



Innovative e-environments for Research on Cities and the Media

Deliverable

I.5 Data Management Plan

Deliverable Lead: CRB

Deliverable due date: 30/09/2016

Actual submission date: <<03/10/2016>>

Version: 1.3



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 693559.

Document Control Page

Title	DI.5 Data Management Plan
Creator	Cinémathèque Royale de Belgique / Davy Hanegreefs
Description	Data Management Plan lists all data that will be collected, processed and generated within the project and details all rules, regulations and provisions connected to handling this data.
Publisher	I-Media-Cities Consortium
Contributors	Davy Hanegreefs, Gabriella Scipione, Véronique Waterschoot, Nicola Mazzanti
Creation date	06/09/2016
Type	Report
Language	en-GB
Rights	copyright "I-Media-Cities Consortium"
Audience	<input checked="" type="checkbox"/> public <input type="checkbox"/> restricted
Review status	<input checked="" type="checkbox"/> Draft <input checked="" type="checkbox"/> WP leader accepted <input checked="" type="checkbox"/> Technical Manager accepted <input checked="" type="checkbox"/> Coordinator accepted
Action requested	<input type="checkbox"/> to be revised by Partners <input type="checkbox"/> for approval by the WP leader <input type="checkbox"/> for approval by the Technical Committee <input type="checkbox"/> for approval by the Project Coordinator <input checked="" type="checkbox"/> to deliver to European Commission
Requested deadline	

Table of Contents

Executive Summary	3
1. Introduction	4
2. Data of I-Media-Cities	5
3. Data Register	6
3.1. Data Register Component 1: Data Summary	7
3.2. Data Register component 2: FAIR data-principles	7
3.2.1. Making Data Findable	8
3.2.2. Making Data Openly Accessible	11
3.2.3. Making Data Interoperable	13
3.2.4. Increase Data Reuse	14
3.3. Data Register Component 3: Data Preservation and Archiving	15
4. Data Quality Assurance	16
4.1 Data Quality Assurance before Collection	16
4.2 Data Quality Assurance during Collection	16
4.3 Data Quality Assurance after Collection	16
5. Resourcing	17
6. Data Security	17
7. Ethical Aspects of Data Management Plan	19
8. Conclusion	20
Annex I. Data Register version 1 (09/20116)	21
Annex II: Terms of Use Template	31
Annex III. Privacy Policy template	38
Annex IV. Disclaimer template	40

Executive Summary

This Data Management Plan (DMP) describes all elements of the data management life cycle of I-Media-Cities. It presents the type of data will be collected, processed and generated by the project and how the I-Media-Cities Consortium plans to manage these datasets in the respect of the FAIR-principles of data management, as needed for a European project participating in Horizon 2020.

It provides information on the type of data the project will collect, process and generate, how this data will be handled during and after the project lifetime and what methods and standards will be applied to the data.

The DMP also includes all rules and provisions on data preservation, security and ethics for I-Media-Cities. These rules and provisions were already listed and in article 6 on IPR Management of D1.1 and D9.1 on the ethics of the project. This deliverable will detail the specific data protection and ethical rules that apply to the different datasets collected, analyzed and generated within the project.

Finally, a Data Register listing all details for all the open datasets, planned for in I-Media-Cities, is provided.

I. Introduction

This Data Management Plan (DMP) provides all the information needed to answer any question that might arise concerning the data collected, processed and generated within the project. There is a multitude of diverse research and user-generated data planned for, which makes such a concrete Data Management Plan an absolute necessity.

This document will list all the datasets that are going to be produced within the project, and will provide an analysis of all the main elements of the data management policy the project will implement to handle those datasets. I-Media-Cities is part of the OPEN Research Data Pilot, a flexible pilot running under Horizon 2020, whose goal it is to improve visibility, reach more people and have a greater impact – for science, society and your own career. Consequently, the DMP will provide information of how the research data of the project follows and adheres to the FAIR-principles of data management. It is also important to note that a key component of the project is that all datasets generated within the project will be made open, as long as there are not limitations due to copyright provisions or privacy laws.

The DMP also develops rules and general principles designed in D 1.1 (Quality Assurance, Risk Management and IPR) D.9.1 POPD – Requirement N° 4.

This Data Management Plan will provide information on the following elements of the Data Management lifecycle:

- Type of Data collected, processed and generated within the project
- The ways the project data complies with the FAIR-principles of data management
- Rules of preservation and archiving
- Ethical rules of data management
- Data Quality Assurance
- Resourcing of data management
- Data security provisions

2. Data of I-Media-Cities

When speaking of the data used, collected, processed and generated within I-Media-Cities, a distinction has to be made between the data and metadata existing before the project (*Background*) and the data and metadata created within the project (*Foreground*), as defined in D 1.1.

I-Media-Cities will use, collect, process and generate both types of data and metadata, but the IPR-related solutions for these two types differ greatly:

A. *The Background of the project*

The Background is any information, in hard copy or in electronic form, including, without limitation, documents, drawings, models, designs, data memoranda, tapes, records, and databases developed before or independent of performance under the project that is necessary for the performance of Project Work and exploitation of its results.

All rights and obligations connected to the background of the project are listed in the Consortium Agreement and in D1.1, article 6.3.¹ The Consortium will ensure that the project complies to these rights and obligations.

Concretely, the Consortium Agreement provides all the details on what constitutes Background for every partner and, to summarize, partners will retain ownership of the Background they contributed to the project.

B. *The Foreground of the project*

The Foreground means the results, including information, materials and knowledge, generated in a given project, whether or not they can be protected. It includes Intellectual property rights, similar forms of protection and unprotected know-how. Thus, foreground includes the tangible and intangible results of the project. Results generated outside the project scope do not constitute the Foreground.

Ownership and IPR of metadata, data and technological solutions created within the project are all dealt with in D1.1² and the Consortium Agreement³, and in general the following rules apply:

- Metadata produced within the project: to be made freely available, but the IP remains with the institution who produced it.
- User-generated content and metadata: IP will be regulated when users register to any I-Media-Cities service or platform.

¹ Nicola Mazzanti, Davy Hanegreefs, Veronique Waterschoot (2016), IMC-D1.1-Quality Control, Risk Management and IPR-v. 1.3-final, I-Media-Cities Consortium, p.36, <https://imediacities.eu/results/>

² Nicola Mazzanti, Davy Hanegreefs, Veronique Waterschoot (2016), IMC-D1.1-Quality Control, Risk Management and IPR-v. 1.3-final, I-Media-Cities Consortium, p.36, <https://imediacities.eu/results/>

³ I-Media-Cities Consortium, IMC-Consortium Agreement-final

- Technological solutions created within the project: IP and ownership remain with the partner(s) who developed it. If there are several partners who developed the technological solution, a shared ownership will be granted.

3. Data Register

Since I-Media-Cities generates a lot of datasets in order to achieve the project's goals, we will classify and detail them in a Data Register, to enable their proper and full management. A Data register lists all datasets of a project and will include their methodology, standards and handling of all data during and after the end of the project. This Data Register has to be understood as a living document, which will be updated regularly during the project lifetime. This article lists and details the possible values of the different datasets and to provide for some context to go along with the Data register.

Below, a preliminary list of all generated datasets of version 1 of the Data Register, completed with their names and the Work Package they are attached to.

Data generated within all Work Packages

- Deliverables that are reports and documents

Every Work Package has a number of deliverables attached to it. Some of these deliverables are reports or documents that need to be considered research data and are vital to understanding the research project as a whole. To this effect, the project will provide information on all documents that are relevant to the research project of I-media-Cities in the Data register. All these documents will be kept at the same place and the details of their data management can be found in the data register in Annex 1.

Datasets generated within WP3

- Performance data of website
- List of email subscribers for I-Media-Cities Newsletter

Datasets generated within WP5

- Account details for end-users
- User-generated metadata
- User-generated data
- Software

Datasets generated within WP6

- I-Media-Cities Data model
- Additional Metadata created for the I-Media-Cities Metadata Repository
- Software

Datasets generated within WP7

- Metadata generated through the use of automatic analysis tools
- Software

Datasets generated within WP8

- Software

See Annex I for the I-Media-Cities Data Register, version 1.

3.1. Data Register Component 1: Data Summary

To keep a structured overview, the first component of every data set description in the Data Register will be a Data Summary. It will deliver the following information in accordance with the guidelines on fair data management in Horizon 2020 projects⁴:

- Data set reference and name: Identifier for the data set.
- Data set purpose: Description of the data that will be collected and generated and its purpose for the project and its objectives.
- Standards and metadata: Specify the types and formats of data generated/collected.
- Data Origin: Specify the origin of the data and whether or not existing data is being re-used.
- Data size: State the expected size of the data (if known).
- Data utility: Specify to whom the data will be useful.

3.2. Data Register component 2: FAIR data-principles

In accordance with the requirements for European H2020-projects, I-Media-Cities will make sure the research generated within the project adheres to the FAIR-principles, which means that there are clear rules on their management that will improve their findability, accessibility, interoperability and reusability.

This Article will go into more detail for each of the different components that help define the FAIR-principles of I-Media-Cities. It will provide further information on all the principles' components and will deal with the questions that need answering in order to comply to the FAIR DATA Guidelines.⁵

⁴ European Commission, Directorate-General for Research and Innovation, H2020 Programme Guidelines on FAIR Data Management in Horizon 2020, version 3.0, 26 July 2016, p. 7

⁵ European Commission, Directorate-General for Research and Innovation, H2020 Programme Guidelines on FAIR Data Management in Horizon 2020, version 3.0, 26 July 2016, p. 7-8

3.2.1. Making Data Findable

To make the collected and generated data of I-Media-Cities findable, every dataset will have specifications on the ways they are made findable through metadata, versioning and other identification mechanisms.

Next to these identification methods, every dataset will also provide the information on the main digital location, used for their storage. To help make the open data of I-Media-Cities discoverable, the project website will include a central catalogue of all open datasets, which will include rules and regulations for accessing these datasets.

In general, the following rules and provisions apply for the different datasets:

a. Reports and other documents

- In accordance with the mandatory properties of the DataCite metadata scheme⁶, which is a part of the DDI-standard for research metadata (version 3)⁷, all reports and documents collected or generated in I-Media-Cities will, at a minimum, have the following metadata attached to it: Persistent and unique identifier (file name), Title, Keywords/tags, Author/Creator, Subject, Time of Publication, Publisher (Always I-Media-Cities Consortium), Version, File size.
- All metadata will be added to the properties of the document itself. For deliverables, next to the metadata being added to the properties, some of the metadata will also be published on the document control and title pages.
- The file name is the unique identifier I-Media-Cities uses for reports and other documents. It is constructed by the following standard identification mechanisms:

For deliverables:

Start with IMC (Project code for I-Media-Cities), followed by the number of the deliverable (can be found in list of deliverables in the Grant Agreement), followed by the title of the deliverable or annex to the deliverable and the version of the document. All these elements are connected by dashes.

Example: IMC-D7.1-Report on state of the art for moving image analysis-v. 1.3-final

For documents that are not deliverables:

Start with IMC, followed by the type of the document (e.g. questionnaire, minutes ...) and the title and the version of the document. All these elements are connected by dashes.

⁶ Joan Starr (2011), DataCite Metadata Kernel, Mappings <http://www.dcc.ac.uk/resources/metadata-standards/ddi-data-documentation-initiative>

⁷ <http://www.dcc.ac.uk/resources/metadata-standards/ddi-data-documentation-initiative>

Example: IMC-Minutes-Kick-off meeting Brussels May 23-24-version 1.2-final

- Tags and keywords will be added to the properties of all documents and will always include the following: I-Media-Cities, Horizon2020, H2020-REFLECTIVE-6-2015-69359.
- Clear Versioning is provided for all documents. All deliverables will have a version number on the title page and all other documents will have a version number in the file name. The last version of a document will always carry the label 'final' at the end of the file name, after the version number of the document.

The final versions of documents and deliverables will be added to the project website, www.imediacities.eu. All the previous versions will be put on the project Work Cloud, B2DROP⁸. The project partners have unlimited access to all versions of all documents and reports.

b. Films, Photographs and text-files (contents) created within the project

User-generated contents, which means any film, photograph and document, created by partners, researchers and the general public, using the I-Media-City services, and then saved in the I-Media-Cities repository. The user-generated content does not belong to the Background of the project.

All rights and obligations concerning user-generated contents will be listed and detailed in the IPR-provisions connected to every single I-Media-City service and must be approved by users, before they are able to save their data in the I-Media-City repository.

Every Film, Photograph and text-file created within I-media-Cities will have a title and a source (metadata elements) attached to them, when they are entered into the I-Media-Cities repository. Once entered into the repository, they will receive a unique numerical identifier, dedicated to the I-media-Cities repository.

c. Metadata collected, processed or created within the project

- The following types of metadata will be collected and in part generated within I-Media-Cities:
 - Technical Metadata: these contain information on the technical aspects of the films, photographs and documents, such as size, format, color ...
 - Descriptive Metadata: these contain information on the subject of the films, photographs and documents, such as title, director/author, summary, keywords ...
 - Administrative Metadata: these contain information on the rights status, creation, preservation and quality control of the films, photographs and documents, such as source/publisher, rights holder.

⁸ <https://b2drop.eudat.eu/login>

- All metadata connected to the films, photographs and text-files (Background or user-generated) of I-Media-Cities, which is created within the project, will be attached to a linked and open metadata model. Since a specific metadata model and repository will be created for I-Media-Cities, the structure and standards used for this model will be detailed and described in D6.1 *Content metadata subsets selection and analysis of metadata schemas and vocabularies* (M 10).
- All metadata connected to the films, photographs and documents (background or user-generated) of I-Media-Cities, can derive from four sources:
 - Previously existing metadata in the databases of the Project partners of I-Media-Cities (this metadata will be considered part of the Background, since it is collected not created)
 - Automatic A/V Analysis Tools
 - Metadata generated by the software
 - Users (both project partners and external users)

Rules of ownership and IPR management of all metadata collected, processed and generated within I-Media-Cities, will be subject to the regulations and provisions detailed in D1.1 Quality Control, Risk Management and IPR, already published on the project website⁹.

The approach I-Media-Cities will follow to maintain the distinction between Metadata that is considered background (already present in the partners' databases) and the metadata generated within and for the purpose of the project, is based upon the definition of the generating source of the metadata.

d. Software and its intellectual property rights (Patents, copyrights and Trademarks)

- All software developed in I-Media-Cities will be governed by the rights and obligations attributed to the results and the software of the project, as listed and detailed in D1.1¹⁰ and the Consortium Agreement¹¹.
- The rights owners of any software developed in I-Media-Cities, will clearly be indicated in the copyright information of the software.

e. Personal data

- Personal data means any information relating to an identified or identifiable natural person.
- An email address will function as the unique identifier for personal data.

⁹ Nicola Mazzanti, Davy Hanegreefs, Veronique Waterschoot (2016), IMC-D1.1-Quality Control, Risk Management and IPR-v. 1.3-final, I-Media-Cities Consortium, article 6, p.36 – 41, <https://imediacities.eu/results/>

¹⁰ Nicola Mazzanti, Davy Hanegreefs, Veronique Waterschoot (2016), IMC-D1.1-Quality Control, Risk Management and IPR-v. 1.3-final, I-Media-Cities Consortium, articles 6.3, 6.4 and 6.5, p36 – 38, <https://imediacities.eu/results/>

¹¹ I-Media-Cities Consortium, IMC-Consortium Agreement-final, article 9.8, p.18-20

- Users can be asked to register an account on I-Media-Cities in order to be able to submit data into the project. When creating their account, users will also have to approve the Terms of Use, the Privacy Policy and the Disclaimer attached to this type of account.
- All users can always have access to, edit or even cancel their own account and personal data, by logging in via their email address and their personal password.
- All rules and regulations connected to personal data, included informed consent for data sharing and long term preservation, are listed and detailed in deliverable 9.1¹² on ethics and in the IPR section of deliverable 1.1¹³.

3.2.2. Making Data Openly Accessible

Openness of data and free access is one of the key concepts of the project. As much as possible, and respecting all Intellectual Property involved, but also taking into considerations the principles of the Public Sector Directive 2003/98/EC, all data will be made openly and freely accessible. All data concerned are: content, metadata (including user-generated), and research produced within the project.

To that avail, all publications, metadata and other content that are produced as a direct result of the project will be made openly available, which is known as a “gold” model. The project will also encourage researchers using its resources, to make their works (or a significant section of it) available via and within the e-environment and eco-system of I-Media-Cities. In other words, research that is produced by using and accessing the content made available via I-Media-Cities, will be made available via the project e-environments.

For every dataset, the following specifications on their accessibility will be detailed in the Data Register:

- a. Is the dataset open available or not.

Open data is defined as follows:

Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and share-alike.¹⁴

As a general rule, the project will openly make available the contents and metadata produced within the project through a “gold” model. This is due to the fact that the consortium strongly wants to encourage other stakeholders in using the material generated by the project, enhancing the replicability of the proposed approach and wider fruition of the technology.

Some datasets generated in I-Media-Cities will not be openly accessible, depending on the copyright protection rules and ethics rules attached to the generated datasets. The

¹² Nicola Mazzanti, Davy Hanegreefs, Veronique Waterschoot (2016), IMC-D9.1- POPD Requirement No.4-v. 1.0-final, I-Media-Cities Consortium. D9.1 is not a public deliverable, but information on the ethics rules on personal data can be ascertained by contacting the project via email on I-Media-Cities@cinematek.be.

¹³ Nicola Mazzanti, Davy Hanegreefs, Veronique Waterschoot (2016), IMC-D1.1-Quality Control, Risk Management an IPR-v. 1.3-final, I-Media-Cities Consortium, article 6, p36 – 41, <https://imediacities.eu/results/>

¹⁴ <http://opendefinition.org/>

accessibility status is listed in the data register and will always be made clear to users before they can start generating datasets.

b. How will the dataset be made available?

An open data license will be applied to all datasets that are openly available. This license will list and determine all intellectual property rights that exist on the datasets, and is conformant with the principles set forth in the Open Definition above. If user-generated data or content need to be anonymized in order to be able to make it openly available, I-Media-Cities will do so.

Open data needs to be technically open as well as legally open. To this end, I-Media-Cities will make sure that all open data is available in bulk in a machine-readable format or through an API and for a reasonable price. The project will also verify which part of the dataset is available for download.

Data will be priced at no more than a reasonable cost for reproduction or even as a free download from the Internet. The open data license will mention the cost of the dataset or will provide instructions on how and where information on the cost for downloading the dataset can be found.

The project website will have clear instructions on how to access the different open datasets and will try to avoid a Data Request Methodology as much as possible. If such a methodology is needed, the project website will also list the full details instructions for this DRM.

c. Methods and software necessary to access the dataset

The methods used to share data will be dependent on a number of factors, such as the type, size, complexity and sensitivity of the dataset.

For easy data availability, the project will use one of the following methods of online publication to share data:

- Files will be provided for download via the project website
- Via ftp servers
- As an API
- Linked metadata structure
- Via 3rd party websites

If the project uses specific software to share the data, the partners will make sure to primarily use open source software. If dedicated software, which is not open source, is necessary to access the data, the software will have clear indicators on its copyright situation, rules of use and will provide a manual for users.

d. The location of the deposited datasets

All open datasets will be available through the project website. The user-generated metadata and content will be deposited in the dedicated I-Media-Cities repositories, which will be

developed as a part of the project. The project website will provide information on how to access these datasets and the used methodology.

The repositories will be hosted at CINECA partner location, in Italy. CINECA will provide access for users to storage resources in the amount agreed with IMC partners. Consequently I-Media-Cities users will be allowed to access the iRODS data repository. Actually two iRODS server instances are installed and running on CINECA PICO HPC cluster¹⁵, intended to enable new "BigData" classes of applications, related to the management and processing of large quantities of data, coming both from simulations and experiments. The storage area is composed by high throughput disks for a total amount of about 4 PB, connected with a large capacity tape library for a total actual amount of 12 PByte (expandable to 16 PByte). Also Virtual Machines have been deployed and equipped with an iRODS server development version in order to perform advanced tests on the infrastructure before going on production.

e. Are there any restrictions on the datasets

If any restrictions apply to a generated dataset, for whatever reason (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related), the Data Register will indicate this restriction and will provide information on the accessibility of the dataset and the request methodology users must follow to gain access. This information will also be made available through the project website.

3.2.3. Making Data Interoperable

Metadata interoperability is fundamental for the project. Interoperability denotes the ability of diverse systems and organizations to work together (inter-operate). In this case, it is the ability to interoperate - or intermix - different datasets.¹⁶ In order to make the data and metadata of I-Media-Cities interoperable with other datasets, the following specifications will be added and detailed to every dataset in the Data Register:

- The open data license applied to the open datasets will allow for the intermixing of the dataset with other open datasets, even if these are external to I-Media-Cities.
- A clear indication on which (meta)data standard was used for the dataset. The project will make largely use of open standards in the case of A/V content, the main metadata standard that is relevant is the one developed following a standardization mandate from the European Commission to the European Committee of Standardization (CEN). CEN produced two European Standards (EN) to facilitate the interoperability of film databases:
 - EN 15744:2009 "Film Identification – Minimum metadata set for cinematographic works";
 - EN 15907:2010 "Film Identification – Enhancing interoperability of metadata – Element sets and structures".

¹⁵ <http://www.hpc.cineca.it/hardware/pico>

¹⁶ <http://opendatahandbook.org/guide/en/what-is-open-data/>

These standards will constitute the base for the activities of the project.

- In order to open the structured information gathered in the project, Open Linked Data will be adopted so that it can be interlinked and become more useful through semantic queries. This type of data builds upon standard Web technologies, such as HTTP, RDF and URI's.
- Specifically for the identification of the A/V contents, the project will foster the adoption of Persistent Identifiers in the project and among the different cultural holding institutions. Among the main mature standards the DOI¹⁷ and the Handles¹⁸ are effective means of identifying an entity over the internet.
- Specifically for the identification of names, the project will foster the adoption of the main mature standards such as ORCID¹⁹ that provide a persistent digital identifier that distinguishes researchers or ISNI²⁰.
- Since I-Media-Cities will create its own metadata ontologies and vocabularies, details on how they are linked to commonly used metadata standards will be provided on the project website. I-Media-Cities strives to use a linked open data model for its database.

3.2.4. Increase Data Reuse

In order to increase the re-use of the open datasets of I-Media-Cities, the project partners will use the dissemination and communication channels to spread the word on the availability of the open datasets:

- I-Media-Cities will promote their available open datasets through the project website and the other available dissemination tools, described in deliverable 3.3, the dissemination and communication plan of I-Media-Cities²¹.
- I-Media-Cities will make every effort to communicate the availability of its open datasets to channels that directly provide information to one or more of the 6 target groups and related projects, as described in D3.3²².

In order to encourage re-use, the special section on the project website that will deal with all the open datasets, will include the following information:

- Catalogue with all open datasets of the project.
- All information on the IPR rules and the open data licenses used for the different datasets. As a general rule, the project will apply the licenses that permit the widest re-use possible.
- All information on the methodology and location needed to access the open datasets.

¹⁷ ISO 26324:2012 "Information and documentation – Digital object identifier system"

¹⁸ <https://www.handle.net/>

¹⁹ <http://orcid.org/>

²⁰ <http://www.isni.org/>

²¹ Nicola Mazzanti, Davy Hanegreefs, Veronique Waterschoot (2016), IMC-D3.3-Dissemination and Communication Plan-v. 1.2-final, I-Media-Cities Consortium, articles 3&4, p.12-19, <https://imediacities.eu/results/>

²² Nicola Mazzanti, Davy Hanegreefs, Veronique Waterschoot (2016), IMC-D3.3-Dissemination and Communication Plan-v. 1.2-final, I-Media-Cities Consortium, article 2, p.7-11, <https://imediacities.eu/results/>

- A clear indication on how long the datasets will remain re-usable.
- A clear contact point for any questions related to the access and re-use of the project datasets.
- If there are any expected difficulties or restrictions in data sharing and re-use, the open data information page will provide this information, along with causes and possible measures to overcome these. Possible strategies to limit restrictions may include: anonymizing or aggregating data; gaining participant consent for data sharing; gaining copyright permissions; and agreeing a limited embargo period.

A key component of the project is the openness of the solution developed. Although some components might involve some IP and require protection, the project is conceived in order to provide a solution that is open and easy to adopt by as many institutions as possible, in as many fields as possible. Concretely, I-Media-Cities will actively seek to involve other research centers and collection-holding institutions to use the tools it will produce, already during the project lifetime and after. The system will be designed to enable more archives, cities and institutions to join the project. Equally encouraged by the partnership is the fact that the concept and the underlying technological solutions are employed by others, on completely different subject matters.

In order to enhance the chance that the open datasets will be re-used, they will also be published on 3rd party data providing websites, such as DataHub²³. Also thanks to the open link to Europeana via the EFG content aggregator, content and data will also be – as much as possible and compatible with the Europeana metadata policies – made available via that channel. This link will be drafted considering the evolution of Europeana portal and its relation with aggregators as EFG, which are now under reflection following Council conclusions²⁴.

Lastly, a key objective of the project is the ability to process queries across language barriers. This will be made possible by defining and implementing all necessary tools, which will be defined in D6.1, such as ontologies, thesauri and authority files. This part is considered as strategic, since I-Media-Cities will function as a pilot and a blueprint to solve this standing roadblock to effective sharing of collections and metadata across European heritage and research institutions. Therefore, it is a strategic objective of the project that this part of the development and the activities is completely open in its concept, results, methods, technical solutions and results.

3.3. Data Register Component 3: Data Preservation and Archiving

The long-term preservation plan for every I-Media-Cities dataset will also require the following information to be added to every dataset of the Data Register:

²³ <https://datahub.io/>

²⁴ <http://data.consilium.europa.eu/doc/document/ST-9643-2016-INIT/en/pdf> - The role of Europeana for the digital access, visibility and use of European cultural heritage - Council conclusions (31 May 2016)

- The retention period of every dataset in the dedicated (Meta)data repositories at CINECA. This period will depend on the type of data and will always be in compliance with the rules set out in D9.1 on ethics for I-Media-Cities and D1.1 on the IPR management plan for the project, but will also be influenced by the results reached in D3.1 Platform Sustainability report with business model scenarios.
- The central preservation location of every dataset will be defined. This will almost always be one of the two repositories foreseen in the project system.
- Clear indications about what will happen with the dataset after the planned retention period of the dataset is finished.

4. Data Quality Assurance

I-Media Cities will apply a Data Quality Assurance methodology to avoid data contamination, incorrect, inaccurate or not recorded data and to generally prevent errors entering a dataset. The Data Quality Plan will apply quality assurance rules to the data before, during and after collection.

4.1 Data Quality Assurance before Collection

To avoid errors, I-media-Cities will strictly define and enforce the format and metadata standards used for the different datasets. The technical solutions will be designed in such a way that automatically enforces the use of the defined standards and formats.

See D6.1 for the I-Media-Cities metadata standards definition.

4.2 Data Quality Assurance during Collection

To ensure user-generated data quality, a specific methodology will be applied to eliminate mistakes during data entry and evaluation. This methodology will be defined in D4.1 on Research Activities and Platform Content Enrichment.

By designing the data and metadata repositories in a specific way, before starting to use them, the project will also enhance the data quality during the collection process. In order to reach this goal, the following rules will be applied during the design phase:

- The use of consistent terminology
- Data is atomized
- Guide data entry through field value definition

See D6.2 on the Metadata repository and D8.1 for the design of the system architecture.

4.3 Data Quality Assurance after Collection

Statistical data could be generally provided by CINECA exploiting statistical tools as for example Elastic Search or Kibana, first producing logs containing information about database content, subsequently parsing log files in order to produce statistical visual representation of such information.

5. Resourcing

Since the design and use of the meta(data) repositories, the preparation of the project data and the application of the FAIR-principles to the datasets were already foreseen during the composition of the project proposal, no extra resources should be needed to fulfill these requirements.

The relevant project partners will make sure that no extra resources are needed to ensure preparation and preservation of the datasets during the lifetime of the project. If there are any costs or charges attached to the preservation of the datasets after the lifetime of the project, these will be planned for and added to the business model scenarios that will be developed through the living lab methodology.

If any unforeseen or unplanned resources are needed to complete and deliver any part of the project plan, the problem resolution mechanism, defined in the Risk Management Plan²⁵, will be started. This means, that the relevant WP Leader will inform the Project Coordinator of the problem, and will analyze and try to manage the problem inside the related project team. In the case the insufficiency of resources can't be resolved or entails risks for the whole project, its solution is referred to the Project Coordinator and Management Committee. The Project Coordinator will discuss the problem at the first Management Committee following their discovery and will devise a solution proposal that must be approved by the Management Committee.

6. Data Security

All the data is stored on secured servers at CINECA. In order to ensure the continuity or the uninterrupted provisioning of the IIA service (the research service at CINECA), CINECA exploits Preventive, Detective and Corrective measures in order to reduce or eliminate various threats/incidents that may occur. The control measures are applied at different levels of CINECA's infrastructure, main control measures adopted include: Uninterruptible power supply (UPS), Server Room conditioning and cooling and different types of connectivity, including firewall and VPN services.

All the equipment is in redundant configuration.

²⁵ Nicola Mazzanti, Davy Hanegreefs, Veronique Waterschoot (2016), IMC-D1.1-Quality Control, Risk Management an IPR-v. 1.3-final, I-Media-Cities Consortium, article 5.4, p31-36, <https://imediacities.eu/results/>

HA clustering performs hardware/software faults detecting, and immediately restarts the application on another system node, making the fault transparent vs system users.

As EUDAT²⁶ partner, CINECA will support access to the CINECA data repository, also providing the web portal with authentication mechanisms implemented by the EUDAT consortium. Technologies developed within EUDAT will be exploited and coupled with novel web tools in order to expose interfaces to allow users browsing and visualizing files from the CINECA repository.

Proper APIs will be included in the IMC web portal interface to activate data access using, for example, EUDAT federated identity management services (i.e. B2ACCESS) or invoking iRODS procedures in order to directly connect to the CINECA iRODS server instance.

Data access and authentication within I-Media-Cities infrastructure will be implemented as a novel Account Service, managing system accounts through the I-Media-Cities web portal interface and mapping I-Media-Cities users to iRODS users so as to maintain the correspondence between accounts on the system in case they pertain to the same physical person.

In EUDAT, access policies for data deposited directly in the Common Data Infrastructure (CDI) through user-facing services (for instance B2SHARE) are determined by the depositor. While the EUDAT CDI encourages open access, depositors can currently choose to keep data private. Consequently access policies for I-Media-Cities data stored in the CDI by a partner community will be determined by CINECA data managers as part of the community and subsequently respected across all copies maintained in the CDI.

Moreover the EUDAT2020 project in the future will develop finer-grained authorisation for data sharing, enabling "group" or "approved reader" authorisation modes. Within the CDI framework, authorisation decisions (the granting of access rights) will remain with the relevant data owners.

The EUDAT approach in terms of Access policies could obviously impact on the I-Media-Cities Security Service architecture and development, even if no specific high-level requirements has been already identified in order to determine how access should be granted to a specific resources within the I-Media-Cities infrastructure.

To make sure no data will be lost, a data recovery plan generally provides basic replication mechanisms conceived in order to preserve data from failures or in case of system takeovers. This kind of approach includes file system backup as well as database support for replication. Upon such kind of basic and very general mechanisms, currently supported from the most part of HPC centres, CINECA offer includes more sophisticated high level mechanisms for data preservation and metadata management. CINECA-"Backup Servers" is an important control measure adopted in order to maintain a backup of the data contained at application and database servers as well as file system servers for both production and staging environment. The Advanced Data Backup mechanism enables recovery of the pertinent service from incident/disasters that cause a complete loss of the data.

²⁶ <https://eudat.eu/>

CINECA also supports high-level data and metadata management tools relying upon services developed within the EUDAT consortium and offers high level support for storage and data preservation. To this extent individual services in the EUDAT CDI provide shorter or longer periods of data preservation depending on Service Level Agreements (SLA). The B2DROP service (which could be easily integrated within the IMC infrastructure) is aimed at short-term sharing of active research data, B2SHARE is aimed at the medium-term storage and sharing of near-publication and published data and B2SAFE is aimed at longer-term preservation and backup of community data repositories.

Backup of data in the EUDAT CDI is performed at two levels: multiple replicas of data can be stored at different sites (also on different geographical locations) through the B2SAFE service, governed by user-defined policies; data may additionally be backed up at an individual site, depending on the nature of the storage systems in use at that site.

I-Media-Cities will apply extra security measures when dealing with the transfer of sensitive data. If such data needs to be transferred, the project will use a File Transfer protocol combined with an SSL Certificate to add an extra layer of security to the transfer of files. Encrypting the data this way is one of the safest ways to make sure that sensitive data is protected during transfer. In fact, CINECA provides file transfer mechanism based on GridFTP which supports x509 proxy certificates as well as SSL encryption. GridFTP is a very effective tool for data transfer: it enhances the standard ftp service making it more reliable and faster. The transfer performance with GridFTP may be easily raised of 4-5 times the ones provided by scp-like tools. GridFTP was developed by the Globus alliance as part of an open-source toolkit for data management. It is a client-server application: on CINECA HPC platforms a server is available.

7. Ethical Aspects of Data Management Plan

All ethics and legal issues connected to data sharing in I-Media-Cities are covered in D9.1 POD Requirement N° 4²⁷ and the IPR-section of D1.1²⁸.

In this Article, we provide a summary of all rules, regulations and provisions planned for to deal with any ethical challenge that might arise with data sharing in I-Media-Cities.

- When dealing with long term preservation or sharing of personal data of any kind, the project will ask the user for informed consent. This informed consent will be achieved through requesting the user to approve the Terms of Use, the Privacy Policy and the Disclaimer, in order to be able to register his account and submit any kind of personal data into the platform or any of its web services. These documents will be added to every web service of I-media-Cities and might differ from each other to fit the specific technical abilities of a specific web service. The Consortium will make sure that all web services will have a clearly marked and dedicated area where users can, at all times during their visit, find and read the terms and policies that apply.

²⁷ Nicola Mazzanti, Davy Hanegreefs, Veronique Waterschoot (2016), IMC-D9.1-POD Requirement N°4-v. 3.0-final, I-Media-Cities Consortium

²⁸ Nicola Mazzanti, Davy Hanegreefs, Veronique Waterschoot (2016), IMC-D1.1-Quality Control, Risk Management and IPR-v. 1.3-final, I-Media-Cities Consortium, article 6, p32-38, <https://imediacities.eu/results/>

Templates for the Terms of Use, the Disclaimer and the Privacy Policy of I-Media-Cities have been added as annexes to this document.

See Annex II for the Terms of Use template

See Annex III for the Privacy Policy template

See Annex IV for the Disclaimer template

- The project will make sure that all legislative standards and acts of ethics are followed when dealing with collecting, processing and sharing data in the project.
- The Terms of Use and the Privacy Policy, attached to any part of the project where data is collected from users, will inform those users on how their data will be used, who it will be shared with and the details of its storage.
- If data needs to be anonymized before it can be made open, the project will do so.

8. Conclusion

This Data Management Plan provides an overview of all aspects of collecting, processing and generating data within I-Media-Cities, as well as all the related challenges and constraints that need to be taken into consideration.

I-Media-Cities has constructed a Data Register that lists all datasets collected, processed and/or generated within the project and details all relevant data management information for each of them. Every dataset has a summary that indicates, among other things, the unique identifiers of the dataset. Since I-Media-Cities is part of the H2020 project and of its Open Research Data Pilot, the Data Register will include the information of how every dataset complies to the FAIR-principles of data management, as well the details on its storage and preservation.

The last three sections of this DMP provides information on the ways data quality is assured throughout and after the project lifetime, the security provisions that protect the datasets, the ethical rules that apply and the way the resources are planned for all parts of data management throughout the project.

A first version of the Data register with all datasets foreseen in I-Media-Cities, containing the detailed data management information for each dataset, has been created and added as an annex to this document. Since the Data Register is a living document, it will be updated throughout the project lifetime, to include previously unforeseen datasets.

Templates for the Terms of Use, Privacy Policy and Disclaimer have also been provided. These documents will be adapted to every web service of I-media-Cities in order to collect the informed consent of users.

Annex I. Data Register version I (09/20116)

1. Documents (deliverables that are reports, reports, other documents)	
Data set reference and name	Every document will have a unique identifier which is constructed following the rules drawn up in Article 3.2.1 of the Data Management Plan.
Data set purpose	Documents that are needed to understand the different aspects of the research aspect of I-media-Cities
Data Origin	Created by project partners of I-media-Cities and responsible publisher is always I-Media-Cities Consortium
Data Size	No document will exceed size of 104Mb
Data Utility	The documents will be useful for the other partners in I-Media-Cities, external people trying to understand the research component of I-media-Cities
Data Findable: Standard and metadata	The used standards for documents are : pdf, Microsoft Word and Excel
Data openly accessible: Open or not, IPR, method of access, location of data, restriction of access	<p>Most documents are openly accessible. Some deliverables are considered for Consortium use only, however, and cannot be accessed by anyone outside of the project.</p> <p>The IPR of all documents remains with the I-Media-City consortium</p> <p>Method of access is download from the I-Media-City website, in the public 'results' section. The deliverables and documents with limited access outside the project, will be stored in the password protected partners-only section of the website.</p>
Data Interoperability Data license, Standards used	Not applicable for documents
Data Reuse Published on 3 rd party website	These documents will not be published on a third party website.
Data preservation and archiving Retention period, after retention period, central preservation location	All documents will be kept for a period of five years after the end of the project lifetime. During the project lifetime they will be kept on the project website www.imediacities.eu and after the project lifetime, they will be stored in a dedicated location on the servers at technical partner CINECA in Italy. After the retention period ends, the documents will only be stored in hard copy at a dedicated location at the Coordinator

2. Performance data of the website	
Data set reference and name	Every version of the performance data report will have a unique identifier which will be constructed using the naming procedures listed at 3.2.1
Data set purpose	Performance data is useful for the project partners and the project coordinator to understand the usability of the website and the central portal of the project and make potential adjustments
Data Origin	Created by project partners of I-media-Cities and responsible publisher is always I-Media-Cities Consortium
Data Size	No report will exceed size of 104Mb
Data Utility	The documents will be useful for the other partners in I-Media-Cities to understand the use of the website
Data Findable: Standard and metadata	The used standards for documents are : pdf, Microsoft Word and Excel
Data openly accessible: Open or not, IPR, method of access, location of data, restriction of access	The performance review reports are not openly available, since they are considered internal data for the consortium and might provide classified information on the use and visitors of the project website(s), the performance review reports are kept on the partners-only portion of the website.
Data Interoperability Data license, Standards used	Not applicable
Data Reuse Published on 3 rd party website	They are not published on a 3 rd party website
Data preservation and archiving Retention period, after retention period, central preservation location	All review reports will be kept on the project website www.imediacities.eu . Since they are only applicable to the website as long as the website is in use, they will be kept for the entire duration of the website they are referring to and on a dedicated location on that website.

3. Email subscribers for I-Media-Cities communication channels	
Data set reference and name	Any personal data used for communication in I-Media-Cities will be connected to a unique identifier, which is always an email address.
Data set purpose	This data set is only used for the communication purposes of the project. Any purpose will be pre-defined in the Informed Consent form.
Data Origin	Information offered by individual persons to the projects for specific purposes.
Data Size	All personal information will be kept at the minimum needed by the project to be able to purposefully communicate with the person
Data Utility	This data is only useful for the communication purposes of the project
Data Findable: Standard and metadata	The only standard used is the standard that constitutes an email address.
Data openly accessible: Open or not, IPR, method of access, location of data, restriction of access	<p>This data is not openly accessible. Rules and regulations of IPR are regulated in the informed consent form, which has to be approved by the individual before his or her personal data can be collected. All personal data used for communication in I-Media-Cities will be stored in the open source communication platform and database at www.mailchimp.com.</p> <p>Access to any personal data, used for communication, is restricted to the owner of the data and the people tasked with communication for the project. These project partners are limited to access to the data as defined in the informed consent and the Ethical rules attached to I-media-Cities (D9.1)</p>
Data Interoperability Data license, Standards used	The standard construction of an email address allows for certain interoperability if the project decides to change tools of communication.
Data Reuse Published on 3 rd party website	Any rules regarding the re-use of the data will be defined in the informed consent form.
Data preservation and archiving Retention period, after retention period, central preservation location	<p>Any personal data will only be kept for as long as it is necessary to perform the function it was collected for. As soon as the function has been fulfilled, all personal data will be destroyed by the Consortium according to the legal rules, set out for such destruction.</p> <p>During the time the data is in use at the project, it will be stored at the servers at www.mailchimp.com.</p>

4. Personal account data for I-Media-Cities	
Data set reference and name	Any personal data collected for use in I-Media-Cities, will be connected to a unique identifier, which is always an email address.
Data set purpose	Any purpose of this dataset will be pre-defined in the Informed Consent form.
Data Origin	Information offered by individual persons to the projects for specific purposes.
Data Size	All personal information will be kept at the minimum needed for specific use in the project
Data Utility	This data is only useful for certain functionalities of the project software solutions and platform
Data Findable: Standard and metadata	The only standard used is the standard that constitutes an email address.
Data openly accessible: Open or not, IPR, method of access, location of data, restriction of access	<p>This data is not openly accessible. Rules and regulations of IPR are regulated in the informed consent form, which has to be approved by the individual before his or her personal data can be collected. All personal account data in I-Media-Cities, will be stored in a database at Cineca.</p> <p>The Consortium's access and allowed use of the data is defined in the informed consent and the Ethical rules attached to I-media-Cities (D9.1)</p>
Data Interoperability Data license, Standards used	The standard construction of an email address allows for certain interoperability if the project decides to change tools of communication.
Data Reuse Published on 3 rd party website	Any rules regarding the re-use of the data will be defined in the informed consent form.
Data preservation and archiving Retention period, after retention period, central preservation location	<p>Any personal data will only be kept for as long as it is necessary to perform the function it was collected for. As soon as the function has been fulfilled, all personal data will be destroyed by the Consortium according to the legal rules, set out for such destruction.</p> <p>During the time the data is in use at the project, it will be stored at the servers at CINECA.</p>

5. User generated metadata	
Data set reference and name	Each set of metadata is related to the content/fragment to which it refers to. The ID will be generated by the system and the structure of the ID will be defined in D6.1.
Data set purpose	Metadata that is generated by users for the purpose of enriching the metadata set accompanying any object of content in the project.
Data Origin	Created by users of the I-Media-Cities platforms and solutions and saved by them in the I-Media-Cities Metadata repository
Data Size	Each type of user generated metadata will have its own size. The specific details will be defined in the data model foreseen in D6.1
Data Utility	This dataset will be useful for the archives participating in the project to enrich the metadata of their content, but will also be useful for every institution interested in metadata, such as other archives, technological partners, developers, commercial companies, etc.
Data Findable: Standard and metadata	The used standards for metadata will be defined in D6.1 on the metadata model of I-Media-Cities, but will be based upon metadata standard CEN.
Data openly accessible: Open or not, IPR, method of access, location of data, restriction of access	<p>User-generated metadata will generally be openly available, but can have restrictions on access and availability. An Informed consent form will define the IPR rules of any user-generated metadata. If there are no restrictions put upon metadata generated by users within any part of the project, this metadata will be made openly available.</p> <p>All user generated metadata will be stored in the metadata repository, located on the servers at Cineca. At a minimum the dataset will be downloadable from the 'results'- section of the project website in a CVS-format, but it might be possible that the dataset is available through use of an API or other software. If this is the case, this information and a handbook to use the software will be available on the 'results'-section of the project website.</p>
Data Interoperability Data license, Standards used	The used standards for metadata will be defined in D6.1 on the metadata model of I-Media-Cities, but will be based upon metadata standard CEN. There will be an open data license connected to this dataset, which will be published as part of the data policy of the dataset
Data Reuse Published on 3 rd party website	This dataset will be published on a third party website. The identity of the third party website where this dataset will be published will be defined in D3.4 Dissemination and Engagement Actions report.

Data preservation and archiving (Retention period, after retention period, central preservation location)	The data policy of this dataset will define the period this dataset will be retained on the servers at CINECA. After this retention period, this dataset will be stored on a central location of the website of ACE or the European Film gateway.
---	---

6. User-generated data (content)	
Data set reference and name	The ID of each item of content created by the users will be generated by the system and the structure of the ID will be defined in D6.1.
Data set purpose	Metadata that is generated by users for the purpose of enriching the metadata set accompanying any object of content in the project.
Data Origin	Created by users of the I-Media-Cities platforms and solutions and saved by them in the I-Media-Cities data repository
Data Size	The specific details of each type of content generated will be defined in the data model foreseen in D6.1.
Data Utility	This dataset will be useful for the archives participating in the project to enrich the metadata of their content, but will also be useful for every institution interested in metadata, such as other archives, technological partners, developers, commercial companies, etc.
Data Findable: Standard and metadata	The used standards for data/content will be defined in D2.3 User requirements – System requirements.
Data openly accessible: Open or not, IPR, method of access, location of data, restriction of access	<p>User-generated data will generally be openly available, but can have restrictions on access and availability. An Informed consent form will define the IPR rules of any user-generated metadata. If there are no restrictions put upon metadata generated by users within any part of the project, this metadata will be made openly available.</p> <p>All user generated data will be stored in the data repository, located on the servers at CINECA. At a minimum the dataset will be downloadable from the 'results'- section of the project website in a CVS-format, but it might be possible that the dataset is available through use of an API or other software. If this is the case, this information and a handbook to use the software will be available on the 'results'-section of the project website.</p>
Data Interoperability Data license, Standards used	The used standards for data will be defined in D2.3 User requirements – System requirements.

Data Reuse Published on 3 rd party website	This dataset will be published on a third party website. The identity of the third party website where this dataset will be published will be defined in D3.4 Dissemination and Engagement Actions report.
Data preservation and archiving (Retention period, after retention period, central preservation location)	The data policy of this dataset will define the period this dataset will be retained on the servers at CINECA. After this retention period, this dataset will be stored on a central location of the website of ACE or the European Film gateway

7. Metadata generated through the use of automatic analysis tools	
Data set reference and name	Each set of metadata is related to the content/fragment to which it refers to. The ID will be generated by the system and the structure of the ID will be defined in D6.1.
Data set purpose	Metadata that is generated through the use of automatic analysis tools, for the purpose of enriching the metadata set accompanying any object of content in the project.
Data Origin	Created by users the I-media-Cities Consortium
Data Size	Each type of generated metadata will have each own size. In any case the specific details will be defined in the data model foreseen in D6.1.
Data Utility	This dataset will be useful for the archives participating in the project to enrich the metadata of their content, but will also be useful for every institution interested in metadata, such as other archives, technological partners, developers, commercial companies, etc.
Data Findable: Standard and metadata	The possible use of standards for metadata will be defined in D6.1 on the metadata model of I-Media-Cities.
Data openly accessible: Open or not, IPR, method of access, location of data, restriction of access	Automatically generated metadata will be openly available. All automatically generated metadata will be stored in the metadata repository, located on the servers at CINECA. At a minimum the dataset will be downloadable from the 'results'- section of the project website in a CVS-format, but it might be possible that the dataset is available through use of an API or other software. If this is the case, this information and a handbook to use the software will be available on the 'results'-section of the project website.
Data Interoperability Data license, Standards used	The used standards for metadata will be defined in D6.1 on the metadata model of I-Media-Cities. There will be an open data license connected to this dataset, which will be published as part of the data policy of the dataset

Data Reuse Published on 3 rd party website	This dataset will be published on a third party website. The identity of the third party website where this dataset will be published will be defined in D3.4 Dissemination and Engagement Actions report.
Data preservation and archiving (Retention period, after retention period, central preservation location)	The data policy of this dataset will define the period this dataset will be retained on the servers at CINECA. After this retention period, this dataset will be stored on a central location of the website of ACE or the European Film gateway

8. Software	
Data set reference and name	Identified by the name of the release of the software following the standard notation.
Data set purpose	The software produced is needed to realize the I-media-Cities services
Data Origin	Created by the technical project partners of I-media-Cities.
Data Size	No limitation is foreseen on the size of the software produced.
Data Utility	This dataset will be useful for the project partners in the project to access and use the I-Media-Cities services
Data Findable: Standard and metadata	Not applicable
Data openly accessible: Open or not, IPR, method of access, location of data, restriction of access	The IPR of all software remains with the partners that produced this foreground. At the moment it is not foreseen for the software developed to be openly accessible. The software is considered for Consortium use only and cannot be accessed by anyone outside of the project.
Data Interoperability Data license, Standards used	Not applicable for the software

Data Reuse Published on 3 rd party website	Not applicable for the software
Data preservation and archiving Retention period, after retention period, central preservation location	The data policy of this dataset will define the period this dataset will be retained on the servers at CINECA. After this retention period, this dataset will be stored and archived in CINECA

9. Metadata model	
Data set reference and name	Identified by the name 'Metadata model I-Media-Cities'
Data set purpose	Metadata model that is generated for the purpose of structuring and unifying all metadata collected, analyzed and generated within the project and organizing the metadata repository
Data Origin	Created by the I-Media-Cities technical team
Data Size	Not applicable
Data Utility	This dataset will be useful for the archives participating in the project to enrich the metadata of their content, but will also be useful for every institution interested in metadata, such as other archives, technological partners, developers, commercial companies, etc. The structure of the model will allow it to be published as an open and linked model, which can be easily re-used by external and internal parties.
Data Findable: Standard and metadata	The used standards and all metadata entities will be defined in D6.1 on the metadata model of I-Media-Cities, but will be based upon metadata standards, such as CEN.
Data openly accessible: Open or not, IPR, method of access, location of data, restriction of access	The model is defined in D6.1 and will be openly accessible. The deliverable and the information on the model will be published on a public place of the project website and can be downloaded there in pdf. There is no restriction of access.

<p>Data Interoperability Data license, Standards used</p>	<p>The used standards for metadata will be defined in D6.1 on the metadata model of I-Media-Cities.</p>
<p>Data Reuse Published on 3rd party website</p>	<p>This dataset will be published on a third party website. The identity of the third party website where this dataset will be published will be defined in D3.4 Dissemination and Engagement Actions report.</p>
<p>Data preservation and archiving (Retention period, after retention period, central preservation location)</p>	<p>All documents will be kept for a period of five years after the end of the project lifetime. During the project lifetime they will be kept on the project website www.imediacities.eu and after the project lifetime, they will be stored in a dedicated location on the servers at technical partner CINECA in Italy. After the retention period ends, the documents will only be stored in hard copy at a dedicated location at the Coordinator</p>

Annex II: Terms of Use Template

Terms of Use

These terms of use are applicable to any use of any part of the web services of I-Media-Cities.

Please read these Terms of Use, the Privacy Policy and the Disclaimer carefully, before starting to use any of the I-Media-Cities web services.

1) Introduction

Welcome to I-Media-Cities. These are the general terms and conditions on which we supply all our services. If you use our services, you agree to abide by these terms.

At I-Media-Cities we supply a lot of different services. Some of them will have specific terms tailored for them. If that is the case, I-Media-Cities' contract with you for that service will be on these terms, supplemented by any terms specific to the service. In the case of any conflict, service specific terms will take precedence.

We process personal data in accordance with the [Open access to research data policy of the European Union](#). As part of this agreement you consent to our doing so. You should read the policy carefully, especially if you have any concerns about your privacy.

Warning: unless we have agreed a particular level of service with you, we make absolutely no promises about the quality or existence of any of our services. Please read the sections below and our general exclusion of liability.

All use of our web services are subject to our terms for web users. There are further terms if you have signed up for an account with one of our services. At the end are some general terms and definitions.

2) [Terms of Use for our web services](#)

Web services include any of our websites as well as **(fill out all web services)**

A. What you agree to

You agree not to use our websites to do any of the following:

- Anything which is illegal either where you are in the world, or where we are.
- Please refrain from the use of any inappropriate or hurtful language. I-Media-Cities in no way condones racism or any other form of prejudice. Please remember that anything you write on the internet can and will be found and read by several different people. I-Media-Cities will remove any hurtful or inappropriate comments and contributions and might deny offenders access to any part of their web services.

- Try to remember that not everyone has the same opinion as you might have. Not everybody considers history in the same way. If something is not clear, always ask for clarification first.
- Cause nuisance to other users of our services
- Interfere with the normal running of our services
- Try to access our systems in a way other than those advertised by us and, in particular, to use a web crawler that does not respect the robots exclusion policy.
- I-Media-Cities retains the right to judge, edit or remove any contribution and comments from its users and visitors. If any of these contributions are considered to be inappropriate by I-Media-Cities, they will be edited or removed without I-Media-Cities having to give formal notice of this change.
- If you consider any of the contributions or comments on any of the web services of I-Media-Cities to be inappropriate or hurtful, please don't hesitate to contact us via I-Media-Cities@cinematek.be.
- I-Media-Cities wants to hear from you. Please submit all your questions, comments and or recommendations via I-Media-Cities@cinematek.be.

B. Inappropriate use of the I-Media-Cities web services

- It is not permitted to use any information on our web services to extract and collect any user's personal data or any other kind of user information.
- Without our express consent, it is not permitted to use any part of our web services to contact any of our users with commercial or other types of messages and comments.
- It is not permitted to use any part of any of our web services to distribute any messages of a commercial or idealistic nature. It is a given that this is also not permitted by distributing them as a contribution or comment.

C. Other websites

Some of our activities are carried out on web platforms provided by third parties. For example the ... (fill out) is hosted on (Fill out). If you make use of any service where that is the case, you are responsible for complying with any terms of service of the third party platform.

The web services of I-Media-Cities might contain hyperlinks to third party websites. I-media-Cities has no claims nor control over these third party websites and the user has to comply to any and all terms of service applicable to this third party platform. I-media-Cities cannot be held accountable or liable for any damages that might derive from the use of these third party websites and platforms.

D. Accounts

Some of our services require you to create an account in order to make certain, or any, use of the service. You can create a personal account on our website, but it is also possible to create an account through (Fill out if applicable). Your personal profile consists of at least your user name and an email address. You can add other information to your profile, such as (fill out if applicable). You as the creator of the account, always has the final decision on what extra information you want to add.

By logging in to your user account, you are able to contribute to certain parts of our web services. Your activities will be linked to your personal account and some or all of them will be shown on your personal profile page. (Scratch if not applicable) You are allowed to remove anything you have added to any part of our web services. Please also carefully check the Privacy Policy of I-media-Cities.

Your profile is protected by a password, which you yourself have chosen. Please keep this password private and never share it with anyone else. Change your password regularly.

All our accounts are subject to the following rules.

Rules for our accounts

- You must be at least .. (Fill out) years old and a human being.
- If asked for any personal details, you must answer truthfully (see our privacy policy for what we do with those details). You must supply us with a valid e-mail address.
- You are responsible for the security of your accounts and making sure that any contact details in the account are kept up to date. If we need to contact you but are unable to do so, for example because your e-mail address is no longer valid, then any consequences of that failure will be your responsibility.
- You must not let anyone else use your account. If pressure is applied to you to do so — for example if an employer demands your username and password — please inform them that their attempt to subvert your agreement with us will mean that they have no permission to use any of our services. We may take action, including criminal prosecution, if they use our services using an account they have obtained in this way.
- You must let us know of any un-authorized use of your account as soon as you are able to after becoming aware of it. Please contact us via I-Media-Cities@cinematek.be
- Unless an account is associated with a paid-for service, we may suspend or terminate it at any time. Equally, you may close your account at any time

Removal of your accounts

You can ask to have your user profile removed at any given time. In order to do so, please send a message with your request to I-Media-Cities@cinematek.be. In case of an account removal, I-Media-Cities retains the right not to remove parts of or entire contributions attached to that account, if I-Media-Cities judges these contributions to be of general use to the web services. If that is the case, I-Media-Cities will take over the copyright and responsibility of these contributions. If expressly asked, I-Media-Cities will also change the user name and source attached.

E. Community Members

Our most important class of account is one you may create in order to become a community member. By registering as a community member, you are acknowledging your connection to us. You are subject to any rules for community members we may publish and we may send you email messages we think appropriate for members, for example in order to poll you on some important issue.

Community membership is not membership in the formal sense of membership of a company limited by guarantee.

F. Content and intellectual property

General Rules

- Users of our web services with a personal account, can contribute to parts of our web services by ... (Fill out what is applicable).

- Contributions you provide to any part of the web services can be indexed and tagged, which will allow them to link to other contributions on and parts of the web services.
- Not All / All (Scratch which is applicable) contributions and comments are public and visible for all users and visitors of the I-Media-Cities web services. You are always responsible for the content of your own contributions and comments.

Intellectual Property

- When you provide any contribution, you guarantee that you are licensed to provide this contribution or comment and that you do not infringe on the intellectual property of any third party. You are at all times obligated to provide a correct mention of sources and rights holders, if applicable to your contributions.
- I-Media-Cities retains the right to edit or remove any contributions and comments any of its users or visitors have made on any part of its web services, if these contributions or comments are not in line with our terms of use or constitute an infringement of any of our terms of use or data policies.
- If I-media-Cities is approached by a third party because of a copyright infringement as a result of one or more of your contributions or comments on one of our web services, you will be hold accountable for any costs, claims and damages I-media-Cities might encounter as a result. I-Media-Cities will do its utmost to limit these claims and costs. I-Media-Cities will immediately edit or remove any of its users' or visitors' contributions that have been reported on as copyright infringement.
- User or visitors that repeatedly infringe on our terms of use or any of our policies and/or licenses, will be permanently or temporarily refused access to any part of our web services.
- Except in cases where a third party has any copyright claims on any part of your contributions or comments, you are considered to be the copyright holder of any comment and contribution you make on any part of our web services. If you feel that you are the copyright holder of any contribution or comment anybody else has provided to the I-Media-Cities services, please contact us immediately via I-Media-Cities@cinematek.be
- I-Media-Cities is and wants to be an open source of information. Information available on our web services must be able to be shared and re-used by third party websites and for educational purposes. I-Media-Cities will do its utmost to provide a possibility to share information through other channels, such as ... (Fill out).

What we do with your content

If you contribute content to any of our services, for example by commenting on a task, or uploading data, then as a general rule you agree to license that content to us under the same license as prevails for that service or website.

For example: (fill out)

The only exception to this policy is where a service we supply to you expressly allows a different license, for example a task run submitted to a given project would contain your data under whatever license — including no license — you choose.

G. Our content

Unless otherwise stated all our services are offered under open content or data licenses and you should refer to the provisions of the license in question to find out what you are allowed to do. Some of our content belongs to third parties. Most third party data is subject to an open license, but we cannot guarantee it. You should refer to the third party if you are in doubt. You are obliged to always

mention I-Media-Cities when referencing any part of us under license and are asked to even provide a link to the correct part of our web services. All parts of the brand, the trademarks and the logos of I-Media-Cities are protected under copyright law and can in no way be licensed or reused by unauthorized parties.

H. Ownership does not change

As a general rule, this agreement will not change the ownership of any intellectual property belonging to either party. Where your content is used by us or vice versa both you and we would do so under a license (see above).

I. Liability

Indemnities where you may owe us

If you breach any of your obligations under this agreement and, as a result, cause us to be sued by anyone else, you will have to compensate us for any loss we have suffered as a result, which includes any costs, such as paying lawyers, or for our own time, we incur defending a claim as well as any damages awarded. If your breach causes you to be sued by someone else, you will not sue us for any loss you suffer as a result.

Unless clearly expressed otherwise, all contributions by any of our users and visitors on any of our web services are the responsibility of our users and visitors. In other words, no rights or claims towards I-Media-Cities can be attributed to any of contribution or comment made by users or visitors on our web services. I-Media-Cities cannot be held accountable for any damages that might occur as a result of any contributions or comments any user has provided to any part of the I-Media-Cities web services.

Exclusion what we do not owe you

We limit our liability in several different ways - all of which we believe to be fair. In case any one of them is found to be unenforceable by a court, each of the following limitations of liability is separate and our liability to you is limited by all of them.

- All exclusions of liability are only in so far as we are allowed to do so by whatever law applies to the situation. For example, (Fill out)
- We will not be liable for any damage that was not reasonably foreseeable at the time we made this agreement.
- We are not liable for any loss which is indirect or consequential. That includes any loss of business or profit.
- We exclude, in so far as we are allowed, any warranties that would be implied by law.

3) Open Licenses

The *I-Media-Cities* research platform uses ... (Fill out if any open source is used) which is released under the ... (Fill out applicable license of this open source) and all of the source code for the applications is hosted on ... (Fill out if any source code is hosted on a third party site).

For datasets, we make a distinction between **source** data and **derived** data. Our crowd-sourcing projects use digital **source** data (for example films, photographs, scanned documents) to allow contributors to complete research tasks such as the tagging and transcription. This **source** data will typically be available through a ... (Fill out license applicable to the data), but in some cases might be placed under ... (Fill out if any other licenses apply) licenses when the institutions owning the archival material have asked that commercial use is prohibited without further consultation (we hope that this will be infrequent).

All the **derived** data, produced with the help of volunteers via use of any of our web services, is licensed under an open ... (Fill out applicable) license, depending on case-by-case agreements with our partners.

Check the project's information pages for further details about which specific licenses have been applied.

4) General conditions

This agreement

We may update these terms and conditions at any time. If we do so, we will announce the change via our social media and main site (www.imediacities.eu). Any changes will be binding on you from the moment we announce them. All terms, policies and rules applying to I-media-Cities will always be accessible and available on a dedicated part of the website.

This agreement is made under the laws of Belgium. I-Media-Cities will always try to solve any problems and differences of opinion amicably and through communication. If all else fails, these differences will be referred to a court of law in Belgium.

Boilerplate

These final "boilerplate" terms should go without saying, but we are saying them anyway just to be clear.

- If any part of this agreement is ineffective (for example because it is unlawful) then the rest of the agreement should be read without it.
- This agreement is between you and us and is not intended to give anyone else any rights.
- We may sometimes fail to enforce our rights under this agreement (for example because we decide not to, or we did not realize you were in breach of contract). Just because we have not enforced any of our rights, does not stop us from doing so in the future.
- Neither party is liable for anything which is beyond their reasonable control.
- If for some reason beyond I-Media-Cities' reasonable control, we are unable to or it would not be commercially viable for us to, continue to supply any of our services, we may cease to supply that service, ending any agreement between us for its supply. If we do so, we will return to you a fair proportion of any sum you have paid us in advance for the supply of that service, taking into account the service we have already supplied to you.

Further Information

I-media-Cities and any of its web services are copyright protected and controlled by the I-Media-Cities Consortium. Any comments, questions and complaints can, at all times, be directed to the address and telephone numbers below.

I-Media-Cities Consortium

Cinematek

Rue Ravenstein 3

1000 Brussels

Belgium

0032/2 551 19 00

Annex III. Privacy Policy template

Privacy Policy

This Privacy Policy is applicable to any use of any part of the web services of I-Media-Cities.

Please read this Privacy Policy, the Terms of Use and the Disclaimer carefully, before starting to use any of the I-Media-Cities web services.

Introduction

The I-media-Cities Consortium considers the privacy of and dealing with personal data of the users of our web services as extremely important. Personal data are carefully and respectfully collected, processed and stored, in accordance with all applicable national, European and international laws.

1. Processing Personal Data

The I-Media-Cities Consortium processes personal data of every user who registers an account on the website or provides a contribution or comment to any part of our web services. In order to be able to provide a contribution or comment, a user is required to at least register a user name and an email address.

If you register an account at I-Media-Cities, you are able to provide the platform with added information, such as ... (Fill out) It is your decision and responsibility which added information you provide to the I-Media-Cities web service. Some or all of your activities on any of the I-Media-Cities web services will be added to your personal workspace.

Profiles, except your email address, your contributions and your comments are / are not (Scratch which does not apply) Your comments and contributions can also be shared on third party sites and web solutions, such as Facebook, Twitter, etc. Because of the general functionalities of the internet, public contributions will be registered by websites such as Google.

2. Purpose of Processing Personal Data

The I-Media-Cities Consortium collects and processes your data in order to optimize, adjust, enhance and develop our web services and their usability.

The I-Media-Consortium can also use your data to offer and suggest relevant information.

Other than the way it was written in the Privacy Policy and restricted to the rules and regulations of all applicable laws, I-Media-Cities will not make any of your personal data public, unless: (a) it is obligated to on the grounds of law, regulations, orders of conduct or ordered to by a court ruling; (b) it is necessary for the sake of security reasons; and/or (c) the I-Media-Cities Consortium has a reasonable suspicion the relevant user is part of any illegal activities.

3. Cookies

Every time a user visits the website, that visit will be registered by means of so-called 'cookies'. Cookies are small files that are stored on the user's hard drive, which allow the I-Media-Cities Consortium to access information about the use of their website and which allows for preferences, user names and passwords to be automatically remembered. This information can also be analyzed to discover trends and draw statistics relating to such things as the parts of the website that are most visited, average time of a visit, etc. In order to do this, I-Media-Cities will use Google Analytics (on which the [privacy policy of Google](#) applies)

The information that is collected through cookies can only be coupled to a computer and not to a specific user. It will therefore not allow the I-Media-Cities Consortium to derive any conformation that reveals the identity of a user. You have the ability to format your browser in such a way that cookies are refused. By doing this, it is possible that the website will no longer function optimally.

4. Safety of Personal Data

The I-Media-Cities Consortium will take every organizational and technical measure deemed reasonable to make sure your personal data are safe, correct and up-to-date. The I-Media-Cities Consortium will safely store your personal data and take all necessary measures to protect them from loss, abuse, illegal access and publication. The Consortium, however, cannot guarantee that personal data transferred over the internet are always 100% protected.

The I-Media-Cities Consortium will take reasonable measures to remove personal data or make them anonymous, as soon as they no longer need to be stored.

5. Access to and Correction of Personal Data

At any given time, you are able to send the I-Media-Cities Consortium a request to receive all personal data they have stored of you. The I-Media-Cities Consortium will make sure that the requested information is send to you within four weeks of them receiving your request.

If you would like to change or correct your own personal data, you can submit a request to do so via I-Media-Cities@cinematek.be. The I-Media-Cities Consortium will reply within four weeks of them receiving your request whether and to what extent your request will be fulfilled.

Annex IV. Disclaimer template

Disclaimer

This Disclaimer is applicable to any use of any part of the web services of I-Media-Cities.

Please read this Disclaimer, the Terms of Use and the Privacy Policy carefully, before starting to use any of the I-Media-Cities web services.

While we are proud of the services we provide to the world for free and try to make them reliable and useful, we make no promises about them. All web services are certain to fail some of the time. We adapt and change our services from time to time, so you may find that something that worked for you may cease to work. We may also stop supplying any service, temporarily or permanently or block access to our services to anyone for any reason.

If you need some guarantee of a particular service level, then please do not hesitate to contact us to discuss whether we could offer you a commercial version of any of our services (assuming one is not available).

We are not a law firm and nothing we do is intended to be taken as legal advice. In particular:

We encourage the use of one of the [open data commons licenses](#) but we cannot give any warranty that they will work in the way expected or should be used for any specific purpose. If in doubt, you should take your own legal advice.