



FOSTERING INDUSTRIAL SYMBIOSIS FOR A SUSTAINABLE RESOURCE  
INTENSIVE INDUSTRY ACROSS THE EXTENDED CONSTRUCTION VALUE CHAIN

# First version of the Data Management Plan

FENIX

**D10.3: First version of the Data Management Plan**  
WP 10, T 10.3

Authors: Petra Novotna (FEN)

H2020-WASTE-2014-two-stage



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 642154.

### Technical References

|                        |   |
|------------------------|---|
| <b>Project Acronym</b> | FISSAC  |
| <b>Project Title</b>   | FOSTERING INDUSTRIAL SYMBIOSIS FOR A SUSTAINABLE RESOURCE INTENSIVE INDUSTRY ACROSS THE EXTENDED CONSTRUCTION VALUE CHAIN |
| <b>Number</b>          | 642154  |
| <b>Start date</b>      | 01/09/2015  |
| <b>End date</b>        | 29/02/2020  |

|  |                          |
|--|--------------------------|
| <b>Deliverable No.</b>                   | D10.3                    |
| <b>Dissemination level</b> <sup>11</sup> | PU                       |
| <b>Work Package</b>                      | WP 10 – Management       |
| <b>Task</b>                              | T 10.3 – Data Management |
| <b>Lead beneficiary</b>                  | FEN                      |
| <b>Contributing beneficiary(ies)</b>     | ACC, ACR+                |
| <b>Due date of deliverable</b>           | 29 February 2016         |
| <b>Actual submission date</b>            | 29 February 2016         |

### Document history

| V | Date       | Beneficiary | Author                            |
|---|------------|-------------|-----------------------------------|
| 1 | 01/02/16   | FEN         | Petra Novotna                     |
| 2 | 09/02/16   | ACR+        | Angeliki Koulouri                 |
| 3 | 15/02/16   | ACC         | Daniel Hiniesto Muñoz de la Torre |
| 4 | 24/02/2016 | ACR+        | Angeliki Koulouri                 |

<sup>1</sup> PU = Public

PP = Restricted to other programme participants (including the Commission Services)

RE = Restricted to a group specified by the consortium (including the Commission Services)

CO = Confidential, only for members of the consortium (including the Commission Services)



## 0 Summary

According to the Guidelines on Open access to Scientific Publications and Research Data for projects funded or co-funded under Horizon 2020, Europe 2020 strategy underlines the central role of knowledge and innovation in growth generation. For these reasons the European Union strives to improve access to scientific information and to boost the benefits of public investment in the research funded under the EU Framework Programme Horizon 2020.

The project FISSAC participates in the Commission's Open Pilot Open Research Data to research data policy for facilitating access, re-use and preservation of research data. Deliverable D10.3 Data Management Plan (DMP) is required for the FISSAC project and will be carried out in T10.3 (WP10), as a structural activity that will identify the results that should be subject of dissemination and exploitation activities. The DMP also analyses the main data uses and users and explore the restrictions related to IPR according with the Consortium Agreement. This DMP deliverable is prepared in compliance with the template provided by the Commission in the Annex 1 of the "Guidelines on Data Management in Horizon 2020".

This Data Management Plan has been prepared to define the data assurance processes that are to be applied during and after the completion of the project.

## Table of content

|             |   |           |
|-------------|---|-----------|
| <b>0</b>    | <b><u>SUMMARY</u></b>                                       | <b>3</b>  |
| <b>1</b>    | <b><u>INTRODUCTION</u></b>                                  | <b>6</b>  |
| <b>2</b>    | <b><u>OPEN ACCESS AND OPEN RESEARCH DATA PILOT</u></b>      | <b>7</b>  |
| <b>2.1.</b> | <b>DISSEMINATION, COMMUNICATION AND OPEN ACCESS</b>         | <b>7</b>  |
| <b>2.2.</b> | <b>OPEN ACCESS TO PEER-REVIEWED SCIENTIFIC PUBLICATIONS</b> | <b>7</b>  |
| <b>2.3.</b> | <b>OPEN ACCESS TO RESEARCH DATA</b>                         | <b>8</b>  |
| <b>3</b>    | <b><u>DMP OBJECTIVE</u></b>                                 | <b>10</b> |
| <b>4</b>    | <b><u>INFORMATION MANAGEMENT AND POLICY</u></b>             | <b>11</b> |
| <b>5</b>    | <b><u>DMP IMPLEMENTATION</u></b>                            | <b>14</b> |
| <b>6</b>    | <b><u>RESEARCH DATA</u></b>                                 | <b>16</b> |
| <b>6.1.</b> | <b>CHARACTERISTICS FOR DATASETS PRODUCED IN THE PROJECT</b> | <b>16</b> |
| <b>7</b>    | <b><u>FISSAC DATA SETS</u></b>                              | <b>17</b> |
| <b>7.1</b>  | <b>COLLECTION AND MANAGEMENT OF FISSAC DATA SETS</b>        | <b>17</b> |
| <b>8</b>    | <b><u>DATA SETS TECHNICAL REQUIREMENTS</u></b>              | <b>21</b> |
| <b>8.1</b>  | <b>GENERAL REQUIREMENTS</b>                                 | <b>21</b> |
| <b>8.2</b>  | <b>PROHIBITED FILE TYPES</b>                                | <b>21</b> |
| <b>8.3</b>  | <b>STATIC GRAPHICAL IMAGES</b>                              | <b>21</b> |
| <b>8.4</b>  | <b>ANIMATED GRAPHICAL IMAGE</b>                             | <b>22</b> |
| <b>8.5</b>  | <b>AUDIO DATA</b>   | <b>22</b> |
| <b>8.6</b>  | <b>TEXTUAL DATA</b>   | <b>22</b> |
| <b>8.7</b>  | <b>NUMERIC DATA</b>   | <b>22</b> |
| <b>8.8</b>  | <b>PROCESS AND TEST DATA</b>                                | <b>22</b> |
| <b>8.9</b>  | <b>MICROSOFT OFFICE APPLICATION SUITE</b>                   | <b>23</b> |
| <b>9</b>    | <b><u>NAMING CONVENTION</u></b>                             | <b>24</b> |
| <b>10</b>   | <b><u>CONCLUSIONS</u></b>                                   | <b>25</b> |
| <b>11</b>   | <b><u>REFERENCES</u></b>                                    | <b>26</b> |
| <b>12</b>   | <b><u>APPENDICES</u></b>                                    | <b>27</b> |
| <b>A1</b>   | <b>DATA TYPES METADATA PARAMETERS</b>                       | <b>27</b> |
| <b>A2</b>   | <b>METADATA DEFINITION TEMPLATES</b>                        | <b>28</b> |
| <b>A3</b>   | <b>METADATA STANDARDS</b>                                   | <b>28</b> |
| <b>A4</b>   | <b>DATA SETS TECHNICAL REQUIREMENTS</b>                     | <b>29</b> |
| <b>A5</b>   | <b>DEFINITIONS</b>  | <b>30</b> |



## List of Tables

|   |    |
|---|----|
| Table 1: FISSAC project partners and their roles .....        | 15 |
| Table 2: Forecast of FISSAC datasets related to each WP ..... | 17 |
| Table 3: Image default sizes and resolution .....             | 22 |

## List of Figures

|  |    |
|--|----|
| Figure 1 Data Management Plan overview.....  | 10 |
| Figure 2: Information Management .....       | 12 |
| Figure 3: FISSAC website .....               | 13 |
| Figure 4: FISSAC organization structure..... | 14 |
| Figure 5: OpenAIRE website .....             | 19 |
| Figure 6: ZENODO repository .....            | 19 |
| Figure 7: Data Entry Box – Summary .....     | 23 |

## Abbreviations and acronyms

|      |                                       |
|------|---------------------------------------|
| WP   | Work Package                          |
| DMP  | Data Management Plan                  |
| GA   | Grant Agreement                       |
| CA   | Consortium Agreement                  |
| CMS  | Content Management System             |
| CMA  | Content Management Application        |
| CDA  | Content Delivery Application          |
| IM   | Information Management                |
| ECM  | Enterprise Content Management         |
| ERM  | Electronic Records Management         |
| BPM  | Business Process Management           |
| GMS  | Government Metadata Standard          |
| PU   | Public                                |
| CO   | Confidential                          |
| CI   | Classified                            |
| JPEG | Joint Photographic Experts Group      |
| JFIF | JPEG File Interchange Format          |
| PNG  | Portable Network Graphics             |
| AVI  | Audio Video Interleaved               |
| WAV  | Waveform Audio File Format            |
| RIFF | Resource Interchange File Format      |
| LPCM | Linear Pulse-Code Modulation          |
| AIFF | Audio Interchange File Format         |
| MIME | Multipurpose Internet Mail Extensions |
| OA   | Open Access                           |

## 1 Introduction

This document constitutes the first issue of Data Management Plan (DMP) in the EU framework of the project FISSAC under Grant Agreement No 642154. The objective of the DMP is to establish the measures for promoting the findings during the project's life. The DMP enhances and ensures relevant project's information transferability and takes into account the restrictions established by the Consortium Agreement. In this framework, the DMP sets the basis for both Dissemination Plan and Exploitation Plan. The first version of the DMP is delivered at M6; later the DMP will be monitored and updated in parallel with the different versions of Dissemination and Exploitation Plans (the progress of the implementation of DMP will be included in the Project Progress Reports, at M18 and M36. It is acknowledged that not all data types will be available at the start of the project. However and whenever important, if any changes occur to the FISSAC project due to inclusion of new data sets, changes in consortium policies or external factors, the DMP will be updated as well in order to fine-tune it to the actual data generated and the user requirements as identified by the FISSAC consortium participants.

FISSAC project comprises seven technical work packages (WP) as follows:

- WP1 - FROM CURRENT MODELS OF INDUSTRIAL SYMBIOSIS TO A NEW MODEL
- WP2 - CLOSED LOOP RECYCLING PROCESSES TO TRANSFORM WASTE INTO SECONDARY RAW MATERIALS
- WP3 - PRODUCT ECO-DESIGN AND CERTIFICATION
- WP4 - PRE-INDUSTRIAL SCALE DEMONSTRATION OF THE RECYCLING PROCESSES AND ECO-INNOVATIVE PRODUCTS
- WP5 - INDUSTRIAL PRODUCTION & REAL SCALE DEMONSTRATION
- WP6 - FISSAC MODEL FOR INDUSTRIAL SYMBIOSIS
- WP7 - INDUSTRIAL SYMBIOSIS REPLICABILITY AND SOCIAL ISSUES

To facilitate the technical work there are three transversal work packages to provide, structure, coordination, integration and communications across all the work packages.

- WP8 - EXPLOITATION AND BUSINESS MODELS FOR INDUSTRIAL SYMBIOSIS
- WP9 - DISSEMINATION
- WP10 - MANAGEMENT

This document has been prepared to describe the data management life cycle for all data sets that will be collected, processed or generated by FISSAC project. It is a document outlining how research data will be handled during FISSAC project, and after the project is completed. It describes what data will be collected, processed or generated and what methodologies and standards are to be applied. It also defines if and how this data will be shared and/or made open, and how it will be curated and preserved.

## 2 Open Access and Open Research Data Pilot

Open access can be defined as the practice of providing on-line access to scientific information that is free of charge to the reader and that is reusable. In the context of research and innovation, 'scientific information' can refer to:

- (i) peer-reviewed scientific research articles (published in scholarly journals) or
- (ii) research data (data underlying publications, curated data and/or raw data).

The EC capitalises on open access and open science as it lowers barriers to accessing publicly-funded research. This increases research impact, the free-flow of ideas and facilitates (innovation in) a knowledge-driven society at the same time underpinning the EU Digital Agenda (OpenAIRE Guide for Research Administrators - EC funded projects). Open access policy of European Commission is not a goal in itself, but an element in promotion of affordable and easy accessible scientific information for the scientific community itself, but also for innovative small businesses.

### 2.1. Dissemination, Communication and Open Access

For the implementation of FISSAC project, there is a complete dissemination and communication set of activities scheduled, with the objectives of raising awareness among non-expert citizens, but potential next users of the FISSAC knowledge and solutions. For instance, e-newsletters, e-brochures, poster or events, are foreseen for the dissemination of FISSAC to key groups potentially related to the project results' exploitation.

Likewise, FISSAC website, webinars, press releases or short videos, for instance, will be developed for a Communication to a wider audience. Details about all those dissemination and communication elements are provided in the Deliverable D9.1 "Dissemination Plan".

Open Access (OA) to scientific information is a complementary element to dissemination and communication, and how this issue is specifically tackled by FISSAC project is described in the present document.

### 2.2. Open Access to peer-reviewed scientific publications

Open access to scientific peer-reviewed publications has been anchored as an underlying principle in the Horizon 2020 Regulation and the Rules of Participation and is consequently implemented through the relevant provisions in the grant agreement.

More specifically, Article 29: "Dissemination of results, Open Access, Visibility of EU Funding" section 2 of FISSAC Grant Agreement (FISSAC, Research & Innovation action, 2014) establishes the obligation to ensure open access to all peer-reviewed articles produced by FISSAC.

#### 29.2 Open access to scientific publications

Each beneficiary must ensure open access (free of charge online access for any user) to all peer reviewed scientific publications relating to its results.

In particular, it must:

- (a) as soon as possible and at the latest on publication, deposit a machine-readable electronic copy of the published version or final peer-reviewed manuscript accepted for publication in a repository for scientific publications;

Moreover, the beneficiary must aim to deposit at the same time the research data needed to validate the results presented in the deposited scientific publications.

- (b) ensure open access to the deposited publication — via the repository — at the latest:
- (i) on publication, if an electronic version is available for free via the publisher, or
  - (ii) within six months of publication (twelve months for publications in the social sciences and humanities) in any other case.
- (c) ensure open access — via the repository — to the bibliographic metadata that identify the deposited publication.

The bibliographic metadata must be in a standard format and must include all of the following:

- the terms “European Union (EU)” and “Horizon 2020”;
- the name of the action, acronym and grant number;
- the publication date, and length of embargo period if applicable, and
- a persistent identifier.

## 2.3. Open Access to research data

Research data is the second type of scientific information that OA is planned for, besides the publications. 'Research data' refers to information, in particular facts or numbers, collected to be examined and considered and as a basis for reasoning, discussion, or calculation. In a research context, examples of data include statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images. The focus is on research data that is available in digital form.

Open Research Data Pilot is a novelty in Horizon 2020 aiming to improve and maximise access to and re-use of research data generated by projects (European Commission, 9 December 2013). Particularly FISSAC is participating in this Open Research Data Pilot programme as issued in Article 29 article 3:

### 29.3 Open access to research data

Regarding the digital research data generated in the action ('data'), the beneficiaries must:

- a) deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the following:
  - i. the data, including associated metadata, needed to validate the results presented in scientific publications as soon as possible;
  - ii. other data, including associated metadata, as specified and within the deadlines laid down in the 'data management plan' (see Annex 1);
- b) provide information — via the repository — about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and — where possible — provide the tools and instruments themselves).

This does not change the obligation to protect results in Article 27, the confidentiality obligations in Article 36, the security obligations in Article 37 or the obligations to protect personal data in Article 39, all of which still apply.

The beneficiaries do not have to ensure open access to specific parts of their research data if the achievement of the action's main objective, as described in Annex 1, would be jeopardised by making those specific parts of the research data openly accessible. In this case, the data management plan must contain the reasons for not giving access to third parties.

### Consortium Agreement - Access Rights

The Parties have identified and agreed on the Background for the Project and have also, where relevant, informed each other that Access to specific Background is subject to legal restrictions or limitations. Anything which has not been identified in the Consortium Agreement shall not be the object of Access Right obligations regarding Background. Any Party can propose to the General Assembly to modify its Background in the Consortium Agreement.



Each Party shall implement its tasks in accordance with the Consortium Plan and shall bear sole responsibility for ensuring that its acts within the Project do not knowingly infringe third party property rights.

- Any Access Rights granted expressly will exclude any rights to sublicense unless expressly stated otherwise.
- Access Rights shall be free of any administrative transfer costs.
- Access Rights are granted on a non-exclusive basis.
- Results and Background shall be used only for the purposes for which Access Rights to it have been granted.
- All requests for Access Rights shall be made in writing.
- The granting of Access Rights may be made conditional on the acceptance of specific conditions aimed at ensuring that these rights will be used only for the intended purpose and that appropriate confidentiality obligations are in place.
- The requesting Party must show that the Access Rights are needed.

Access Rights to Results and Background Needed for the performance of the own work of a Party under the Project shall be granted on a royalty-free basis, unless otherwise agreed for Background in Consortium Agreement.

Access Rights to Results if needed for Exploitation of a Party's own Results shall be granted on Fair and Reasonable conditions to be agreed in writing among the Parties concerned.

Access rights to Results for internal non-commercial research activities shall be granted on a royalty-free basis.

Access Rights to Background if Needed for Exploitation of a Party's own Results, including for research on behalf of a third party listed in Attachment 3, shall be granted on Fair and Reasonable conditions.

A request for Access Rights may be made up to twelve months after the end of the Project or after the termination of the requesting Party's participation in the Project.

Affiliated Entities have Access Rights under the conditions of the Grant Agreement if they are identified in attachment "Identified Affiliated Entities" to this Consortium Agreement. Such Access Rights must be requested by the Affiliated Entity from the Party that holds the Background or Results. Alternatively, the Party granting the Access Rights may individually agree with the requesting Party to have the Access Rights include the right to sublicense to the latter's Affiliated Entities. Access Rights to Affiliated Entities shall be granted on Fair and Reasonable conditions and upon written bilateral agreement.

Affiliated Entities which obtain Access Rights in return they should fulfil all confidentiality and other obligations accepted by the Parties under the Grant Agreement or this Consortium Agreement as if such Affiliated Entities were Parties.

Access Rights may be refused to Affiliated Entities if such granting is contrary to the legitimate interests of the Party which owns the Background or the Results.

Access Rights granted to any Affiliated Entity are subject to the continuation of the Access Rights of the Party to which it is affiliated, and shall automatically terminate upon termination of the Access Rights granted to such Party.

Upon termination of the status as an Affiliated Entity, any Access Rights granted to such former Affiliated Entity shall lapse.

Further arrangements with Affiliated Entities may be negotiated in separate agreements.

### 3 DMP Objective

The purpose of FISSAC Data Management Plan (DMP) is to provide a management assurance framework and processes that fulfil the data management policy that will be used by the FISSAC project participants with regard to all the dataset types that will be generated by the FISSAC project.

The aim of the DMP is to control and ensure quality of project activities, and to effectively/efficiently manage the material/data generated within the FISSAC project. It also describes how data will be collected, processed, stored and managed holistically from the perspective of external accessibility and long term archiving.

All aspects of procedures that are associated with the quality control of data management internal to the project is the subject of a separate deliverable, D10.2 “Quality Assurance Plan”.

The content of the DMP is complementary to other official documents that define obligations under the Grant Agreement (GA) and associated annexes, and shall be considered a living document and as such will be the subject of periodic updating as necessary throughout the lifespan of the project.

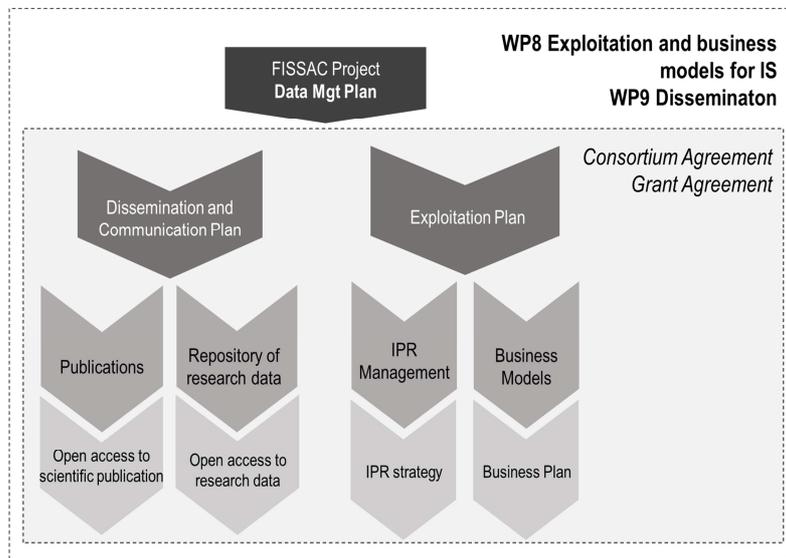


Figure 1 Data Management Plan overview

## 4 Information Management and Policy

The information available to different stakeholders will be managed and stored in a Content Management System (CMS) taking advantage of existing information management open sources that could be adaptable to project data dissemination needs. CMS offers different levels of accessibility depending on the degree of confidentiality of the information. It includes both, Publications and Repository of other research data. Open access to research data refers to right to access and re-use digital research data under the terms and conditions set out in the Grant Agreement.

### Content Management System

A content management system is a computer application that allows publishing, editing, modifying, organizing, deleting, and maintaining content from a central interface. Such systems of content management provide procedures to manage workflow in a collaborative environment. These procedures can be manual steps or an automated cascade. CMSs have been available since the late 1990s.

The function of CMS is to store and organize files, and provide version-controlled access to their data. CMS features vary widely. Simple systems showcase a handful of features, while other releases, notably enterprise systems, offer more complex and powerful functions. Most CMSs include Web-based publishing, format management, (version control), indexing, search, and retrieval. The CMS increases the version number when new updates are added to an already-existing file. Some content management systems also support the separation of content and presentation. A CMS may serve as a digital asset management system containing documents, movies, pictures, phone numbers, scientific data. CMSs can be used for storing, controlling, revising, semantically enriching and publishing documentation. Distinguishing between the basic concepts of user and content. The CMS has two elements:

- **Content Management Application (CMA)** is the front-end user interface that allows a user, even with limited expertise, to add, modify and remove content from a Web site without the intervention of a Webmaster.
- **Content Delivery Application (CDA)** compiles that information and updates the Web site.

### Information Management

Information Management (IM) is the collection and management of information from one or more sources and the distribution of that information to one or more audiences. This sometimes involves those who have a stake in, or a right to that information. Management means the organization of and control over the structure, processing and delivery of information.

Information includes both electronic and physical information. The organizational structure must be capable of managing this information throughout the information lifecycle regardless of source or format (data, paper documents, electronic documents, audio, social business, video, etc.) for delivery through multiple channels that may include cell phones and web interfaces. The focus of IM is the ability of organizations to capture, manage, preserve, store and deliver the right information to the right people at the right time.

Information management environments are comprised of legacy information resident in line of business applications, Enterprise Content Management (ECM), Electronic Records Management (ERM), Business Process Management (BPM), Taxonomy and Metadata, Knowledge Management (KM), Web Content Management (WCM), Document Management (DM) and Social Media Governance technology solutions and best practices.

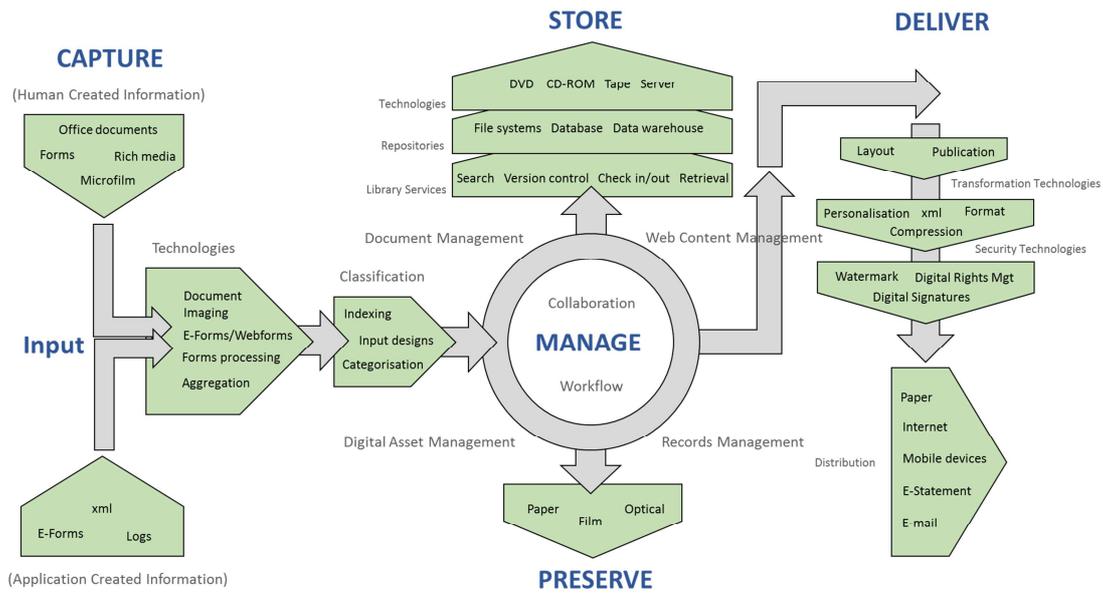


Figure 2: Information Management

### FISSAC project website

Project website will be used for storing both public and private documents related to project and dissemination, the website is meant to be live for the whole project duration and minimum 2 years after the project ends.

- Public section of the project website: public deliverables, brochure, poster, presentations, scientific papers, videos, etc.
- Private section of the project website: confidential deliverables, work packages related documentation, etc.

The website [www.fissacproject.eu](http://www.fissacproject.eu) was launched on 15<sup>th</sup> of January 2016. The website was designed by a subcontractor and will be managed by ACR+. It will be dynamic and interactive in order to ensure a clear communication and wide dissemination of project news, activities and results. The website is of primary importance due to the expected impact on the target audiences. It was designed to give quick, simple and neat information. The website will be regularly updated with news and articles. It will also provide access to the FISSAC platform and FISSAC model, once they are online. All partners are responsible for feeding the project website with news and relevant information. The website will remain at least two years after the end of the project (February 2020). The website will be available in English and in the languages of the project partners (Czech, French, German, Hungarian, Italian, Spanish, Swedish and Turkish). However, the information will be selectively translated where needed in the various languages of the partnership, specifically for hosting regional workshops, webinars and for disseminating local news.



### About the Project

The FISSAC project involves stakeholders at all levels of the construction and demolition value chain to develop a methodology, and software platform to facilitate information exchange, that can support industrial symbiosis networks and replicate pilot schemes at local and regional levels.

The model will be based on three sustainability pillars:

- environmental (with a lifecycle approach)
- economic
- social (taking into consideration stakeholders engagement and impact on society).

Our ambition is that the model we create can be replicated in other regions and other value chain scenarios.

FISSAC aims to demonstrate the effectiveness of the processes, services and products at different levels.

[FISSAC PLATFORM](#)



### 08 FISSAC takes part in Clustering and Networking event organised by the EASME

Dec 2015

Brussels, Belgium

Acciona Infrastructures, the FISSAC project coordinator was invited to participate in the Clustering and Networking event 'Boosting synergies on EU WASTE' organised by the Executive Agency for Small and Medium Enterprises (EASME) on December 8th 2015. The discussion was joined by a number of coordinators of diverse European funded projects in the field of WEEE and raw materials, construction and building sectors, urban mining, and industrial competitiveness.

Figure 3: FISSAC website

## 5 DMP Implementation

The organizational structure of the FISSAC project was created in order to address an effective project direction and management through the communication flow and methods for reporting, monitoring, management of intellectual properties, background and foreground generated among the project. Moreover, according to Project Quality Assurance Plan to be developed (see WP 10 management), communication aspects and information generated in the project will be monitored taking also into consideration management of gender equality and risks analysis regarding financial, legal, administrative and technical co-ordination and mitigation actions aspects. If new risks appear along the project, new mitigation actions will be launched.

The FISSAC project is partly coordinated by the Scientific and Technical Committee and Innovation Management Committee. The project has a structured governance and management framework that controls and directs decisions during the project. This is organised as shown in Figure 4 below. The DMP is issued as project deliverable D10.3 under the work package 10 and will be administrated by the Technical Coordination as shown in Figure 3 below.

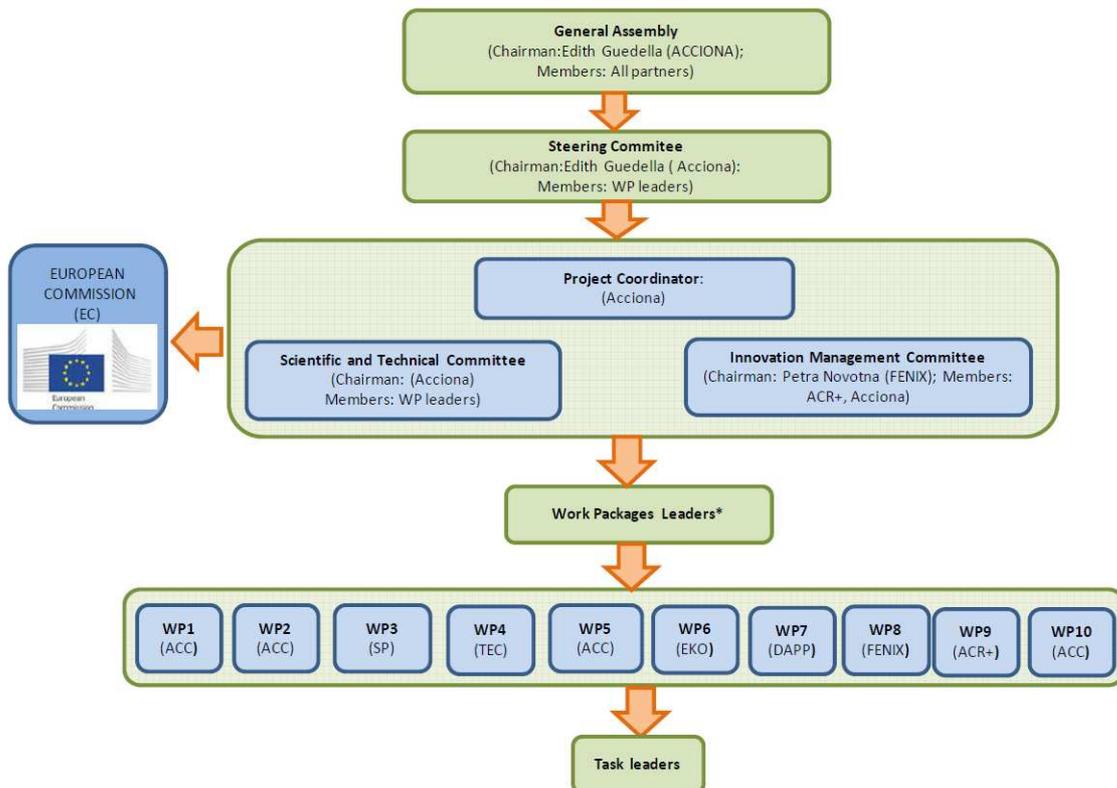


Figure 4: FISSAC organization structure

Table 1: FISSAC project partners and their roles

| Partner short name | Partner legal name  | Partner role in FISSAC project  |
|--------------------|---|---|
| 1. ACC             | ACCIONA INFRAESTRUCTURAS S.A.   | Project coordinator, participating in the development and demonstration of FISSAC implemented technologies and FISSAC model.  |
| 2. ACR+            | ASSOCIATION DES CITES ET DES REGIONS POUR LE RECYCLAGE ET LA GESTION DURABLE DES RESSOURCES | Dissemination leader, Stakeholders network, analysis of IS model and social aspects.  |
| 3. AEN             | ASOCIACION ESPAÑOLA DE NORMALIZACION Y CERTIFICACION  | Standardization tasks   |
| 4. CSIC            | AGENCIA ESTATAL CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS                             | Re-formulation of ceramic tiles composition and determination of measurable reduction of raw materials consumption by introducing waste in the ceramic tiles composition formula, participation in the design of new materials able to provide practical demonstration of FISSAC implemented technologies and FISSAC model.   |
| 5. AKG             | AKG GAZBETON ISLETMELERI SANAYI VETICARETCARET AS   | Participation in the development of new products based on secondary raw materials and demonstration of FISSAC implemented technologies and products.  |
| 6. BEF             | BEFESA SALZCHALACKE GMBH  | Active industrial partner as secondary raw material supplier.   |
| 7. BGM             | BRITISH GLASS MANUFACTURERS CONFEDERATION LIMITED   | Contribution to IS replicability activities and social issues.  |
| 8. CBI             | CBI Betonginstitutet AB   | Contribution in pre-industrial demonstration and real scale demonstration.  |
| 9. CSM             | CENTRO SVILUPPO MATERIALI SPA   | Contribution in eco-design and certification activities.  |
| 10. DAP            | D'APPOLONIA SPA   | Participation in development of the software platform, FISSAC methodology and business model for IS, and will lead demonstration of the replication of FISSAC model.  |
| 11. EKO            | EKODENGE MUHENDISLIK MIMARLIK DANISMANLIK TICARET ANONIM SIRKETI                            | Development of the software platform tool.  |
| 12. FAB            | FUNDACION AGUSTIN DE BETANCOURT   | Participation in the development and demonstration of FISSAC implemented technologies and products.   |
| 13. FEN            | FENIX TNT SRO   | Exploitation leader, business modelling, IPR management, Data Management.   |
| 14. FER            | FERALPI SIDERURGICA S.p.A.  | Active industrial partner as secondary raw material supplier.   |
| 15. GEO            | GEONARDO ENVIRONMENTAL TECHNOLOGIES LTD   | Participation in developing the software platform tool.   |
| 16. GTS            | GLASS TECHNOLOGY SERVICES LIMITED   | Active R&D partner as secondary raw material supplier.  |
| 17. TRI            | INGENIEURBUERO TRINIUS GMBH   | Eco-design and certification activities.  |
| 18. HIF            | HIFAB AB  | Contribution in the demonstration of the replication of FISSAC model, exploitation & business model for IS.   |
| 19. KER            | KERABEN GRUPO SA  | Participation in the development of new products based on secondary raw materials and demonstration of FISSAC implemented technologies and products.  |
| 20. OVA            | OPENBARE VLAAMSE AFVALSTOFFENMAATSCHAPPIJ   | Member of ACR+. As a competent (regional) government body with experience in the development and follow-up of policies, business models, partnerships offers insight and steering during the research process.  |
| 21. RIN            | RINA SERVICES SPA   | Contribute in Environmental Technology Verification tasks.  |
| 22. SP             | SP SVERIGES TEKNISKA FORSKNINGINSTITUT AB   | Eco-design and certification activities leader, LCA and LCC methods, responsible for ecological and economic evaluation of the developed processes. Evaluation of non-technical opportunities and obstacles for different business models in order to create better instruments and development towards greater sustainability. Contribution with the analysis of circular business models. |
| 23. SYM            | SIMBIOSY SIMBIOSI INDUSTRIAL SL   | Demonstration of the replication of FISSAC model, exploitation & business model for IS, IS model trends.  |
| 24. TCM            | TURKIYE CIMENTO MUSTAHSILLERI BIRLIGI   | Participation in the development of new products based on secondary raw materials and demonstration of FISSAC implemented technologies and products.  |
| 25. TEC            | FUNDACION TECNALIA RESEARCH & INNOVATION  | Active R&D partner participating in setting the basis for the IS concerning innovative solutions for the use of by-products of steel and ceramic industries in environmental-friendly products and efficient applications for the construction sector. Validation at pre-industrial scale to demonstrate the efficiency of the solutions and products.                                      |
| 26. VAN            | VANNPLASTIC LTD   | Participation in the development of new products based on secondary raw materials and demonstration of FISSAC implemented technologies and products.  |

## 6 Research data

'Research data' refers to information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion, or calculation. In a research context, examples of data include statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images. The focus is on research data that is available in digital form.

### 6.1. Characteristics for datasets produced in the project

As indicated in the Guidelines on Data Management in Horizon 2020 (European Commission, Research & Innovation, October 2015), scientific research data should be easily:

#### 1. DISCOVERABLE

The data and associated software produced and/or used in the project should be discoverable (and readily located), identifiable by means of a standard identification mechanism (e.g. Digital Object Identifier).

#### 2. ACCESSIBLE

Information about the modalities, scope, licenses (e.g. licencing framework for research and education, embargo periods, commercial exploitation, etc.) in which the data and associated software produced and/or used in the project is accessible should be provided.

#### 3. ASSESSABLE and INTELLIGIBLE

The data and associated software produced and/or used in the project should be easily assessable for and intelligible to third parties in contexts such as scientific scrutiny and peer review (e.g. the minimal datasets are handled together with scientific papers for the purpose of peer review, data is provided in a way that judgments can be made about their reliability and the competence of those who created them).

4. USEABLE beyond the original purpose for which it was collected The data and associated software produced and/or used in the project should be useable by third parties even long time after the collection of the data (e.g. the data is safely stored in certified repositories for long term preservation and curation; it is stored together with the minimum software, metadata and documentation to make it useful; the data is useful for the wider public needs and usable for the likely purposes of non-specialists).

#### 5. INTEROPERABLE to specific quality standards

The data and associated software produced and/or used in the project should be interoperable allowing data exchange between researchers, institutions, organisations, countries, etc.

## 7 FISSAC Data Sets

### 7.1 Collection and Management of FISSAC Data Sets

#### Types of data

The types of data to be included within the scope of the FISSAC Data Management Plan shall as a minimum cover the types of data that is considered complementary to material already contained within declared project deliverables.

#### Data Collection & Definition

The responsibility to define and describe all non-generic data sets specific to an individual work package shall be with the WP leader. The WP leader shall formally review and update the data sets related to his WP on a six-monthly basis. All modifications/ additions to the data sets shall be provided to the FISSAC Coordinator (ACCIONA) for inclusion in the DMP, and shall be prepared in accordance with the metadata capture table template contained in Appendix 2.

*Table 2: Forecast of FISSAC datasets related to each WP*

| WP num. | WP name   | WP leader | Dataset reference | Dataset name         |
|---------|---|-----------|-------------------|----------------------|
| WP1     | FROM CURRENT MODELS OF INDUSTRIAL SYMBIOSIS TO A NEW MODEL                      | ACC       | FISSAC_WP1        | INDUSTRIAL SYMBIOSIS |
| WP2     | CLOSED LOOP RECYCLING PROCESSES TO TRANSFORM WASTE INTO SECONDARY RAW MATERIALS | ACC       | FISSAC_WP2        | RECYCLING PROCESSES  |
| WP3     | PRODUCT ECO-DESIGN AND CERTIFICATION  | SP        | FISSAC_WP3        | ECO-DESIGN           |
| WP4     | PRE-INDUSTRIAL SCALE DEMONSTRATION OF THE RECYCLING                             | TEC       | FISSAC_WP4        | PREINDUSTRIAL DEMO   |
| WP5     | INDUSTRIAL PRODUCTION & REAL SCALE DEMONSTRATION                                | ACC       | FISSAC_WP5        | REAL SCALE DEMO      |
| WP6     | FISSAC MODEL FOR INDUSTRIAL SYMBIOSIS   | EKO       | FISSAC_WP6        | FISSAC MODEL         |
| WP7     | INDUSTRIAL SYMBIOSIS REPLICABILITY AND SOCIAL ISSUES                            | DAPP      | FISSAC_WP7        | REPLICABILITY        |
| WP8     | EXPLOITATION AND BUSINESS MODELS FOR INDUSTRIAL SYMBIOSIS                       | FEN       | FISSAC_WP8        | EXPLOITATION         |
| WP9     | DISSEMINATION   | ACR+      | FISSAC_WP9        | DISSEMINATION        |
| WP10    | MANAGEMENT  | ACC       | FISSAC_WP10       | MANAGEMENT           |

#### Data set reference and name

All data sets within this DMP have been given a unique field identifier and are listed in the table contained in Appendix 1.

#### Data Set Description

A data set is defined as a structured collection of data in a declared format. Most commonly a data set corresponds to the contents of a single database table, or a single statistical data matrix, where every column of the table represents a particular variable, and each row corresponds to a given member of the data set in question. The data set may comprise data for one or more fields. For the purposes of this DMP data sets have been defined by generic data types that are considered applicable to the FISSAC project. For each data set, the characteristics of the data set have been captured in a tabular format as enclosed in Appendix 1.

## Standards & Metadata

Metadata is defined as “data about data”. It is “structured information that describes, explains, locates, and facilitates the means to make it easier to retrieve, use or manage an information resource”. This is especially relevant in the distributed data network environment that exists within FISSAC. Meta Data shall be considered as the formal means by which data is defined and by which the meaning of information is established. All data-sets generated within the project shall be defined such that “data about data” is specified.

Metadata can be categorised in three types:

- Descriptive metadata describes an information resource for identification and retrieval through elements such as title, author, and abstract.
- Structural metadata documents relationships within and among objects through elements such as links to other components (e.g., how pages are put together to form chapters).
- Administrative metadata manages information resources through elements such as version number, archiving date, and other technical information for the purposes of file management, rights management and preservation.

There are a large number of metadata standards which address the needs of particular user communities. More details about these standards can be found in Annex 3.

## Data Sharing

During the period when the project is live the sharing of data shall be defined by the configuration rules defined in the access profiles for the project participants as described in the FISSAC Quality Assurance Plan (D10.2). Each individual project data set item shall be allocated a 3 character “dissemination classification” for the purposes of defining the data sharing restrictions. The classification shall be an expansion of the system of confidentiality applied to deliverables reports provided under the FISSAC Grant Agreement.

PU: Public (data can be shared outside the consortium without restriction)

CO: Confidential, only for members of the consortium (including the Commission Services)

CI: Classified, as referred to in Commission Decision 2001/844/EC

The three above levels are linked to the “Dissemination Level” specified for all FISSAC deliverables.

All material designated with a PU dissemination level shall be deemed uncontrolled.

Data will be shared when the related deliverable or paper has been made available at an open access repository. The normal expectation is that data related to a publication will be openly shared. However, to allow the exploitation of any opportunities arising from the raw data and tools, data sharing will proceed only if all co-authors of the related publication agree. The Lead author is responsible for getting approvals and then sharing the data and metadata on Zenodo ([www.zenodo.org](http://www.zenodo.org)), a popular repository for research data. The Lead Author will also create an entry on OpenAIRE ([www.openaire.eu](http://www.openaire.eu)) in order to link the publication to the data.

OpenAIRE is a service that implements the Horizon 2020 Open Access mandate for publications and its Open Research Data Pilot and may be used to reference both the publication and the data. A link to the OpenAIRE entry will then be submitted to the FISSAC Website Administrator (ACR+) by the Lead Author.

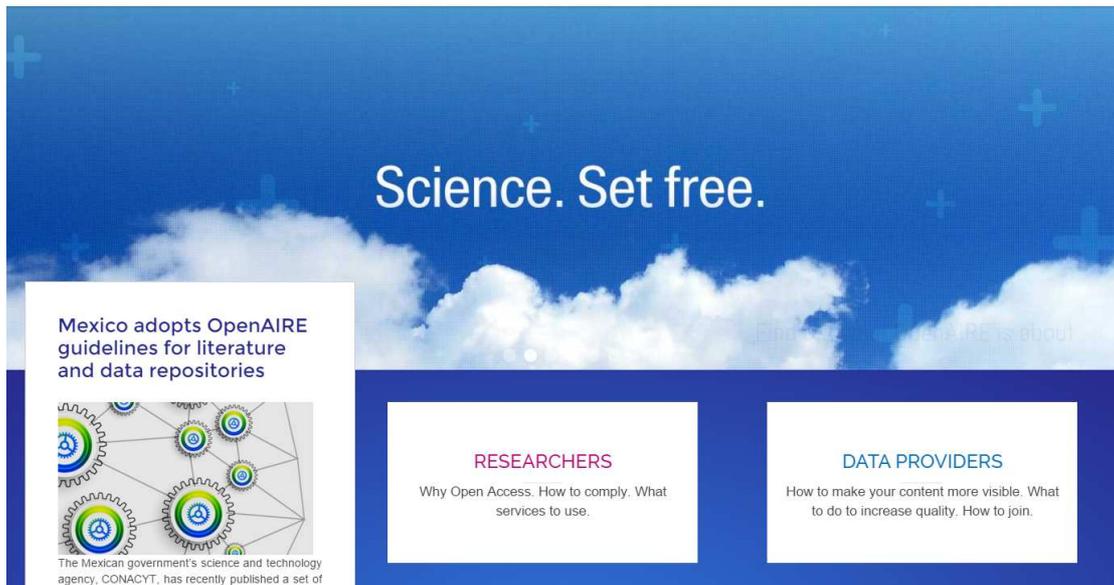


Figure 5: OpenAIRE website

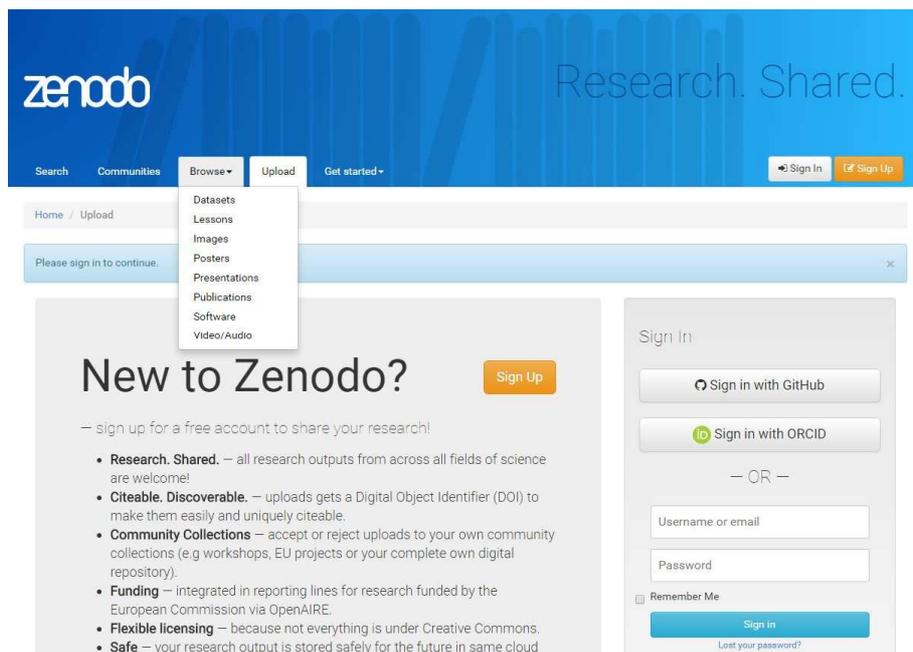


Figure 6: ZENODO repository

### Data archiving and preservation

Both Zenodo and OpenAIRE are purpose-built services that aim to provide archiving and preservation of long-tail research data. In addition, the FISSAC website, linking back to OpenAIRE, is expected to be available for at least 2 years after the end of the project. At the formal project closure all the data material that has been collated or generated within the project and classified for archiving shall be copied and transferred to a digital archive.



The document structure and type definition will be preserved as defined in the document breakdown structure and work package groupings specified. At the time of document creation the document will be designated as a candidate data item for future archiving. This process is performed by the use of codification within the file naming convention (see Section 8). The process of archiving will be based on a data extract performed within 12 weeks of the formal closure of the FISSAC project.

The archiving process shall create unique file identifiers by the concatenation of “metadata” parameters for each data type. The metadata index structure shall be formatted in the metadata order as listed in Appendix 1. This index file shall be used as an inventory record of the extracted files, and shall be validated by the associated WP leader.

## 8 Data Sets Technical Requirements

### 8.1 General requirements

The applicable data sets are restricted to the following data types for the purposes of archiving. The technical characteristics of each data set are described in the following sections. The copy rights with respect to all data types shall be subject to IPR clauses in the GA, but shall be considered to be royalty free.

### 8.2 Prohibited file types

The use of file compression utilities, such as “WinZip” is prohibited. No data files shall be encrypted.

### 8.3 Static Graphical Images

Graphical images shall be defined as any digital image irrespective of the capture source or subject matter. Images should be composed such to contain only objects that are directly related to FISSAC activity and do not breach IPR of any third parties.

#### Image file formats

Image file formats are the standardised means of organising and storing digital images. Image files are composed of digital data and can consist be of two primary formats of “raster” or “vector”. It is necessary to represent data in the rastered state for use on a computer displays or for printing. Once rasterised, an image becomes a grid of pixels, each of which has a number of bits to designate its colour equal to the colour depth of the device displaying it. The FISSAC project shall only use raster based image files of one of the two formats described below and shall be selected based on the technical needs and the format characteristics described below. The two allowable static image file formats are JPEG and PNG (detailed description in Annex 4).

#### Image file sizes & file compression

There is normally a direct positive correlation between image file size and the number of pixels in an image, the colour depth, or bits per pixel used in the image. Compression algorithms can create an approximate representation of the original image in a smaller number of bytes that can be expanded back to its uncompressed form with a corresponding decompression algorithm. Considering different compressions, it is common for two images of the same number of pixels and colour depth to have a very different compressed file size. With some compression formats, images that are less complex may result in smaller compressed file sizes. This characteristic sometimes results in a smaller file size for some lossless formats than lossy formats. The use of compression tools shall not be used unless absolutely necessary. A digitally stored image has no inherent physical dimensions. Some digital file formats record a DPI value, or more commonly a PPI (pixels per inch) value, which is to be used when printing the image. This number provides information to establish the printed image size, or in the case of scanned images, the size of the original scanned object.

Resolution refers to the number of pixels in an image. Resolution can be expressed by the width and height of the image as well as the number of pixels in the image. For example, an image that is 2048 pixels wide and 1536 pixels high (2048X1536) contains 3,145,728 pixels. As the megapixels in the pickup device increases so does the possible maximum size image that can be produced. File size is determined by the number of pixels. The image default sizes and resolution shall be as shown in Table 1. The image default size shall be A4.

Table 3: Image default sizes and resolution

| PPI | Pixels      | mm       | Paper size | Size (Greyscale) | Size (RGB) |
|-----|-------------|----------|------------|------------------|------------|
| 300 | 11114x14008 | 840x1186 | A0         | 155.7MB          | 467MB      |
| 300 | 7016x11114  | 594x840  | A1         | 78MB             | 234MB      |
| 300 | 4961x7016   | 420x594  | A2         | 34.8M            | 104.4MB    |
| 300 | 3508x4961   | 297x420  | A3         | 17.4MB           | 52.2MB     |
| 300 | 2480x3508   | 210x297  | A4         | 8.7MB            | 26.1MB     |
| 300 | 1748x2480   | 148x210  | A5         | 4.3MB            | 13MB       |
| 300 | 1240x1748   | 105x148  | A6         | 2.2MB            | 6.5MB      |
| 300 | 874x1240    | 74x105   | A7         | 1.08MB           | 3.25MB     |
| 300 | 614x874     | 52x74    | A8         | 0.54MB           | 1.6MB      |

## 8.4 Animated graphical image

Graphic animation is a variation of stop motion and possibly more conceptually associated with traditional flat cell animation and paper drawing animation, but still technically qualifying as stop motion consisting of the animation of photographs (in whole or in parts) and other non-drawn flat visual graphic material. The two allowable animated graphical image file formats are AVI and MPEG (detailed description in Annex 4). The WP leader shall determine the most suitable choice of format based on equipment availability and any other factors.

## 8.5 Audio data

An audio file format is a file format for storing digital audio data on a computer system. The bit layout of the audio data (excluding metadata) is called the audio coding format and can be uncompressed, or compressed to reduce the file size, often using lossy compression. The data can be a raw bitstream in an audio coding format, but it is usually embedded in a container format or an audio data format with defined storage layer. Detailed description of audio data types is in Annex 4.

## 8.6 Textual data

A text file is structured as a sequence of lines of electronic text. These text files shall not contain any control characters including end-of-file marker. In principle the least complicated form of textual file format shall be used as the first choice. Detailed description of textual data types is in Annex 4.

## 8.7 Numeric data

Numerical Data is information that often represents a measured physical parameter. It shall always be captured in number form. Other types of data can appear to be in number form i.e. telephone number, however this should not be confused with true numerical data that can be processed using mathematical operators.

## 8.8 Process and test data

Standard Test Data Format (STDF) is a proprietary file format originating within the semiconductor industry for test information, but it is now a Standard widely used throughout many industries. It is a commonly used format produced for/by automatic test equipment (ATE). STDF is a binary format, but can be

converted either to an ASCII format known as ATDF or to a tab delimited text file. Software tools exist for processing STDF generated files and performing statistical analysis on a population of tested devices. FISSAC innovation development shall make use of this file type for system testing.

## 8.9 Microsoft Office Application Suite

FISSAC participants shall use the currently MS supported operating system and convert from any previous obsolete releases.

### Microsoft Office Application Data files

The types of specific applications available within the current Microsoft Windows operating system shall be used to support all project activities in preference to any other software solutions. The data file types associated with these applications shall be saved in the default format and be in accordance with the file naming convention as specified in Section 8.

### Microsoft Office Configuration

At the Microsoft Office Application level the “file properties” shall be configured using the “document properties” feature. This is accessed via “Info” dropdown within the “File” menu. The “properties” and “advanced properties” present a data entry box under the “Summary” as shown in Figure 4.

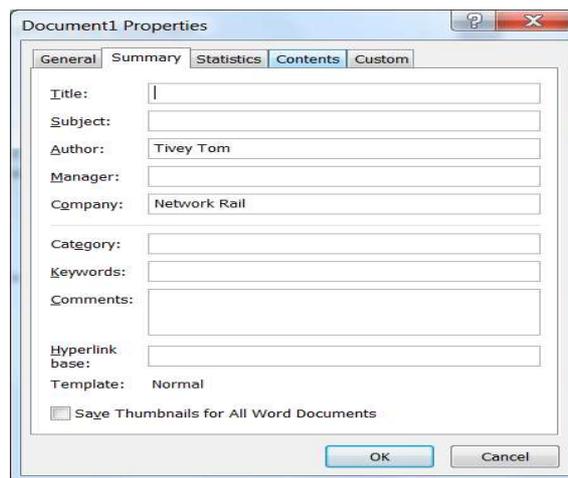


Figure 7: Data Entry Box – Summary

- Title: Duplication of the name used for the data file name
- Subject: Identifier for FISSAC work package discrimination and shall be of the following format FISSAC\_WPxx where xx is the work package number in the range 01 to 10.
- Author: Name of the person creating the document and be formatted to have the surname stated first as follows: surname\_firstname\_secondname
- Manager: Name of the author’s immediate line manager and be formatted to have the surname stated first as follows: surname\_firstname\_secondname
- Company: Company name of the author to be stated as follows: companyname\_FISSAC participant number
- Keywords: Free format text and should contain key words that would be relevant and useful to future data searches. The keywords should all be in lower case and separated with commas
- Comments: Description of file contents in free format text
- Hyperlink base: Blank
- The tickbox indicating “Save Thumbnails for All Word Documents” shall be untagged.

## 9 Naming Convention

All files irrespective of the data type shall be named in accordance with the following document file naming convention:

FISSAC\_Dx.x\_Deliverable short title\_Px\_yyyymmdd\_Status

“Dx.x”: Deliverable number according to the DoA

“Px”: Lead beneficiary number

“yyymmdd”: Year/month/day

“Status”: Short name of the last reviewer (beneficiary short name)

*Example: FISSAC\_D10.2\_Quality Plan\_P1\_20150511\_Acc*

Appendix files will be referred to the main document according to the following rule:

FISSAC\_Dx.x\_Deliverable short title\_Appx\_Px\_yyyymmdd\_Status

Where “Appx” is the Appendix letter

*Example: FISSAC\_D10.2\_Quality Plan\_AppA\_P1\_20150511\_Acc*

When the document has been approved by the EC, the status in the file name will be changed to “Final” while a copy of the file in PDF format will be uploaded on the webpage.

The file naming convention contains the 7 following sections:

[PROJECT]\_[WORKPACKAGE]\_[TASK]\_[TITLE]\_[VERSION]\_[DISSEMINATIONCLASS]\_[ARCHIVE]

Where:

- [PROJECT] is FISSAC for all document types;
- [WORKPACKAGE] is the FISSAC project work package number, with WP as a prefix;
- [TASK] is the FISSAC project task number, this is two numbers where numbers less than 10 have a leading zero;
- [TITLE] represents the description of the data item contents excluding capitalisation and punctuation characters;
- [VERSION] is the version number consisting of integer numbers only without leading zeros, prefixed with V;
- [DISSEMINATIONCLASS] is the dissemination classification allocated to a document type that define the data access post archiving, consists of the characters CO and a suffix of a single number in the range 1 to 3;
- [ARCHIVE] this is a single character defining the allocation of the data item for future archiving and is represented by a Y or N ;

## 10 Conclusions

This report contains the first release of the Data Management Plan (DMP) and represents the status of the mandatory quality requirements at the time of deliverable D10.3.

This first version of the DMP establishes the measures for promoting the findings during the project's life. The DMP enhances and ensures relevant project's information transferability and takes into account the restrictions established by the Consortium Agreement. In this framework, the DMP sets the basis for both Dissemination Plan and Exploitation Plan. The first version of the DMP is delivered at M6; later the DMP will be monitored and updated in parallel with the different versions of Dissemination and Exploitation Plans (the progress of the implementation of DMP will be included in the Project Progress Reports, at M18 and M36).

This report should be read in association with all the referenced documents, appendix material and including the EC Grant /Consortium Agreement, annexes and guidelines.

## 11 References

2 <http://www.dcc.ac.uk/resources/briefing-papers/standards-watch-papers/what-are-metadata-standards>

Guidelines on Data Management in Horizon 2020:

[https://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)



## 12 Appendices

### A1 Data Types Metadata Parameters

Metadata summarises basic information about data, which can make future data analysis more effective and efficient. The following table contains metadata definitions for the allowable data types in the FISSAC project.

| Metadata Fields       | JPEG/<br>JFIF | PNG | AVI | WAV | AIFF | MPEG | ASCII | MIME | TXT |
|-----------------------|---------------|-----|-----|-----|------|------|-------|------|-----|
| Author                | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Client ID             |               |     |     |     |      |      |       |      |     |
| Comments              | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Company               | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Date accessed         | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Date archived         | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Date completed        |               |     |     |     |      |      |       |      |     |
| Date created          | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Date last modified    | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Date last saved       |               |     |     |     |      |      |       |      |     |
| Description           |               |     |     |     |      |      |       |      |     |
| Device category       |               |     |     |     |      |      |       |      |     |
| Document ID           |               |     |     |     |      |      |       |      |     |
| File description      |               |     |     |     |      |      |       |      |     |
| File format           | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| File format version   | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| File name             |               |     |     |     |      |      |       |      |     |
| File version          |               |     |     |     |      |      |       |      |     |
| Folder path           |               |     |     |     |      |      |       |      |     |
| Height                | ✓             | ✓   |     |     |      |      |       |      |     |
| Horizontal resolution |               |     |     |     | ✓    | ✓    |       |      |     |
| Keywords              | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Language              |               |     |     |     |      |      |       |      |     |
| Length                |               |     |     |     |      |      |       |      |     |
| Line manager          | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Name                  | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Owner                 | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Priority              |               |     |     |     |      |      |       |      |     |
| Read status           |               |     |     |     |      |      |       |      |     |
| Recording time        |               |     |     |     | ✓    | ✓    |       |      |     |
| Sensitivity           | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Sharing status        |               |     |     |     |      |      |       |      |     |
| Size                  | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Source                |               |     |     |     |      |      |       |      |     |
| Status                |               |     |     |     |      |      |       |      |     |
| Subject               | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Tags                  |               |     |     |     |      |      |       |      |     |
| Title                 | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| Type                  | ✓             | ✓   | ✓   | ✓   | ✓    | ✓    | ✓     | ✓    | ✓   |
| URL                   |               |     |     |     |      |      |       |      |     |
| Vertical resolution   |               |     |     |     |      | ✓    |       |      |     |
| Width                 | ✓             | ✓   |     |     |      |      |       |      |     |

## A2 Metadata Definition Templates

Any new data types introduced shall be submitted for updates to this DMP on the form shown below.

| Metadata Definition Form |        |                    |          |                |        |
|--------------------------|--------|--------------------|----------|----------------|--------|
| Data Set Name            | [Name] | Data Set Reference | [DMPxxx] | WP(s) involved | [WPxx] |
| Data Set Description     |        |                    |          |                |        |
| Standards & Metadata     |        |                    |          |                |        |
| Data Sharing             |        |                    |          |                |        |
| Archiving & Preservation |        |                    |          |                |        |

## A3 Metadata standards

The first three profiled below primarily support discovery and access. They are progressively more complex to implement and more specialised to particular domains. The last, PREMIS, has been developed specifically to support digital preservation activities.

### Dublin Core Metadata Element Set

The Dublin Core Metadata Element Set (ISO Standard 15836) is a basic standard which can be easily understood and implemented and as such is one of the best known metadata standards. It was originally developed, in 1995, as a core set of elements for describing the content of web pages and enabling their search and retrieval. The Dublin Core Metadata Element Set consists of 15 elements which address the most basic descriptive, administrative and technical elements required to uniquely identify a digital resource. The emphasis is now on supporting resource discovery across domains. The Dublin Core Metadata Initiative develops and maintains a suite of inter-related standards. It coordinates a number of working groups who collaborate to develop a metadata registry which supports extended and qualified profiles of Dublin Core, tailored to the needs of a number of different communities or functions, e.g. Dublin Core Collection Description Application Profile (for describing whole collections) and Dublin Core Library Application Profile (for describing published library holdings). Most resource discovery metadata standards can be mapped to the Dublin Core Metadata Element Set, enabling basic federated searching across metadata created using a number of different standards, without detracting from richer metadata held elsewhere. A draft specification for expressing Dublin Core in XML is available from the Dublin Core Metadata Initiative.

### e-GMS (e-Government Metadata Standard)

The UK Government is committed to enabling consistency across public sector information and providing better access to public services. As part of this commitment they have developed e-GMS, a metadata standard for government information resources, to enable consistency across government and public sector organisations. Its use is compulsory within the sector and is part of the wider e-GIF (the e-Government Interoperability Framework) which defines technical policies and specifications to enable interoperability and easy access to information across the sector. The standard is currently at version 3 (2004) although version 3.1 will be released soon and a complete overhaul to version 4 is planned. The 15 elements of Dublin Core makes up the core of the standard and it can be readily mapped to 5 other standards if interoperability across metadata records from other disciplines is required. The further 10 elements take account of records management functions, Data Protection and Freedom of Information legislation and basic preservation information. A cut down version of the standard, e-GMS for websites (currently at version 3), is available for those creating metadata for websites.

### ISO 19115: 2003(E) — Geographic Information: Metadata

ISO 19115 was developed by the geospatial community to address specific issues relating to both the description and the curation of spatial data. This complex metadata standard can be used for describing digital or physical objects or datasets which have a spatial dimension. There are over 400 elements in the Data Dictionary, which are divided into 14 metadata packages. Each package supports a particular function, some are specific to spatial data and some deal



with general description and data curation issues. Abstract models written in UML (Unified Modelling Language) are provided for most of the packages to help the implementer understand how the elements interrelate. The standard also includes methodologies for creating application profiles, metadata extensions and hierarchical metadata and provides implementation examples. Geospatial professionals have developed a number of profiles of this standard to fit particular uses. One of these is UK GEMINI which defines an element set for discovery level metadata. It is also compliant with e-GMS and was developed collaboratively by the UK Association of Geographic Information (AGI) and the Cabinet Office e-Government Unit.

#### PREMIS: Data Dictionary for Metadata Preservation

The Preservation Metadata: Implementation Strategies (PREMIS) international working group was set up by OCLC and RLG in 2003 to define a core set of preservation metadata elements, which could be applied broadly across the preservation community, and to examine a number of practical application issues. In 2005 the group published their final report which included version 1 of the PREMIS Data Dictionary. The accompanying XML schema allows PREMIS compliant metadata to be expressed consistently in XML. PREMIS is rapidly gaining community acceptance and is maintained by the Library of Congress. It won the Digital Preservation Coalition's 2005 Digital Preservation Award. The PREMIS data model builds on the Open Archival Information System (OAIS) Reference Model (ISO 14721), and defines relationships between five digital preservation activities which it calls entities: Intellectual Entities, Objects (divided into three types: representation, file and bitstream), Events, Agents and Rights. 108 sub-entities and further qualifiers are defined for describing preservation activities of the latter four entities. Only 8 of these are mandatory. The PREMIS Data Dictionary's scope is restricted to the digital preservation activities of: maintaining viability, renderability, understandability, authenticity and identity. It assumes metadata will be auto-generated as much as possible. Implementers are expected to use other applicable metadata standards to describe Intellectual Entities, the characteristics of Agents, Rights relating to access and/or distribution, details of media and hardware, and the business rules of a repository.<sup>2</sup>

## A4 Data Sets technical requirements

### Static graphical image

The JPEG (**Joint Photographic Experts Group**) images are stored in the JFIF (JPEG File Interchange Format) file format. This format supports eight-bit grayscale images and 24-bit colour images (eight bits each for red, green, and blue). It should be noted that the JPEG applies lossy compression to images, which can result in a significant reduction of the file size and applications can determine the degree of compression to apply, however the degree of compression affects the visual quality of the result and it should be noted that repeated editing and compression can have significant impact of image quality.

The PNG (**Portable Network Graphics**) file format was developed as a free, open-source alternative to file formats and generically supports eight-bit paletted images (with optional transparency attributes for all palette colours) and 24-bit tricolor (16 million colours) or 48-bit tricolor definition.

Compared to JPEG, PNG is particularly beneficial when the image has large, uniformly coloured areas and shall be specifically used during image editing process because of its lossless compression properties. PNG is designed to work well in online viewing applications such as web browsers and can be fully streamed with a progressive display option. PNG is robust, providing both full file integrity checking and simple detection of common transmission errors. PNG shall therefore be used where there is a strong likelihood that the image will be incorporated in an internet context.

### Animated graphical image

**Audio Video Interleaved (AVI)** is a multimedia container format introduced by Microsoft in November 1992 as part of its Video for Windows software. AVI files can contain both audio and video data in a file container that allows synchronous audio-with-video playback. AVI files also use the file format extensions subsequently developed by "Matrox" and are supported by Microsoft under the designation "AVI 2.0".

**MPEG-1 or MPEG-2 Audio Layer III** is an audio coding format for digital audio which uses a form of lossy data compression. It is a common audio format for audio streaming and storage, as well as a de facto standard of digital audio compression. The use of lossy compression is designed to greatly reduce the amount of data required to

<sup>2</sup> <http://www.dcc.ac.uk/resources/briefing-papers/standards-watch-papers/what-are-metadata-standards>

represent the audio recording. An MP3 file can be constructed at higher or lower bit rates, with resulting higher or lower quality. For this reason it is a non-preferred solution for combined audio-visual data capture.

### Audio data

**Waveform Audio File Format (WAV)** is a Microsoft and IBM audio file format standard for storing an audio bitstream on PCs. It is an application of the Resource Interchange File Format (RIFF) bitstream format method for storing data. It is the main format used on Windows systems for raw and uncompressed audio. The usual bitstream encoding is the linear pulse-code modulation (LPCM) format.

**Audio Interchange File Format (AIFF)** is used for storing sound data for computers and other electronic audio devices. The format was developed by Apple and is most commonly used on this type of computer systems. Most AIFF files are uncompressed pulse-code modulation (PCM). These type of AIFF files use significantly more disk space than lossy formats like MP3 - about 10 MB for one minute of stereo audio at a sample rate of 44.1 kHz and a bit depth of 16 bits. Consideration shall be given to data requirements where there are quality and/or data volume compromises to be made.

### Textual data

The **ASCII** standard allows ASCII-only text files to be freely interchanged and readable on Unix, Macintosh, Microsoft Windows, DOS, and other systems. These differ in their preferred line ending convention and their interpretation of values outside the ASCII range (their character encoding).

The file format **.txt** is for files consisting of text usually containing very little formatting (e.g., no bolding or italics). The precise definition of the .txt format is not specified, but shall be matched to the format accepted by the system in question or a simple text editor. Files with the .txt extension can be considered as the least complex, and hence can easily be read or opened by any program that reads text and are considered universal (i.e. platform independent). The ASCII character set is the most common format for English-language text files, and is generally assumed to be the default file format.

Text files usually have the Multipurpose Internet Mail Extensions (**MIME**) type "text/plain", usually with additional information indicating an encoding. Under the Microsoft Windows operating system, a file is regarded as a text file if the suffix of the name of the file (the "extension") is ".txt". However, many other suffixes are used for text files with specific purposes. For example, source code for computer programs is usually kept in text files that have file name suffixes indicating the programming language in which the source is written.

## A5 Definitions

**Royalty-free Access Rights** - refers to the right to use copyrighted material or intellectual property without the need to pay royalties or license fees for each use or per volume sold, or some time period of use or sales.

**Affiliated Entities** - an affiliated entity is:

1. an organization that has a legal existence separate from the University, and
2. which is formed or operated to support or complement the mission of the University, and
3. (a) was created by the University, or (b) is controlled or strongly influenced by the University, or (c) receives significant financial support from the University, or (d) uses University resources, name, or identity.