Policy	Internal.
Archiving and Storage	This dataset will be stored in DMS LogMeIn account, and destroyed when the account is deleted.
Preservation Time	Five years beyond the end of the project.
Additional preservation cost	None.

Table 8: Webinar engagement reports

### Course engagement reports

Dataset Reference	DMS-4B(y2)
Relevant Work Package	WP4 Monitoring and analysis
Type	Collected
Origin	SOTON account in Canvas
Scale (Approx. End Volume)	Small (~100Mb)
Description	Data about the interactions of participants in online courses. Relevant fields are: <ul style="list-style-type: none"> <li>• Name</li> <li>• Email</li> <li>• Engagement measurements</li> </ul>
Useful to	The WP4 leader for monitoring and analysing the performance of the services, and producing deliverables
Methodology & format	Automatically generated in Canvas. Downloaded by WP4 leader for aggregate analysis. Name and email fields deleted immediately after download.
Data Sharing Policy	Internal.
Archiving and Storage	This dataset will be stored in Soton Canvas account, and destroyed when courses in the account are deleted.
Preservation Time	Five years beyond the end of the project.
Additional preservation cost	None.

Table 9: Course engagement reports

### **3. Data security**

All data will be safely stored in the password-protected accounts of the platforms where the data is held (GoToWebinar, Typeform, and Canvas). When the data is generated, it will be safely stored in the dedicated partners' servers. All these data security measures are compliant with the GDPR regulation.

The coordinator makes sure to regularly update the access list of personnel within the teams of the partners to the files and sensitive information contained on them.

### **4. Ethical aspects**

As a GDPR compliant data management plan, personal data will be requested only when strictly needed, and only for the purposes stated when data is requested. Name and corporate email will be the only fields requested to participants. When requesting personal data, disclaimers will be shown, with a clear statement of the purpose of collecting and keeping such information.