



INVITE

Open Innovation 2.0 LAB

DATA MANAGEMENT PLAN - INITIAL VERSION

November, 2017



This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N° 763651.

Main authors	
Name	Organisation
Mr Kevin Trendafili, Mr Kostas Giagtzoglou, Mr Kostas Bougiouklis	Q-PLAN

LEGAL NOTICE

The information and views set out in this report are those of the authors and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

© INVITE Consortium, 2017

Reproduction is authorised provided the source is acknowledged.

INVITE

Project Information

Title	“Co-designing and piloting demand-driven mechanisms, skill-sets and measures for stimulating and facilitating open innovation across European innovation systems” (GA Number 763651)
Duration	September 2017 – August 2020 (36 months)
Website	www.invite-project.eu / www.oi2lab.com
Coordinator	Q-PLAN INTERNATIONAL ADVISORS PC
Contact person	Mr Kostas Giagtzoglou (+30 2310 411 191, giagtzoglou@qplan.gr)
Project overview	<p>INVITE is set on co-creating a well-connected European Open Innovation (OI) ecosystem, in which knowledge meaningfully flows across borders and is translated into marketable innovations, bringing increased socio-economic benefits to EU citizens. To this end, INVITE will co-design, pilot and demonstrate a pan-European service platform, the Open Innovation Lab, aiming to: (i) better link the currently fragmented innovation systems of the EU by facilitating meaningful cross-border knowledge flows; (ii) empower EU businesses with the skill-sets required to tap into Europe’s knowledge-base and turn it into value; and (iii) increase the participation of private investors in OI and collaborative innovation projects. To this end, the Open Innovation Lab will experiment with novel bottom-up collaborative models of innovation and leverage OI support services and ICT tools to stimulate and support OI across Europe, building a vibrant community of OI actors and stakeholders (including academia, government, industry and civil society) along the way. The valuable knowledge, evidence and experiences gained through the experiments of the Open Innovation Lab will be diffused across the EU so as to fuel their replication and scale-up for the benefit of the European economy and society as a whole.</p>
Consortium	<ol style="list-style-type: none"> 1. Q-PLAN INTERNATIONAL ADVISORS PC www.qplan-intl.com – Greece 2. STEINBEIS INNOVATION GGMBH www.steinbeis-europa.de – Germany 3. EUROPE UNLIMITED S.A. www.e-unlimited.com / www.techtour.com – Belgium 4. RTC NORTH LIMITED www.rtcnorth.co.uk - United Kingdom 5. NINESIGMA EUROPE BVBA www.ninesigma.com – Belgium 6. INTRASOFT INTERNATIONAL SA www.intrasoft-intl.com - Luxembourg 7. CENTRE FOR RESEARCH & TECHNOLOGY- HELLAS www.iti.gr – Greece 8. WIRTSCHAFTSFOERDERUNG REGION STUTTGART GMBH www.region-stuttgart.de – Germany 9. THE DURHAM, GATESHEAD, NEWCASTLE UPON TYNE, NORTH TYNESIDE, NORTHUMBERLAND, SOUTH TYNESIDE AND SUNDERLAND COMBINED AUTHORITY www.nelep.co.uk – United Kingdom

Table of Contents

EXECUTIVE SUMMARY	2
1. INTRODUCTION	3
2. DATA MANAGEMENT PRINCIPLES	5
2.1 Data Summary.....	5
2.2 FAIR data principles.....	6
2.2.1 <i>Standards and Metadata</i>	6
2.2.2 <i>Sharing, Re-use and Interoperability</i>	6
2.2.3 <i>Preservation and Open Access</i>	6
2.3 Data security	7
2.4 Ethical aspects.....	7
3. DATA MANAGEMENT PLAN	8
3.1 Overview	8
3.2 Market gaps and opportunities (WP1).....	9
3.3 User needs and requirements (WP1).....	10
3.4 Co-creation workshop outcomes (WP1)	11
3.5 Pilot monitoring, co-evaluation and validation (WP3).....	12
3.6 Open Innovation Lab user data (WP3).....	14
3.7 Business model validation and improvement (WP4).....	15
3.8 Dissemination and communication results (WP6).....	16
4. CONCLUSIONS AND WAY FORWARD	17
ANNEX.....	18

Executive Summary

This document constitutes the initial version of the Data Management Plan (DMP) and has been elaborated as a deliverable (D7.2) in the framework of the INVITE project. INVITE aims at co-creating a better connected Open Innovation (OI) ecosystem across Europe. It envisions an OI ecosystem in which knowledge meaningfully flows across borders and is translated into marketable innovations, bringing increased socio-economic benefits to European citizens.

To this end, INVITE will co-design, pilot and demonstrate a suite of well-tailored mechanisms, skill-sets and measures to be deployed by a pan-European service platform, the Open Innovation Lab, with a view to better linking the currently fragmented innovation systems of the EU by facilitating meaningful cross-border knowledge flows; empowering EU businesses with the skill-sets required to tap into Europe's knowledge-base and turn it into value; and increasing the participation of private investors in OI as well as collaborative innovation projects.

The Open Innovation Lab will experiment with novel bottom-up collaborative models of innovation and leverage OI support services and ICT tools to stimulate and support OI across Europe, building a vibrant community of OI actors and stakeholders (including academia, government, industry and civil society) along the way. The valuable knowledge, evidence and experiences gained through the experiments of the Open Innovation Lab will be diffused across the EU so as to fuel their replication and scale-up for the benefit of the European economy and society as a whole.

Under this light, it becomes evident that INVITE entails several activities within its framework which involve the collection, production and /or processing of data, with a view to generating meaningful insights that will feed into the project and fuel the co-creation and delivery of truly demand-driven and evidence-based results.

In this context, the initial version of INVITE's DMP sets out the overall methodological principles pertaining to the management of the data that that will be collected, processed and/or generated in the framework of INVITE, safeguarding sound, FAIR and ethical data management along the entire duration of the project. Moreover, it provides a first, yet still meaningful overview of INVITE's datasets, as identified in this early stage of the project, along with information on the specific methodology pertaining to their management on a dataset by dataset basis.

The initial version of the DMP is the first of the three versions of INVITE's Data Management Plan to be produced in the course of the project and will serve as living document. Along these lines, the DMP will be updated and further elaborated during the project in order to reflect an accurate, up-to-date and ultimately comprehensive plan for managing the data that will be collected, and/or generated by the project across their entire life cycle, both during and after the completion of INVITE.

1. Introduction

The current document, titled Data Management Plan – Initial Version (DMP), has been elaborated within the framework of the **INVITE** project which has received funding by the European Union’s Horizon 2020 Research and Innovation programme under Grant Agreement No 763651.

INVITE is set on co-creating a well-connected European Open Innovation (OI) ecosystem. It envisions an OI ecosystem in which knowledge meaningfully flows across borders and is translated into marketable innovations, bringing increased socio-economic benefits to EU citizens. To this end, INVITE will co-design, pilot and demonstrate a pan-European service platform, the **Open Innovation Lab**, aiming to better link the currently fragmented innovation systems of the EU by facilitating meaningful cross-border knowledge flows; empower EU businesses with the skill-sets required to tap into Europe’s knowledge-base and turn it into value; and increase the participation of private investors in OI and collaborative innovation projects.

The Open Innovation Lab will experiment with novel bottom-up collaborative models of innovation and leverage **OI support services** and **ICT tools** to stimulate and support OI across Europe, building a vibrant community of OI actors and stakeholders (including academia, government, industry and civil society) along the way. The valuable knowledge, evidence and experiences gained through the experiments of the Open Innovation Lab will be diffused across the EU so as to fuel their replication and scale-up for the benefit of the European economy and society as a whole.

To this end, INVITE has brought together a well-balanced and complementary **consortium**, that consists of **9 partners across 5 different European countries**, as presented in the following table.

Table 1: INVITE partners

No	Name	Short name	Country
1	Q-PLAN INTERNATIONAL ADVISORS PC	Q-PLAN	Greece
2	STEINBEIS INNOVATION GGMBH	SEZ	Germany
3	EUROPE UNLIMITED SA	E-UNLIMITED	Belgium
4	RTC NORTH LBG	RTC NORTH	United Kingdom
5	NINESIGMA EUROPE BVBA	NINESIGMA	Belgium
6	INTRASOFT INTERNATIONAL SA	INTRASOFT	Luxembourg
7	CENTRE FOR RESEARCH AND TECHNOLOGY HELLAS	CERTH/ITI	Greece
8	WIRTSCHAFTSFOERDERUNG REGION STUTT GART	WRS	Germany
9	THE DURHAM, GATESHEAD, NEWCASTLE UPON TYNE, NORTH TYNESIDE, NORTHUMBERLAND, SOUTH TYNESIDE AND SUNDERLAND COMBINED AUTHORITY	NELEP	United Kingdom

In this context, **all partners of INVITE's consortium will adhere to sound data management** in order to ensure that the meaningful data that will be collected, processed and/or generated throughout the duration of the project are well-managed, archived and preserved, taking into account the [Guidelines on Data Management in Horizon 2020](#).

Along these lines, the **objectives of this initial version of the DMP** are to:

- Provide an overview of the principles underpinning the data management methodology that will be employed with a view to safeguarding the sound management of the data to be collected, processed and/or generated in the framework of INVITE, while also making them Findable, Accessible, Interoperable and Re-usable (FAIR).
- Identify the data that will be collected, processed and/or generated during the project as well as present meaningful information on how they will be handled, the methodology and standards applied to each one as well as how they will be curated and preserved during and after the project (if applicable) on a dataset by dataset basis.

With the above in mind, the initial version of **the DMP is structured in 4 distinct chapters**, as follows:

- **Chapter 1** provides introductory information about the initial version of the DMP, the context in which it has been elaborated as well as about its objectives and structure.
- **Chapter 2** describes the principles that are applied in the framework of INVITE in order to safeguard the effective management of data across their entire lifecycle, in line with the guidelines of the Commission.
- **Chapter 3** presents a description of the data that will be collected, processed and/or generated, addressing important relevant aspects such as methodology and metadata as well as data sharing, archiving and preservation on a dataset by dataset basis.
- **Chapter 4** outlines concludes on the next steps of the project with respect to data management.

Finally, the **Annex** of this document, includes the template for collecting data management information per dataset in the framework of INVITE.

The DMP is not a fixed document. In fact, this is its **initial version** and it will evolve during the lifespan of the project. In particular, the DMP will be **further elaborated and updated twice throughout the duration of INVITE (i.e. as D7.3 at M12 and as D7.4 at M36)** as well as ad hoc (if necessary), in order to include new datasets, better detail and/or reflect changes in the methodology applied or other aspects pertaining to the already identified datasets (such as costs for making data FAIR, size of dataset, metadata, etc.), changes in consortium policies and plans or other potential external factors. Q-PLAN is responsible for the elaboration of the DMP and with the support of all partners will update and enrich it when required.

2. Data management principles

This chapter of the DMP presents the **overall methodological principles pertaining to data management** in the framework of INVITE. Further details with respect to the specific methodology applied for each of the different datasets to be collected, processed and/or generated over the course of INVITE (as identified in this early stage of the project) are provided in Chapter 3. With that in mind, the DMP's second chapter starts by providing a general summary of the data to be collected, processed and/or generated by the project as well as their types and versioning. It proceeds and concludes by presenting the overall approach employed for making data FAIR, ensuring data security and ultimately taking into account ethical aspects in this context.

2.1 Data Summary

INVITE will produce several datasets during the lifetime of the project. The data included within these **datasets may be quantitative, qualitative or a blend** of those in nature and will be **analysed from a range of methodological perspectives** with a view to producing meaningful insights that will feed the activities of the project and fuel the delivery of evidence-based results. These datasets will be **available in a variety of easily accessible formats**, including post scripts (e.g. pdf, xps, etc.), spreadsheets (e.g. xlsx, csv, etc.), text documents (e.g. docx, rtf, etc.), compressed formats (e.g. rar, zip, etc.) or any other format required depending on the objectives and methodology of the activity within the frame of which they are produced.

Moreover, in order to facilitate the reference of the datasets that will be collected and/or generated during INVITE, a **standard naming and versioning structure** will be employed, as follows:

INVITE _ [Name of Study] _ [Issue Date]

- **INVITE:** The name of the project.
- **Name of Study:** A short version of the name of the study for which the dataset is created.
- **Issue Date:** The date on which the latest version of the dataset was modified (DD.MM.YYYY).

With the above in mind, some **indicative examples** to showcase the naming structure applied in the context of INVITE are provided below:

- **INVITE_Needs&Requirements_31.10.2017** – A dataset generated within the framework of the survey conducted to identify the needs and requirements of diverse open innovation stakeholders. This is the version of the dataset that was last modified on the 31st of October 2017 (31/10/2017).
- **INVITE_BMValidation_01.02.2018** – A dataset created in the process of validating and improving the business models developed for the Open Innovation Lab with a view to feeding the elaboration of the business plan that will guide its market rollout beyond the end of the project. The last modification of this dataset was on the 1st of February 2018 (01/02/2018).

2.2 FAIR data principles

The [Guidelines on Data Management in Horizon 2020](#) of the Commission emphasise the importance of making the data produced by projects funded under Horizon 2020 **Findable, Accessible, Interoperable as well as Reusable (FAIR)**, with a view to ensuring their sound management. This means using standards and metadata to make data discoverable, specifying the data sharing procedures and which data will be open, allowing data exchange via open repositories as well as facilitating the reusability of the data. With that in mind, the following sections of the DMP lay out the principles followed in the framework of INVITE with respect to the standards and metadata required to make data findable, the sharing procedures foreseen to make the data accessible and safeguard their interoperability as well as ensure their preservation and open access, making them easily reusable by interested stakeholders.

2.2.1 Standards and Metadata

Any open datasets produced by INVITE will be accompanied by data that will facilitate their understanding and re-use by interested stakeholders. These data may include basic details that will assist interested stakeholders to locate the dataset, including its format and file type as well as meaningful information about who created or contributed to the dataset, its name and reference, date of creation and under what conditions it may be accessed. Complementary documentation may also encompass details on the methodology used to collect, process and/or generate the dataset, definitions of variables, vocabularies and units of measurement as well as any assumptions made. Finally, wherever possible consortium partners will identify and use existing standards.

2.2.2 Sharing, Re-use and Interoperability

The Project Coordinator (Q-PLAN) in collaboration with the respective Work Package Leaders (WPL) and any other involved project partners, will determine whether and how the data collected and/or produced in the framework of INVITE's different activities will be shared and/or re-used either by other project partners or by external interested stakeholders, within and without the framework of the project. This includes the definition of access procedures as well as potential embargo periods along with any necessary software and/or other tools which may be required for data sharing and re-use. In case the dataset cannot be shared, the reasons for this will be clearly mentioned (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related). An informed consent form will be requested from all external data providers in order to allow for their data to be analysed and shared, while all such data will be anonymised before sharing (for more details in this respect see Section 2.4 of the current document).

2.2.3 Preservation and Open Access

Any datasets that will be deemed open for sharing and re-use will be deposited to an open data repository and will be made accessible to all interested stakeholders, ensuring their long-term preservation and accessibility beyond the lifetime of the project. At the moment, we consider the use of Zenodo (www.zenodo.org) as one of the best online and open services to enable open access to INVITE's datasets,

but similar repositories will also be considered and an appropriate decision will be timely made at a future update of the DMP.

Q-PLAN will be responsible for uploading all open datasets to the repository of choice, while all partners will be responsible for disseminating them through their professional networks and other media channels.

2.3 Data security

INVITE will handle any collected / generated data securely throughout their entire lifecycle. In this context, the project partner responsible for collecting / generating, processing and/or storing the data will ensure that they are protected and any necessary data security controls have been implemented, so as minimize the risk of information leak and destruction. Overall, data will be stored within the private server of the project partner responsible for the respective dataset and will be backed-up frequently to ensure their security.

2.4 Ethical aspects

INVITE entails activities which involve the collection of meaningful data from selected individuals (e.g. interviews with users and stakeholders of the Open Innovation Lab, etc.). The collection of data from participants in these activities will be based upon a process of informed consent. Any personal information will be handled according to the principles laid out by the Directive 95/46/EC of the European Parliament and of the Council on the “Protection of individuals with regard to the processing of personal data and on the free movement of such data” (24 October 1995) and its revisions as well as with relevant national regulations and laws. The participants’ right to control their personal information will be respected at all times (including issues of confidentiality). The Project Coordinator (Q-PLAN) will regulate and deal with any ethical issue that may arise during the project in this respect, in cooperation with the Steering Committee of the project.

The ethics aspects pertaining to the collection and processing of data in the framework of INVITE will be addressed in further detail within the ethics deliverables of the project namely “D8.1: H - Requirement No. 1” and “D8.2: POPD - Requirement No. 2”, both due for M6 of the project.

3. Data management plan

3.1 Overview

INVITE places special emphasis on the management of the valuable data that will be collected and/or generated throughout its activities. In this respect, the table below provides a list of the datasets identified by INVITE consortium members at this stage of the project, indicating the name of the dataset, its linked Work Package and the respective leading consortium member (i.e. Work Package Leader).

Table 2: List of INVITE datasets

No	Dataset Name	Linked Work Package	Work Package Leader
1	Market gaps and opportunities	WP1	CERTH/ITI
2	User needs and requirements	WP1	CERTH/ITI
3	Co-creation workshop outcomes	WP1	CERTH/ITI
4	Pilot monitoring, co-evaluation and validation	WP3	RTC NORTH
5	Open Innovation Lab user data	WP3	RTC NORTH
6	Business model validation and improvement	WP4	Q-PLAN
7	Dissemination and communication results	WP6	E-UNLIMITED

With the above in mind, **the current chapter provides meaningful information per each dataset**, including:

- The name of the dataset and the type of study in the frame of which it is produced.
- A concise description of the dataset as well as its format and volume.
- The methodology and tools employed for collecting/generating the data.
- Any standards that will be used (if applicable) as well as metadata to be created.
- Potential external stakeholders for whom the data may prove useful.
- Provisions regarding the confidentiality of the data.

The information provided within this section reflects the current views and plans of INVITE at this early stage of the project (M3) and will be further elaborated in future versions of the DMP (e.g. through the inclusion of more elaborate descriptions of the datasets, standards and metadata, how the datasets may be preserved, accessed and re-used in the long-term, etc.). The template employed for collecting the information from project partners is annexed to this document.

3.2 Market gaps and opportunities (WP1)

Dataset name	Analysis of Market Gaps and Opportunities (INVITE_Gaps&Opportunities).
Type of study	This study involves an evaluation of three existing Open Innovation (OI) platforms/service providers: Enterprise Europe Network (EEN), NineSigma and Steinbeis. The data that supports this evaluation will be collected in two phases. Phase 1 involves a desk review of secondary data sources that are in the public domain, whereas Phase 2 involves a series of semi-structured interviews with 6-8 respondents who work for either an EEN consortium member, NineSigma or Steinbeis for the Open Innovation Lab.
Dataset description	The data collected will consist of a combination of information extracted from the secondary data sources and information provided by the respondents during the depth interviews. In both cases, the data collected will be mainly of a qualitative nature and will be recorded in plain text in English. Both the secondary and primary data sources will be used as the basis of a report which will summarise the strengths and weaknesses of the three existing OI platforms/service providers and identify potential market gaps and opportunities.
Methodologies for data collection / generation	A semi-structured interview guide will be used to collect the data during the depth interviews. The interviews will be conducted face-to-face, by telephone or via Skype/WebEx (or similar online videoconferencing application).
Storage and volume of the dataset	In the case of the secondary data sources, copies of the data will not be stored. However, references and/or web links to the data sources consulted will be included as an appendix to the final report. In the case of the primary data collected during the depth interviews, the dataset will be stored in standard Word format (.docx) on RTC NORTH's in-house server and will be preserved for 5 years following the end of the project, before eventually being deleted. For security reasons, the dataset will be backed up at the end of each night on removable tape and will be stored in a fireproof safe.
Metadata and standards	The dataset will be accompanied with basic descriptive metadata (i.e. title, author, date created and keywords).
For whom might the dataset be useful?	The dataset will be used by INVITE partners to help design and pilot the Open Innovation Lab, based on the needs and requirements of a range of open innovation stakeholders.
Confidentiality	The primary data collected via the depth interviews will not be shared and/or re-used (outside the framework of the project and/or beyond its completion) to ensure the confidentiality of the interviewees and their responses.

3.3 User needs and requirements (WP1)

Dataset name	Analysis of user needs and requirements (INVITE_Needs&Requirements)
Type of study	<p>This qualitative study to be conducted in the framework of INVITE will follow an interview-based survey of user needs and requirements. In particular, a stratified, purposeful sampling will be employed, in order to include diverse stakeholders of a quadruple helix innovation system (e.g. private sector, academic organisations & research institutes, governmental and public services and civil society), and aimed at revealing their views regarding Open Innovation practices across Europe and informing the demand-driven development of INVITE’s Open Innovation Lab. In this case, the sample will be stratified in the sense that participants will vary according to stakeholder sector and level of engagement and purposeful as sample participants will be recruited through key organizations/businesses, which the project partners will identify as impactful to the project’s expected outcomes.</p>
Dataset description	<p>The data collected will contain the responses (mainly qualitative – plain text in English) provided by interviewees who will participate in the interview-based survey, addressing their different views of the current state-of-play in Open Innovation across Europe. No secondary data or other third-party sources will be used.</p>
Methodologies for data collection / generation	<p>Data will be collected via semi-structured questionnaires administered to study participants in the frame of interviews. Participants are to be recruited from all actors of the quadruple helix and invited to participate in the interviews over the phone, e-mail, or a face-to-face meeting.</p>
Format and volume of the dataset	<p>Private versions of the dataset will be accommodated at the data management portal created and maintained by CETH/ITI, while links to the portal will exist at the INVITE website with access restricted to the INVITE partners. The dataset, comprised of interview responses, will be stored in standard spreadsheet format (.xlsx). At least 50 interviews will be conducted across Europe and the same number of records will be collected and stored in the dataset. Furthermore, in order to avoid data losses, RAID and other common backup mechanisms will be utilized ensuring data reliability and performance improvement. The archiving system of CETH/ITI will contain the initial data as sent to the INVITE repository. The dataset will remain at the data management portal for the whole project duration, as well as at least 2 years after the end of the project. The volume of data is estimated to be approximately 10 MB for all interview responses collected and the analysis report submitted. Finally, after the end of the project, the portal is going to be accommodated with other portals at the same server for 5 years following the end of the project, before eventually being deleted, so as to minimize the cost of</p>

	maintenance. For security reasons, the dataset will be backed up in an external hard-drive every month.
Metadata and standards	Basic descriptive metadata (such as title, author, date created and keywords) will accompany the dataset.
For whom might the dataset be useful?	The dataset will be used by CERTH/ITI in order to analyse and extract the topics of interest that will inform INVITE partners in the process of designing the INVITE Co-creation Workshop and ultimately the first version of the Open Innovation Lab according to the needs and requirements of open innovation stakeholders.
Confidentiality	The dataset will remain confidential. Participant names will be protected and replaced by an abbreviated version of their sector, a sequential two-digit number and the initials of the project partner who conducted the interview (e.g. ARxxQPLN for Academia/Research sector, PSxxQPLN for Public Sector, BSxxQPLN for the Private Sector, CSxxQPLN for Civil Society Sector, where xx is a sequential two-digit number corresponding to the interviewee). The raw data collected through the interview-based survey will not be shared and/or re-used (outside the framework of the project and/or beyond its completion) to ensure the confidentiality of the interviewees and their responses.

3.4 Co-creation workshop outcomes (WP1)

Dataset name	Co-creation workshop outcomes (INVITE_Co-creation).
Type of study	A co-creation workshop whereby all invited participants will explore and provide feedback on the various aspects and features of the Open Innovation Lab to be co-created.
Dataset description	A record of feedback and suggestions expressed by participants of the workshop.
Methodologies for data collection / generation	All participants will be introduced to a series of ideas and concepts and along with them INVITE partners will co-create and conclude on user-driven characteristics for the Open Innovation Lab and interconnected pilot programme of the project. Different workshop sessions will be dedicated to discussing and co-creating the different components of the Open Innovation Lab (i.e. service, capacity building and ICT components) while at the same time collecting interesting ideas and feedback with respect to the design of INVITE pilots. Participants will be engaged in creative brainstorm and ideation sessions that will also include gamified exercises, which will serve as ice breakers allowing them to get to know each other better, break down their inhibitions and thus unlock their imagination. Simple illustrators will be employed to visually capture ideas and help them grow into concrete concepts.

	These will be recorded and be made available to INVITE partners for the design of the Open Innovation Lab.
Format and volume of the dataset	In the case of the primary data collected, including feedback and ideas resulting from the co-creation process, the dataset will be stored in standard Word format (.docx) on RTC NORTH's in-house server and will be preserved for 5 years following the end of the project, before eventually being deleted. For security reasons, the dataset will be backed up at the end of each night on removable tape and will be stored in a fireproof safe.
Metadata and standards	The dataset will be accompanied with basic descriptive metadata (i.e. title, author, date created and keywords).
For whom might the dataset be useful?	The dataset will be used by INVITE partners to help better design and shape the Open Innovation Lab to its users, based on the resulting feedback and ideas of the co-creation workshop participants.
Confidentiality	The primary data collected via the co-creation workshop will not be shared and/or re-used (outside the framework of the project and/or beyond its completion) to ensure the confidentiality of the participants and their contributions.

3.5 Pilot monitoring, co-evaluation and validation (WP3)

Dataset name	Pilot monitoring, co-evaluation and validation (INVITE_PilotMCV)
Type of study	The study will involve deploying the pilots of INVITE through the Open Innovation Lab in two rounds, in order to address the needs of users and stakeholders of the open innovation ecosystem as well as to collect data and feedback from them, so as to improve and fine tune the design of the pilots and the Open Innovation Lab.
Dataset description	Data and feedback provided by users (mainly SMEs) and other stakeholders participating in the pilots of INVITE based on a custom-made multi-layer framework for monitoring the operation, performance and results of the pilots. Indicative core themes to be monitored and measured include the integration of open innovation in the users' business model, external knowledge and technology search and acquisition, collaboration with other stakeholders, occasional vs continuous engagement in open innovation activities, disruptive vs incremental innovation, internal innovation capability, time-to-market, level of proficiency gained in collaborative innovation, scale achieved in terms of outreach (volume, sectoral and geographical), fundraising capacity, staff impact, organizational impact, cost-benefit and overall satisfaction. Moreover, the dataset will include qualitative data

	<p>on the perceived most significant change which pilots have brought about within the organizations of pilot users.</p>
<p>Methodologies for data collection / generation</p>	<p>Users and stakeholders who will participate in the pilots will be required to provide feedback as part of the ongoing monitoring framework that will be established to keep track of the performance and results of the pilots through a semi-structured questionnaire designed to mine both qualitative and quantitative and administered via a blend of face-to-face, telephone and/or digital means.</p>
<p>Storage and volume of the dataset</p>	<p>The dataset will be stored in either standard Word (.docx) or standard Excel (.xlsx) format on RTC NORTH's on-premise server and will be preserved for 5 years following the end of the project, before eventually being deleted. For security reasons, the dataset will be backed up at the end of each night on removable tape and will be stored in a secure place.</p>
<p>Metadata and standards</p>	<p>The dataset will be accompanied with basic descriptive metadata (i.e. pilot description, pilot participant name, organisation, etc.).</p>
<p>For whom might the dataset be useful?</p>	<p>The dataset will be used by INVITE partners to validate and fine-tune the Open Innovation Lab with the help of its Advisory Board members, based on the needs and requirements of a range of open innovation users and stakeholders who will have taken part in its interconnected pilot programme. Moreover, the dataset would be useful for open innovation intermediaries as well as (e-)training providers who design and/or offer relevant open innovation services.</p>
<p>Confidentiality</p>	<p>The data and feedback collected will not be shared and/or re-used (outside the framework of the project and/or beyond its completion) to ensure the confidentiality of the pilot programme participants and their input into the iterative pilot deployment, validation and fine-tuning phase. Still, some selected meaningful and properly anonymised aggregated data and results from pilot activities will be integrated within the INVITE Replication and Scale-up Guide that will be openly disseminated, with a view to providing interested stakeholders with insights and practical guidelines on how to replicate the results of INVITE within their context.</p>

3.6 Open Innovation Lab user data (WP3)

Dataset name	Open Innovation Lab user data (INVITE_OI2Labdata)
Type of study	Study aimed at informing the validation and fine-tuning of the Open Innovation Lab and its ICT infrastructure (including the online collaborative space, open multi-sided marketplace, crowdfunding tool and e-learning environment) as well as of the business models designed to guide its market rollout and scale up.
Dataset description	The dataset will contain demographic data of the users registered to Open Innovation Lab as well as data stemming from the use of the functionalities offered by the ICT tools integrated within its platform.
Methodologies for data collection / generation	Users and stakeholders wishing to employ the ICT tools of the Open Innovation Lab will be required to provide basic data through a dedicated online registration form built-in the ICT infrastructure to be implemented by INTRASOFT. Moreover, the platform will automatically keep track of all necessary data pertaining to the online activities of the users who will access the Open Innovation Lab via their unique username-password combination.
Storage and volume of the dataset	The data collected will be stored on a secure database to be employed by the ICT infrastructure of the Open Innovation Lab. INTRASOFT will be the main responsible partner to manage the server and to assign specific user roles with access and administration privileges. Administration privileges can be assigned only to the consortium members. There will be a secure environment per user, in order to guarantee the integrity of the user's personal data.
Metadata and standards	Descriptive metadata (i.e. title, author and keywords) will be created to accompany the dataset.
For whom might the dataset be useful?	The dataset will be used by INVITE consortium members to gain insights, analyze and improve the experience of users within the ICT infrastructure of the Open Innovation Lab and thus provide the evidence required to fuel its demand-driven validation and fine-tuning.
Confidentiality	In order to ensure the privacy of the users of the Open Innovation Lab as well as safeguard the project partners' commercial interests in view of its post-project market roll-out, the dataset will remain confidential. The data will not be shared and/or reused outside the framework of the project and/or beyond the time of its completion.

3.7 Business model validation and improvement (WP4)

Dataset name	Business model validation and improvement (INVITE_BMCValidation).
Type of study	A study blending both quantitative and qualitative elements based on an interview-based survey of users and stakeholders participating in the pilot activities of the Open Innovation Lab as well as members identified as potential early adopters and lead users from the Advisory Board of the project. The main target of the study is to validate and improve the business models developed for the Open Innovation Lab with a view to feeding the elaboration of the business plan that will guide its market rollout beyond the end of the project.
Dataset description	The dataset will be comprised of both qualitative as well as quantitative data regarding the different elements (e.g. revenue streams, value propositions, collaborators, etc.) of the business models designed for the Open Innovation Lab.
Methodologies for data collection / generation	Data will be collected via interviews utilising semi-structured questionnaires that will balance the flexibility of open-ended questions and the focus of restricted ones that can be easily quantified. The questionnaires will be administered either face-to-face or via electronic means of communication (e.g. web conferencing apps, telephone, etc.).
Storage and volume of the dataset	The dataset will be stored in a simple spreadsheet format (such as .xlsx or similar) which will be kept at the private server of Q-PLAN. The number of the records within the dataset will depend on the methodology which will be designed at a later stage of the project and utilised in the framework of the study. For security reasons, the dataset will be backed-up on a daily basis automatically from Q-PLAN's server in an external hard disk drive owned by the company and kept in its premises.
Metadata and standards	Descriptive metadata will be created to accompany the dataset (i.e. title, author and keywords).
For whom might the dataset be useful?	The dataset will be of great use to INVITE consortium members, enabling them to validate and refine the business models of the Open Innovation Lab, paving the way for its commercial exploitation after the completion of the project.
Confidentiality	In order to ensure the privacy of the study's participants as well as safeguard the project partners' commercial interests that may arise from the exploitation of the Open Innovation Lab, the dataset will remain confidential. The data will not be shared and/or reused outside the framework of the project and/or beyond the time of its completion.

3.8 Dissemination and communication results (WP6)

Dataset name	Dissemination and communication results (INVITE_Dissemination).
Type of study	Study aimed at monitoring and assessing the results of the different dissemination and communication activities of the project using an appropriate framework with quantitative metrics.
Dataset description	Data collected with a view to measuring and assessing the results of the project in terms of dissemination and communication with measurable metrics tailored to the different respective channels and tools employed as well as the activities deployed.
Methodologies for data collection / generation	Primary data will be collected through the dissemination activity reports of project partners regarding media products, project events, external events, general publicity, etc. Third party tools will be employed as well (e.g. Google analytics for metrics related to the website, social media statistics, information reported by the partners about published articles in media, outreach at third party events etc.). Only aggregated data will be collected in this dataset.
Storage and volume of the dataset	The dataset will be stored in standard spreadsheet format (.xlsx) and its size will evolve throughout the course of the project. Moreover, the dataset is being kept at the private server of the project's Dissemination Manager (E-UNLIMITED) and will be shared with the project partners for reporting and progress assessment purposes. It will be preserved for 5 years following the end of the project, before eventually being deleted and for security reasons, it will be backed up in an external hard-drive every week.
Metadata and standards	Descriptive metadata will be provided (such as title, type of data, data collection method and keywords).
For whom might the dataset be useful?	The dataset will be used by INVITE partners for reporting and progress assessment purposes. The data might also be used for dissemination purposes to promote achievements of the project.
Confidentiality	The raw data collected will not be shared and/or re-used (outside the framework of the project and/or beyond its completion). Still, some selected anonymous extracts might be included in project reports and deliverables which are public and therefore become publicly available. No personal information shall be included in the dataset and therefore no personal information shall be made public.

4. Conclusions and way forward

This initial version of INVITE's Data management Plan (DMP) has set the stage for the implementation of a sound data management methodology in the framework of INVITE. To this end, it has laid out the overall methodological principles to be followed by project partners in this respect with a view to making the data which they will collect, process and or generate as FAIR as possible, while also taking into account data security and ethical aspects.

Moreover, it has provided an initial, yet still meaningful overview of the valuable datasets that are expected to be created within the context of the project along with a description of each dataset emphasizing the methodology to be followed for their management during the lifespan of the project and beyond. The information provided in this respect will be enriched as the activities of the project progress to provide a more accurate description at later stages of INVITE's implementation.

Indeed, the DMP is a living document to be updated throughout the course of INVITE based on the latest developments and available project results. In fact, its initial version will serve as the basis for producing two additional, further elaborated versions, on M12 and M36 of the project, with a view to delivering an accurate, up-to-date and comprehensive data management plan before the completion of INVITE.

Annex

The following table constitutes the template that is employed in the framework of INVITE in order collect data management information from consortium partners.

Table 3: Template for collecting data management information per dataset

Dataset name	
Type of study	
Dataset description	
Methodologies for data collection / generation	
Format and volume of the dataset	
Metadata and standards	
For whom might the dataset be useful?	
Confidentiality	