



ARCHES

H2020 - 693229

Deliverable D1.2 "Data Management Plan – 1st version"

Lead Author: Javier Gutiérrez Meana, Anne Fay With contributions from: All partners

Reviewers: Jonathan Rix

Deliverable nature:	R: Report
Dissemination level: (Confidentiality)	PU: Public
Contractual delivery date:	03-2017
Actual delivery date:	03-2017
Version:	1.1
Total number of pages:	36
Keywords:	Data Management Plan, accessibility, interoperability, re-use, security, ethics



Abstract

This deliverable presents the first version of the Data Management Plan (DMP) of ARCHES. It deals with the identification of research data that will be handled all along the lifespan of the project, as well as how it will be created, stored and backed up. It also addresses privacy, ethical and security issues related to data management. This includes among others who owns data and is responsible for its preservation and which data will be shared and preserved for the future.

The structure of this deliverable is based on the document "Guidelines on FAIR Data Management in Horizon 2020" published by the European Commission (EC) in July 2016 [1]. The content will continuously be updated so as to describe new datasets (if any), refine the strategy and include additional considerations before the final version is produced in September 2019.



Document Information

IST Project Number	H2020 – 693229	Acronym	ARCHES		
Full Title	Accessible Resources for Cultural Heritage EcoSystems				
Project URL	http://arches-project.eu				
Document URL	http://arches-project.eu				
EU Project Officer	Luis García Domínguez				

Deliverable	Number	D1.2	Title	Da	ta Management Plan	– 1 st version
Work Package	Number	WP1 Title Pro		Pro	oject management an	d coordination
Date of Delivery	Contractual	ntractual M6		Actual	M6	
Status	Version 1.1		final ⊠			
Nature	Prototype ☐ Report ☐ Demonstrator ☐ Other ☐ Ethics ☐ ORDP ☒			ORDP ⊠		
Dissemination level	Public ☑ Confidential □					

Authors (Partner)	Treelogic (Treelogic (Tree) – The Wallace Collection (WC)		
	Name	Javier Gutiérrez	E-mail	javier.gutierrez@treelogic.com
Responsible Author		Anne Fay		anne.fay@wallacecollection.org
Responsible Author	Dartner	Tree	Dhono	+34 985 966 136
	Partner	WC	Phone	+44 (0) 207 563 9500

Abstract (for dissemination)	This deliverable presents the first version of the Data Management Plan (DMP) of ARCHES. It deals with the identification of research data that will be handled all along the lifespan of the project, as well as how it will be created, stored and backed up. It also addresses privacy, ethical and security issues related to data management. This includes among others who owns data and is responsible for its preservation and which data will be shared and preserved for the future.
	The structure of this deliverable is based on the document "Guidelines on FAIR Data Management in Horizon 2020" published by the European Commission (EC) in July 2016 [1]. The content will continuously be updated so as to describe new datasets (if any), refine the strategy and include additional considerations before the final version is produced in September 2019.
Keywords	Data Management Plan, accessibility, interoperability, re-use, security, ethics.

		Version Log	
Issue Date	Rev. No.	Author	Change
01-03-2017		Javier Gutiérrez Meana	Template
01-03-2016		Javier Gutiérrez Meana	Structure according to the
			guidelines of the European
			Commission
10-03-2017	0.1	Javier Gutiérrez Meana	Initial contents
14-03-2017	0.2	Anne Fay	Updated contents
20-03-2017	0.3	Javier Gutiérrez Meana	Integration of partner's
20-03-2017	0.3	Javier Gutierrez Mearia	contributions
23-03-2017	0.4	Jonathan Rix	Review
28-03-2017	0.5	Javier Gutiérrez Meana	Integration of comments from
			other partners



29-03-2017	0.6	Andreas Reichinger	Minor changes
30-03-2017	1.0	Javier Gutiérrez Meana	Final version
21-06-2017	1.1	Javier Gutiérrez Meana	Website link updated on page 14
			(Thyssen)



Table of Contents

Doc	ument Information	4
Tab	le of Contents	6
List	of figures	8
List	of tables	9
Abb	previations	10
1	Introduction	11
2	Data summary	12
2.1	Purpose	12
2.2	Types and formats	12
2.3	Origin	13
2.4	Re-use	14
2.5	Expected size	15
2.6	Utility	16
3	FAIR data	18
3.1	Making data findable, including provisions for metadata	18
3.2	Making data openly accessible	18
3.3	Making data interoperable	19
3.4	Increase data re-use	20
4	Allocation of resources	21
5	Data security	22
6	Ethical aspects	2 3
6.1	Principles of participation	24
6.2	Principles of consent	24
6.3	Principles of security	24
6.4	Principles of privacy	25
6.5	Deliverables	25
7	Conclusions	26
Ref	erences	27
Ann	nex A Ethics requirements	29
A.1	Information sheets	29
	A.1.1 Information sheet for people with learning disabilities	29
	A.1.2 Information sheet for people with sensory disabilities	32



A.2	li	nformed consents	34
	A.2.1	Informed consent for people with learning disabilities	34
	A.2.2	Informed consent for people with sensory disabilities	36



List of figures

Figure 1: EDM class hierarchy [18]	. 20
Figure 2: ARCHES Information Letter for people with learning disabilities – general description	. 29
Figure 3: ARCHES Information Letter for people with learning disabilities – the research process (1/2)	. 30
Figure 4: ARCHES Information Letter for people with learning disabilities – the research process (2/2)	. 31
Figure 5: ARCHES Information Sheet for people with sensory disabilities (1/2)	. 32
Figure 6: ARCHES Information Sheet for people with sensory disabilities (2/2)	. 33
Figure 7: ARCHES Consent Form for people with learning disabilities (1/2)	. 34
Figure 8: ARCHES Consent Form for people with learning disabilities (2/2)	. 35
Figure 9: ARCHES Consent Form for people with sensory disabilities.	. 36



List of tables

Table 1: Types and formats of data generated/collected	13
Table 2: Examples of data provided by the museums	15



Abbreviations

API: Application Programming Interface

ARCHES: Accessible Resources for Cultural

Heritage EcoSystems

BERA: British Educational Research Association

BSL: British Sign Language

CA: Consortium Agreement

CH: Cultural Heritage

CIDOC: International Committee of

Documentation

CMS: Central Management Server

Coprix: (partner short name) Coprix Media

CRM: Conceptual Reference Model

DMP: Data Management Plan

DoA: Description of Action

EC: European Commission

EDM: European Data Model

EEAB: External Expert Advisory Board

FAIR: Findable, Accessible, Interoperable and Re-

usable

FLG: (partner short name) Fundación Lázaro

Galdiano

GA: Grant Agreement

H2020: Horizon 2020

ICOM: International Council of Museums

IPR: Intellectual Property Rights

KHM: (partner short name) Kunst-historisches

Museum Wien

LIDO: Lightweight Information Describing Objects

MBBAA: (partner short name) Centro Regional de

Bellas Artes de Oviedo

MN: (partner short name) Moritz Neumüller -

ArteConTacto

ORDP: Open Research Data Pilot

ORO: Open Research Online

OU: (partner short name) The Open University

SignTime: (partner short name) Sign Time GmbH

Tree: (partner short name) Treelogic Telemática y

Lógica Racional para la Empresa Europea S.L.

Thyssen: (partner short name) Fundación-

Colección Thyssen-Bornemisza

UBAH: (partner short name) University of Bath

VCS: Version Control System

VRVis: (partner short name) VRVis Zentrum für

Virtual Reality und visualisierung forschungs -

GmbH

V&A: (partner short name) Victoria & Albert

Museum

WC: (partner short name) The Wallace Collection

WP: Work Package



1 Introduction

ARCHES [2] ("Accessible Resources for Cultural Heritage Ecosystems") is a Horizon 2020 (H2020) project which aims to generate more inclusive environments at museums and cultural heritage (CH) sites, so that people with differences and difficulties associated with perception, memory, cognition and communication can easily access and understand art. To this end, ARCHES brings together three key aspects: a participative research methodology with the active involvement of the previously cited target audiences organised in the so called exploration groups; the re-use of available digital resources provided by the partners and external sources; and the development of innovative technologies that will be implemented and fine-tuned by the technology partners with the feedback from the exploration groups. As a consequence, the generation, collection, re-use and preservation of data is deemed to be crucial for the smooth running of ARCHES and will need the collaboration of all the partners within the consortium: Treelogic (Tree), University of Bath (UBAH), The Open University (OU), Sign Time GmbH (SignTime), Neumüller Moritz – ArteConTacto (MN), Centro Regional de Bellas Artes de Asturias – Fine Arts Museum of Asturias (MBBAA), Coprix Media (Coprix), VRVIs Zentrum für Virtual Reality and Visualization Forschungs (VRVis), KHM-Museumsverband -Kunsthistorisches Museum Wien (KHM), The Wallace Collection (WC), Fundación Colección Thyssen Bornemisza – Thyssen Museum (Thyssen), Fundación Lázaro Galdiano – Lázaro Galdiano Museum (FLG) and Victoria & Albert Museum (V&A).

This document, the first version of the ARCHES Data Management Plan, describes the research data that will be generated and collected during the project, as well as the data (i.e. the partner museums digital assets) that will be used throughout the project. The strategy to make data FAIR (findable, accessible, interoperable and re-usable) is commented in the following sections. Nevertheless it is worth noting that updates and more details will be added as ARCHES progresses so as to provide the guidelines and strategies to follow beyond the completion for exploiting data collected and generated within ARCHES. A final version will be released in month 36 of the project.



2 Data summary

2.1 Purpose

The purpose of the data collection/generation is the support of the ARCHES technical developments and complementary activities to validate the proposed solutions in real environments and organise demonstration activities open to everyone according to the objectives and guidelines defined in the document Description of Action (DoA). In particular, data collection and generation are crucial to achieve the following specific objectives:

- To develop and evaluate strategies which enable an exploration of the value, form and function of mainstream technologies by and for people with differences and difficulties associated with perception, memory, cognition and communication.
- To develop and evaluate the use of mainstream technologies to enable the inclusion of people with such disabilities as museums visitors and consumers of art.
- To identify sources Internet, internal archives, libraries, etc. that can provide digital cultural resources and take advantage of their possibilities to integrate this content into innovative tools, applications and functionalities.
- To validate the technological outcomes in operational environments based on a participatory research methodology consisting of three pilot exercises in museums.
- To promote the tools and applications developed in ARCHES by means of on-site demonstration activities all around Europe.

Likewise, data collection/generation is in close connection with the tasks scheduled in the work plan, especially in the context of the technical work packages – i.e., WP3 "Development of an accessible software platform", WP4 "Development of applications for handheld devices" and WP5 "On-site multisensory activities" – paving the way for the system validation and pilot exercises in WP6. Therefore, the DMP will be a living document where new datasets may be incorporated at any time all along the lifespan of the project following the principles of FAIR data.

2.2 Types and formats

ARCHES will handle different types and formats of data according to the multiple research areas and technical developments. In principle, these are grouped into four categories:

- Multimedia research data: A combination of sound, images, video and text components to represent
 the works of art in different ways and support the integration of digital surrogates into the web and
 mobile applications as well as the generation of complementary resources and tools. Photographs,
 3D models, sketches, interactive guides, avatars and serious games are also considered within this
 category.
- Metadata: Data associated to the multimedia content used to better describe and understand each
 of the works of art. This includes information about the author, style, period and other details
 regarding the history and characteristics of the paintings, sculptures and other CH assets. Subtitles
 are other form of metadata that will be generated together with the videos to become accessible to
 people with hearing difficulties.
- User data: Credentials to access the platform and applications to interact with the community as well
 as user preferences to personalise how the content is presented. These preferences will not be based



- on the identification of impairments but on the selection of the desired functionalities from a wide range of options, avoiding the management of potential sensitive information.
- **Exploration data**: Data generated for/by the exploration groups regarding their involvement in the project, from the recruitment phase in the UK, Spain and Austria to the interaction with the research teams to provide feedback on the methodology, activities and technologies. Therefore, information sheets, consent forms and multiple materials to assess the outcomes and allow expressing opinions and feelings e.g. surveys, questionnaires, interviews, etc. will be generated, stored and preserved under this category.

The above information is summarised in Table 1 with the most common formats employed by the partners when managing these datasets. We will decide the ones that best fit the depending on the input sources and the applications. This will also be aligned with the definition of system requirements and architecture in the deliverables D3.1 and D4.1.

Type of data **Description Formats** JPG, PNG, TIF, **Image** GIF, BMP High-resolution photograph TIF, PNG, RAW PLY, DXF, COLLADA, STL, Multimedia research data 3D model OBJ, PDF, U3D, **FBX** MP3, MP4, MOV, Audio and video AVI, WMV, MPG, MPEG, 3GP Subtitles, captions SRT, VTT, QT Metadata Metadata JSON, plain text Credentials **TBD User data Preferences TBD** DOC, DOCX, PDF Information sheet Informed consent DOC, DOCX, PDF **Exploration data** DOC, DOCX, XLS, Feedback XLSX, PDF, JPG, PNG, MP4, 3GP

Table 1: Types and formats of data generated/collected.

2.3 Origin

We identify four different origins for the data collected/generated in ARCHES:

Consortium partners: Information and digital CH assets mainly gathered by the six participating
museums, i.e., MBBAA, KHM, WC, Thyssen, FLG and V&A, so as to present them to the exploration
groups in Spain, Austria and the UK. This data will be obtained from their digital archives, guides,
library, etc. and exploited by means of state of the art technologies – e.g. augmented reality, avatars,
relief printers and models, context-sensitive tactile audio guides and advanced image processing
techniques.

ARCHES may also acquire new image (photographs, models, drawings) and video datasets (e.g. when enough documentation is not available or the development of accessible resources require additional



multimedia content) that can be used in the project. In these cases, image and video datasets will be produced at the participating museums using the most common file formats. Text files relating to information about museum objects and images will be generated in formats such as DOC, PDF or TXT, although it may be processed and converted to other schemes.

- Exploration groups: ARCHES will create new data based on the weekly activity of the exploration groups involved in the activities prepared for them at the six ARCHES museums. It will inform how the platforms, apps and activities are designed and developed with the purpose of improving their functionalities and accessibility. Qualitative and quantitative data resulting from diverse methods e.g. interviews and discussions with participants, along with direct observation will be collected in audio, video and written formats taking the differences and difficulties of the individuals into account as done with the consent forms.
- **Users**: In order to register and log in to the platform, they will provide a name and valid email address. The system will store the options they select for the customised presentation and visualisation to enable the corresponding functionalities.
- External sources: Repositories of CH assets of several institutions such as Europeana [3], DBpedia [4] and the Rijksmuseum [5] that publicly and freely release digital content on the Internet, useful to test the functionalities of the new tools and complement the online experience. Images, videos, audio and metadata will be collected through the corresponding Application Programming Interfaces (API). The list of CH institutions will continuously grow in parallel with the organisation of the sessions where the exploration groups evaluate the types of contents to be included in the platform and applications according to their needs and expectations.

2.4 Re-use

As stated in the DoA and described in the previous sections, a significant part of the research data will consist of existing data coming from museums, CH sites and other related sources to be used for the development of technology such as apps, websites and sensory activities (e.g. 3D reliefs). This is clearly aligned with the expected impacts listed in the work programme under the topic Reflective-6-2015 [6], where promoting "the use of digital cultural heritage allowing its reinterpretation towards the development of a new shared culture in Europe" and exploiting "the rich and diverse European digital cultural heritage in a sustainable way" are two of the cornerstones.

On the one hand, ARCHES will re-use data already stored on the Central Management Server (CMS) of each of the partner museums. For example, in case of the WC, this is MuseumPlus and eMuseumPlus. Therefore, there will be no problem in providing specific digital resources to the technological developers once these are selected by the exploration groups with the support of OU, UBAH as well as the educators and experts working in the museums. This will enable the generation of tactile images among others. In addition, some images and complementary information may be directly retrieved from the public websites, in particular when referring to masterpieces:

- MBBAA: http://www.museobbaa.com/en/collection/permanent-collection/
- KHM: https://www.google.com/culturalinstitute/beta/partner/kunsthistorisches-museum-vienna-museum-of-fine-arts?hl=en
- WC: http://wallacelive.wallacecollection.org/eMuseumPlus
- Thyssen: https://www.museothyssen.org/coleccion



- FLG: http://www.flg.es/museo/la-coleccion/bases-de-datos
- V&A: https://www.vam.ac.uk/collections

On the other hand, different options will be considered to collect and re-use already-existing digital CH resources, making use of the APIs released by the multiple institutions as pointed out in the previous section. These APIs provide direct access to a great deal of artworks in several formats, from images to 3D models together with videos, text and audio tracks. Europeana [3] – MBBAA, KHM, Thyssen and FLG are Europeana contributors – the Rijksmuseum [5] and the Finnish National Gallery [7] will be explored during the initial phase. Furthermore portals like Google Arts and Culture [8] – part of the KHM, Thyssen, FLG and V&A collections can be found here – WikiArt (The Visual Art Encyclopaedia) [9] or DBpedia [4] also make digitisations of paintings, 3D objects and associated metadata available to the general public.

The consortium will pay special attention to Intellectual Property Rights (IPRs) regarding data re-use from external and internal sources. We will respect the terms and conditions described on the corresponding websites and/or the copyright licenses under Creative Commons [10]. Similarly, neither the WC nor the V&A will allow their digital assets to be re-used for commercial purposes and, specifically, the WC will not allow any re-use of the assets by third parties for any reason without permission. The other participating museums in Spain and Austria will take a decision on how the digital content is shared once the artworks are selected by the exploration groups.

2.5 Expected size

Multiple factors should be taken into account to estimate the size of the data. Based on the types identified in section 2.2, we expect multimedia research data to represent the highest percentage of storage needs. Format, resolution, size, length and other variables will be analysed throughout the duration of ARCHES to ensure that the hardware resources are enough to store and preserve data, especially when the platform and apps become public and new institutions are sought to share their contents.

Some examples of data provided by the participating museums can be found in Table 2. Although metadata will not require much space, images (especially those employed to design 3D reliefs) have to be in high resolution for accurate results, ranging between 5 and 100 MB each. Likewise, data related to the generation of tactile reliefs can vary from many MB to several GB. In addition, the consortium is currently analysing the reproduction of 3D models with different materials, colours and dimensions taking advantage of existing or new scans. Depending on the method, the size could be several GB and up to TB.

3D object **Painting Fabric Title** Celadon dish Madame de Pompadour Strawberry thief **Photo** Place of origin Paris (France) Zheijiang (China) London (UK) 14th century **Date** 1758 1883 William Morris / Morris Artist/maker François Boucher Unknown & Co.

Table 2: Examples of data provided by the museums.



	Painting	3D object	Fabric
Materials/techniques	Oil on canvas	Stoneware, glazed	Indigo-discharged and block-printed cotton
Credit Line	Bequeathed by John Jones	Bequeathed by Mr Arthur Hurst	Given by Morris & Co.
Museum No.	487-1882	C.1-1940	T.586-1919
Gallery location	Europe 1600-1815 (Room 3)	Ceramics galleries (Room 145)	British galleries (Room 125)
Description	Madame de Pompadour was the official mistress of King Louis XV. She was also an influential patron of the arts and a leader of taste. S devoted supporter of the Sèvres porcelain factory and keen collector of Japanese lacquer, she furnished her residences with fine furniture and porcelain. []	Green-glazed stonewares from Zhejiang were the most common type of Chinese ceramics exported to the Middle East before 1400. This dish was thrown and carved before being given a thick green 'celadon' glaze, which has pooled in the incised decoration and carved fluting. []	Morris was inspired to draw this design after finding thrushes stealing fruit in his garden. This complicated and colourful pattern is printed by the indigo discharge method and took a long time to produce. Consequently, it was expensive to buy. []
Other		Audio description available: http://bit.ly/2nRAfmM	

User data should not be a problem in terms of size, since it will be restricted to contact details, username and password as well as the configuration options.

Because exploration data will mainly consist of written documents as well as photographs and videos usually taken with a smartphone, they will all account for several GB.

2.6 Utility

Collected/generated data will be useful for different target audiences:

- People with differences and difficulties associated with perception, memory, cognition and communication: Co-creation is a crucial concept in ARCHES where end users are continuously involved in the value chain from the very beginning. Multimedia research data will allow them to explore new forms of interaction with CH resources through the Internet or at the museum. Metadata will be a good support to multimedia content so as to make the available information accessible to all. Besides, in addressing the combination of state of the art technologies and digital CH assets, the exploitation and re-use of data will build bridges not only between cultures but also between the heritage communication and a historically silenced audience, traditionally sidelined by mainstream cultural and intellectual activity.
- Researchers: Social researchers may take advantage of the documents to obtain feedback from the
 participants as well as information sheets and consent forms for people with learning difficulties and
 sensory impairments. It will facilitate the communication and pave the way to replicate the approach
 and methodology in other contexts, especially in the field of education. From a technological



perspective, high-resolution photographs, 3D scans and other datasets may be of interest to design new functionalities and applications in the field of CH.

- Museums: Both multimedia research data and exploration data will allow museums to define more inclusive strategies in the future building on top of the ARCHES results. Moreover, the generated material may be employed for other purposes, such as marketing and preservation, attracting new audiences to the facilities. Likewise, the generated digital content may be integrated in their own websites and other networks to boost dissemination and the re-use of these surrogates.
- **General public**: Although specifically designed for the aforementioned target groups, access to multimedia data through the online platform and smartphone applications will be granted to all on a free basis. In addition, the outputs will be available at the participating museums at least during the open days to be organised at the end of the project.



3 FAIR data

This section deals with four key aspects within the DMP, i.e., findable, accessible, interoperable and re-usable data.

3.1 Making data findable, including provisions for metadata

The collection and generation of consistent and accurate metadata significantly contribute to the improvement of data search and the identification of digital resources. In this particular context where CH assets are the main concern, metadata will also facilitate the finding of similar content across different online collections and repositories, such as the external sources cited in the previous sections. In order to ensure that interoperability among different approaches and structures, the adoption of a common standard is recommended. Therefore, the consortium will analyse the schemas used by the participating museums as well as the most relevant actors in the CH field.

The implementation of a metadata scheme will be exploited for the search functionality in the online platform and applications for handheld devices to allow the user to find specific CH assets by date, data, author, title, origin, type of artwork, etc. Keyword-based search will also be developed to detect specific words and terms free-text fields, such as description, history and others.

Likewise, the consortium will define a unique identifier for each of the digital resources to facilitate the handling of videos, images, etc. among professionals. The methodology will enable the storage and preservation of, for example, multiple photographs of the same artwork keeping their connection for further re-use, studies and cross search.

For data that requires versioning, the consortium will take advantage of Version Control Systems (VCS). For example, source code will be maintained in GIT and SVN repositories; executable object code can be versioned on MyGet.org. Furthermore, the documents can also be versioned in repositories during design and authoring.

3.2 Making data openly accessible

The ARCHES project proposes a general strategy in which different alternatives will be developed to make data openly accessible. Aside from the envisioned platform, apps and multisensory activities where multimedia content will be available for specific actions, different repositories will contain the four different types of data generated/collected by the partners.

When referring to the final version of multimedia research data and associated metadata, a dedicated repository managed by the coordinator (Tree) will be used to deposit this information. This has already been considered in the system architecture as pointed out in the deliverable D3.1 "Report on system architecture definition". A REST-API will facilitate the sharing of the stored digital resources in an easy way by means of an HTTP browser that will call API methods.

User data will not be accessible for people outside the consortium for privacy and security reasons. Even within the consortium, only the person in charge of the technical development of the platform at Tree will be allowed to handle this data. Passwords will be hashed.

Regarding the exploration data, when it does not rise ethical and privacy issues, the consortium will benefit from the available platform Open Research Online (ORO) [11], the OU's repository of research publications and other research outputs. ORO is an open access resource that can be searched and browsed freely by the general public.



Open access to the content of the repositories will be subjected to IPRs. As a consequence, each partner will carefully assess on a case-by-case basis which results – and to what extent – can be made public. In particular, museum data may be retained by the project partners in accordance to the Grant Agreement (GA) and Consortium Agreement (CA). For example, WC digital assets, such as high-resolution images, cannot be made openly available to third parties without permission. On the contrary, other WC digital assets, such as low-resolution images and video content is already (and will be) publically available online on different websites [12], [13] and can be used by third parties for non-commercial purposes. Similarly, V&A digital assets are openly available to third parties for non-commercial purposes. The other participating museums will decide on the strategy to adopt with respect to making data openly accessible once the digital CH assets to be included in the solution are selected in the exploration sessions planned for the end of 2017 and 2018.

Technology developers will also evaluate the outputs and intermediate results that can become public since most developments may contain traces of copyrighted material (e.g. use of high-resolution photographs) and, thus, developers might not have the full copyright on the derivative outputs.

In the context of communication and dissemination actions, different activities are planned to achieve the maximum impact from the outset by taking advantage of the generated data. These are described in deliverable D7.2 "Communication plan, activities and publications – 1st version" and include a dedicated website [2] – where reports about research and technological results that are part of the public deliverables will be available – social networks, mass media, networking in different events, etc. Moreover, the consortium will encourage publication in open access journals and the adoption of green or gold models for open access. Again, ORO may be used to this end as well as the website.

3.3 Making data interoperable

The use of the most common and standardised formats for each of the data types (see Table 1) is the first step to make data interoperable with other systems and for professionals in other disciplines. Free online converters can be found browsing the web to easily obtain a compatible file.

Regarding the vocabularies and ontologies, ARCHES will proceed as outlined in section 3.1, i.e. analysing the schemas and methodologies used by the participating museums and other partners with the purpose of developing a joint strategy to ensure interoperability among their systems. The following paragraphs describe some popular schemas and how they are connected to provide data to the Europeana network [3] (the other way around is possible too). This is an initial suggestion that should be studied in depth by the consortium.

The Conceptual Reference Model (CRM) [14] developed by the International Committee of Documentation (CIDOC) of the International Council of Museums (ICOM) is a well-known and extensively used semantic model based on earlier standards. CIDOC-CRM establishes relationships among implicit and explicit concepts for CH documentation to transform isolated and inhomogeneous metadata into a valuable and coherent global resource. CRMdig [15] has extended CIDOC-CRM in the framework of the 3D-Coform project [16].

The Lightweight Information Describing Objects (LIDO) is an XML harvesting scheme intended to act as a gateway and provider of museum object metadata to online databases and repositories. Therefore, it does not replace CIDOC-CRM but builds on top of this and other data schemas. The strength of LIDO lies in "its ability to support the full range of descriptive information about museum objects" [17].

The Europeana Data Model (EDM) used by the Europeana network [3] aims to guarantee the preservation of the original data from the diverse metadata schemas, while accommodating the rich variety of community standards for museums, archives and digital libraries. In particular, LIDO is one of the standard intermediary



schemas that can be mapped to EDM in a straightforward manner. The EDM class hierarchy is presented in Figure 1.

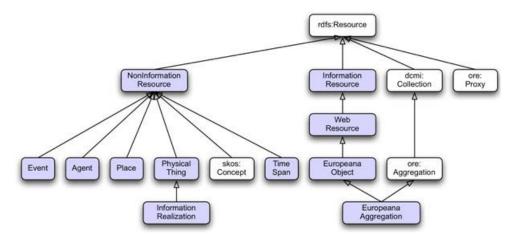


Figure 1: EDM class hierarchy [18].

3.4 Increase data re-use

ARCHES is committed to the exploitation and sharing of the outcomes generated (knowledge, technologies, data, etc.) in the framework of the activities defined in the work plan. Free applications, public deliverables, open publications and reports or access to research data are indicative of this commitment. However, all the aspects related to IPR should be thoroughly assessed for each result and addressed according to the clauses of both the GA and the CA and considering the particular interests and policies of each of the partners. Data re-use is not an exception.

The platform and smartphone applications will be kept updated and running for a period of at least two years after the completion of the project, allowing the consortium to improve the functionalities, add new content and initiate the commercial exploitation in accordance with the business plan developed in WP7. During this period, the access to the repositories will be supported as usual, so that third parties will be able to re-use data following the conditions defined and agreed by the partners.

Because multimedia research data and metadata are mostly related to the digital CH assets owned by the six museums, the re-use by third parties will be subjected to their particular business strategies. Data will be licensed following the same principles each museum apply to its digital collection available on the Internet through its website, Google Art, Europeana and other means all along the execution and after the completion of ARCHES. Outputs derived from the use of digital CH assets will also be released/protected accordingly in agreement with the technical developers (if any) involved. To this end, access to datasets in our repositories may be conditioned to the registration of the researchers and entities willing to take advantage of this content. In order to obtain the necessary credentials, they will be asked to provide a valid e-mail and additional contact details as well as some relevant information concerning the exploitation actions to be carried out. If approved by the consortium, a username and password (or alternatively an API key) will be generated.

Some designs – such as the relief printer – and tools will not be released since the owner will apply for a patent. Making the data public and allowing its re-use would be a violation that would invalidate the process. Therefore, an embargo period is foreseen in this particular situation.

The consortium will jointly work towards the achievement of the maximum quality of the results that will be tested by the exploration groups in the museums.



4 Allocation of resources

The estimated costs for making data FAIR were already considered when drafting the budget for the project before being submitted to the European Commission. The implementation of dedicated spaces in data repositories was deemed to be an important part of the approach as well as the access to the collected and generated data. Consequently, the consortium does not need to make any distinction or to consider additional resources to fulfil the expectations in this regard.

Depending on the repository, there will be different people responsible for the storage, access and curation of data:

- The contributions of ARCHES to the ORO repository of the OU will be managed by the Research Manager Prof Jonathan Rix (OU).
- The multimedia/metadata repository for the interaction with the online platform and applications for handheld devices will be managed by the Technical Manager Ms Ana Belén Rodríguez (Tree).
- Intermediate and final data to generate 3D reliefs, 3D printer design and tactile audio guide will be managed by Mr Andreas Reichinger (VRVis).

Costs associated to the long term preservation of data generated in the project will be assumed by the aforementioned partners as part of their daily operating costs.



5 Data security

All data will be kept in compliance with the "Data protection act" (1998) [19] and the "Freedom of information act" (2000) [20]. Research notes and visual records along with interview material and transcripts will be kept in secure conditions. ARCHES is registered with the Faculty of Data Protection Officer at OU. Therefore, any personal information will be kept on an OU secure server.

- OU will aim to keep collected datasets resulting from the participation of the exploration groups in the pilot exercises and validation separated from personal identity information. Any key linking codes to identity information such as names, addressed and telephone numbers would then be kept secure and separate from the dataset accessible only to the investigators. This, however, may not be acceptable to the exploration groups and the methods they choose to develop. Consequently, partners will be flexible in balancing their needs for privacy with the needs for representation. This is a key aspect of the research approach being adopted. The methodology was submitted to the Open University Data Protection Officer and was recorded on 23rd December 2016.
- Tree will deal with different types of data. In particular, passwords to log in to the user's profile will be hashed so that nobody (except for the user) can see this information. Regarding the infrastructure for multimedia research data and metadata storage, backups will be made on a daily basis following the company's policy. Access to these backups will be restricted to the persons working in the project.
- Although most data generated for the multisensory activities will be intended to be used only by the
 developers, VRVis has a reserved space on a secure data server in its facility, which only members of
 VRVis can access. An easily browse-able folder hierarchy is used: Each museum and each object has
 its own main directory. There, VRVis collects all input data and data generated during the project.



6 Ethical aspects

The research carried out in ARCHES adheres to the "Ethical guidelines for educational research" [21] edited by the British Educational Research Association (BERA) and the "Data protection act" (1998) [19]. It will also follow the OU policy documents "Ethical principles for research involving human participants" and "The code of practice for research and those conducting research". The OU will insist that all explorations groups operate under these guidelines as well as any local or national policies which are relevant. In fact, the research protocol for the ARCHES project already received a favourable opinion by the Open University Human Research Ethics Committee.

Given the possible vulnerable nature of the participants in the exploration groups, the consortium recognises the need to be sensitive to several key aspects when referring to data management and how participants get involved in the pilots and validation (e.g. information sheets and information consents), collaborate (e.g. access to data and audio visual content generation) and provide feedback (e.g. surveys and questionnaires):

- The need to be constantly alert to the potential breaches of confidentiality between exploration group members.
- Consent as an ongoing, unfolding process, particularly in relation to people with learning difficulties. This is particularly relevant given that the project is endeavouring to give these participants a research voice.
- The need to balance our desire to gather data. Our presence as observers to practices rooted in
 everyday relationships elsewhere means we will engage in discussion with relevant management or
 services at the earliest opportunity if we evidence practices which cause us concern about individual
 well-being.
- Issues with privacy in relation to subsequent use of data beyond the initial confines of the exploration groups. We regard images and other audio visual footage to be the property of the individual. If subsequently we wish to use any material, we will need to seek further specific permission.

WP8 "Ethics requirements" focuses on the activities tackling ethics in relation to the exploration groups. The university research teams (OU and UBAH) have considerable personal expertise within the field of ethics, but this will be reinforced by the presence of an independent, international ethical specialist on the External Expert Advisory Board (EEAB). The work package will be executed throughout the duration of the project to ensure the informed, consensual and secure involvement of people with intellectual and sensory impairments. In this context, the key objectives will be to:

- Develop a range of mechanisms to inform participants about the project and to assure we are working with their consent throughout.
- Maintain ethical standards in a manner and form appropriate to that laid out by national and institutional documents and committees.
- Operate with key participatory principles, agreed with the exploration groups that will underpin research practices and relationships throughout the project.
- Ensure all participants understand the importance of rigorously applying these ethical principles.

In order to achieve these objectives, ARCHES will be organised and managed in line with principles of participation, consent, security and privacy. The main aspects connected with the DMP are outlined in the following subsections (and the full version can be found in the DoA).



6.1 Principles of participation

- ARCHES is enabling the research voice of the members of the exploration groups.
- ARCHES must ensure the members of the exploration groups are active, recognised and willing participants.
- Individual members of the exploration groups will join the project on a volunteer basis and will be able to leave whenever they so wish.
- Data collection methods will also be developed in collaboration with the exploration group so they can best identify, capture and record their experiences and views.
- Diverse forms of communication must be used to engage with authentic user perspectives and the
 diverse forms of evidence that this produces must be valued and treated as significant markers of
 certainty.
- Reports will be provided twice a year to the EEAB on ethical issues, in relation to principles of participation, consent, security and privacy.

6.2 Principles of consent

- Consent and assent is an ongoing, unfolding process, to which the research teams need to be alert at all times. It will be demonstrated by engagement as well as through verbal or signed agreement.
- Consent will be made via the communication medium in which the person is most adept (verbal
 and/or augmented communication), and recorded with the person's initials (or alternative if
 necessary), witnessed by an advocate.
- The prospective members of the exploration groups need to meet the researchers before the project begins. Informed consent will be sought following this meeting. Agreement to participate will be viewed as provisional consent.
- Consent is provisional upon the research being conducted within the outlined framework, continuing
 to develop within participant expectations, and there being no adverse change in the person's ability
 to give consent.
- Participants will be encouraged to share information with people they trust, who in turn will be encouraged to ask questions.
- Consent and assent materials must be accessible to people with the range of sensory and intellectual impairments, to ensure all participants are consensually involved in the project.
- Information will be given verbally, supported by sign/symbols/illustrations and repeated on more than one occasion.
- Supporters and other professionals involved must give their informed consent to participate using the agreement form, which will be completed prior to any data collection taking place.

6.3 Principles of security

- There is a need to be constantly alert to the potential for breaches of confidentiality and trust between exploration group members, impacting upon personal and collective well-being.
- Written records or audio recordings of consent will be maintained and updated as appropriate when new members join or if additional consents are sought.



The UK researchers will hold appropriate Disclosure and Barring certificates.

6.4 Principles of privacy

- Research notes and visual records along with interview material and transcripts will be kept in secure conditions.
- The project will be registered with the Open University Faculty Data Protection Officer.
- Images and other audio visual footage are property of the individual. Each individual will be informed
 in person of the possible use of photography and other data collection methods as part of ARCHES
 research sessions. If subsequently we wish to use any material, we will need to seek further specific
 permission.

The aspects described in section 5 will also be taken into account in this field.

6.5 Deliverables

- Full ethical clearance for the collection of personal data from the appropriate committee at the
 participating universities was acquired prior to the end of the third month of the project, involving
 oversight of all consent materials. This was reflected in deliverable D8.2 "POPD Requirement No.
 3".
- A range of accessible consent material in English (see Annex A) was developed by the end of the second month of the project in deliverable D8.3 "POPD H Requirement No. 4". Versions in Spanish and German will be available two months prior to the start of the research groups in Spain and Austria.
- All participants are being provided with an accessible letter (see Annex A) and access to an online video one version for British Sign Language (BSL) speakers [22] and other version for the general public explaining the research process and offering and opt-out option [23].
- Detailed information on the informed consent procedures implemented were provided in deliverable D8.1 "H – Requirement No. 2".



7 Conclusions

In this deliverable the initial strategy to deal with data collection, generation, management and exploitation has been described. More information will continuously be included by the consortium partners until the final version is released in September 2019. This will allow us to improve and fine tune the preliminary actions (including the identification of datasets) as well as implement new measures if necessary.

The section dealing with the data summary presented the purpose, types and formats, re-use, origin, expected size and utility of the data collected and generated in ARCHES. Different categories were defined based on the diverse characteristics and needs of the project, involving the consortium partners as well as external sources.

The approach to make data findable, accessible, interoperable and re-usable was drafted in the corresponding section, where the use of a common data scheme to facilitate searches among different collections and the implementation of APIs to favour interoperability among different systems were commented. In addition, it was clearly stated that the re-use of existing digital CH assets fell within the objectives of the project.

No extra resources will be allocated for the actions described in the DMP since they fall within the common activities of the partners in charge of data management. Similarly, the security measures adopted for the ARCHES project will be aligned with their own policies in this field – in particular, OU has submitted its plan to The Open University Data Protection Officer. However, this does not prevent from considering complementary measures if sensitive data is handled.

Ethical aspects will be monitored in the framework of WP8 "Ethics requirements". The information sheets and informed consents provided in Annex A will be translated into Spanish and German before the exploration groups in these countries are built and following the same principles detailed in section 6 and applied for the first exploration group in the UK. The proposed protocol has already received the confirmation from The Open University Ethics Committee.



References

- [1] European Commission, "H2020 programme. Guidelines on FAIR Data Management in Horizon 2020", version 3.0, 26th July 2016. Available online at (28th March 2017): http://ec.europa.eu/research/participants/data/ref/h2020/grants-manual/hi/oa-pilot/h2020-hi-oa-data-mgt-en.pdf
- [2] ARCHES project website. Available online at (28th March 2017): http://www.arches-project.eu/
- [3] Europeana website. Available online at (28th March 2017): http://www.europeana.eu
- [4] DBpedia website. Available online at (28th March 2017): http://es.dbpedia.org/
- [5] Rijksmuseum API. Available online at (28th March 2017): https://www.rijksmuseum.nl/en/api
- [6] European Commission, "Innovation ecosystems of digital cultural assets". Available online at (28th March 2017): https://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/reflective-6-2015.html
- [7] The Finnish National Gallery API. Available online at (28th March 2017): http://kokoelmat.fng.fi/api/v2support/docs/#/overview
- [8] Google Arts and Culture website. Available online at (28th March 2017): https://www.google.com/culturalinstitute/beta/
- [9] The Visual Art Encyclopaedia website. Available online at (28th March 2017): https://www.wikiart.org/
- [10]Creative Commons website. Available online at (28th March 2017): https://creativecommons.org
- [11]Open Research Online website. Available online at (28th March 2017): http://oro.open.ac.uk/
- [12]The Wallace Collection website. Available online at (28th March 2017): http://www.wallacecollection.org/
- [13]YouTube website, "The Wallace Collection channel". Available online at (28th March 2017): https://www.youtube.com/user/TheWallacecollection
- [14]CIDOC-CRM website. Available online at (28th March 2017): http://www.cidoc-crm.org/
- [15]M. Doerr, M. Theodoridou (2011), "CRMdig: A generic digital provenance model for scientific observation", 3rd USENIX Workshop on the Theory and Practice of Provenance, Heraklion (Greece).
- [16]Cordis website, "3D-COFORM project". Available online at (28th March 2017): http://cordis.europa.eu/
 .project/rcn/89256 en.html
- [17]E. Tsalapati, N. Simou, N. Drosopoulos, R. Stein (2012), "Evolving LIDO based aggregations into Linked Data", CIDOC 2012, Helsinki (Sweden).
- [18]M. Olensky (2010), "Semantic interoperability in Europeana. An examination of CIDOC CRM in digital cultural heritage documentation", *IEEE Technical Committee on Digital Libraries*, 6(2). ISSN 1937-7266.
- [19]The National Archives, "Data protection act 1998". Available online at (28th March 2017): http://www.legislation.gov.uk/ukpga/1998/29/contents
- [20]The National Archives, "Freedom of information act 2000". Available online at (28th March 2017): http://www.legislation.gov.uk/ukpga/1998/29/contents
- [21]British Educational Research Association (2011), "Ethical guidelines for educational research". Available online at (28th March 2017): https://www.bera.ac.uk/wp-content/uploads/2014/02/BERA-Ethical-Guidelines-2011.pdf



[22]YouTube website, "Welcome to ARCHES" (with BSL). Available online at (28th March 2017): https://www.youtube.com/watch?v=tH73dR6TnPg

[23]YouTube website, "Welcome to ARCHES". Available online at (28th March 2017): https://www.youtube.com/watch?v=tH73dR6TnPg



Annex A Ethics requirements

A.1 Information sheets

A.1.1 Information sheet for people with learning disabilities





Faculty of Wellbeing, Education and Language
Studies
Department of Education
The Open University
Walton Hall
Milton Keynes
United Kingdom

MK7 6AA
Tel +44 (0) 1908 652901

ARCHES Information Letter

We want to work with people with learning difficulties and other differences. We will also work with technology companies. Together we will make museums easier to visit.









We need to give museums and technology companies good advice.





We will write reports and make films for museums and technology companies





If you want more information about joining the ARCHES project contact Helena at:

Helena.Garcia-Carrizosa@open.ac.uk or on 07881 512507.





Figure 2: ARCHES Information Letter for people with learning disabilities – general description.







Faculty of Wellbeing, Education and Language
Studies
Department of Education
The Open University
Walton Hall
Milton Keynes
United Kingdom
MK7 6AA
Tel +44 (0) 1908 652901

How is the research done?



· We will take photographs and make films.





We will make notes



· We will have questionnaires





We will talk about the visit





Figure 3: ARCHES Information Letter for people with learning disabilities – the research process (1/2).







Faculty of Wellbeing, Education and Language
Studies
Department of Education
The Open University
Walton Hall
Milton Keynes
United Kingdom
• MK7 6AA
Tel +44 (0) 1908 652901

A typical visit to the Victoria and Albert Museum and the Wallace Collection.





Arrive at the museum at 11 o'clock.





· Spend an hour going around the museum



· Have lunch and talk about the visit



Share our thoughts



If you want more information about joining the ARCHES project contact Helena at:

Helena.Garcia-Carrizosa@open.ac.uk





Figure 4: ARCHES Information Letter for people with learning disabilities – the research process (2/2).



A.1.2

Information sheet for people with sensory disabilities





Faculty of Wellbeing, Education and
Language Studies
Department of Education
The Open University
Walton Hall
Milton Keynes
United Kingdom
MK7 6AA
Tel +44 (0) 1908 652901

ARCHES Information Sheet

This information is about taking part in the ARCHES research project.

If you want more information or to join the ARCHES project contact Helena at Helena.Garcia-Carrizosa@open.ac.uk

ARCHES (Accessible Resources for Cultural Heritage EcoSystems) is a Horizon 2020 funded project. ARCHES involves partners in Heritage and Technology across Europe. The 3-year project ends in October 2019. It will develop online resources, software applications and multisensory technologies to enable access to Cultural Heritage Sites within and beyond the project. The project has three phases. Phase 1 will involve developing new technologies. Phase 2 will involve testing and redeveloping these technologies. Phase 3 will involve checking the new technologies are ready for other people to use.

The overall project is led by TreeLogic, whilst the research for ARCHES is being led by the Open University with Bath University. We are setting up participatory research groups with people with differences and difficulties associated with perception, memory, cognition and communication.

Participants will have a role to play in collecting and understanding the things we learn, and in deciding what we learn about. We will need to ensure that we have accessible data and accessible data analysis. Our research methods will be developed in collaboration with the research group so they can best identify, capture and record their experiences and views. We will consider using a wide variety of ways of recording information. We could for example use photographs, audio recordings, video recordings, interviews, questionnaires, note taking, drawings and different types of software. You will be able to withdraw consent if you do not wish to have photographs or other recordings used in public forums.

The research groups will be in London, Madrid, Oviedo and Vienna. At each site we will set up a group with around 20 members plus their supporters. The intention is for the group to meet every week over the lifetime of the project to undertake some kind of activity associated with the cultural heritage site or the technology being developed within ARCHES. The meetings are likely to be on an agreed day of the week and to run from mid-morning to mid-afternoon. They will consist, for example, of visiting parts of the museum, to examine specific exhibits or rooms using different technologies or of undertaking separate workshop activities to explore a specific issue of access. The group may meet all together or work in smaller groups. Some people may want to come along every week, some people may want to attend less often. We will need to be flexible to support each other in the best way. There will be many different opportunities for participants to take a leading role. We will provide participants with an agreed sum for Travel. There will also be a light meal at the meetings. As much as possible meetings will be arranged to suit the needs of the participants.

Figure 5: ARCHES Information Sheet for people with sensory disabilities (1/2).





All data will be kept in compliance with Data Protection Act (1998) and Freedom of Information Act (2000). We will agree with participants the level of anonymity they wish to maintain, and if they do not express a view we will ensure their names and any other identifying features are removed. We will consider supporters and other accompanying adults and friends to be participants within the research. All participants will have the project explained to them. Consent will be sought in a manner appropriate to a participant's preferred means of communication. We understand that giving consent is an ongoing process. Since this research will be a an evolving process, all participants will need to recognise the on-going, evolving, unfolding nature of the informed consent process throughout the project. All members of the ARCHES Participatory Research Groups can withdraw from the research at any time.

I understand that if I have any concerns or difficulties I can contact the leaders of the research at the Open University or TreeLogic. Either

Jonty Rix
Jonathan.rix@open.ac.uk
Tel: 01908 655903
The Open University
Stuart Hall Building
Walton Hall
Milton Keynes
MK7 6AA
UK

Javier Gutierrez Meana
javier.gutierrez@treelogic.com
Tel: (0034) 985 966 136
TreeLogic
Parque Tecnológico de Asturias ·
Parcela 30
E33428 Llanera ·
Asturias
Spain

If you want more information or to join the ARCHES project contact Helena at Helena.Garcia-Carrizosa@open.ac.uk.

2

Figure 6: ARCHES Information Sheet for people with sensory disabilities (2/2).





Informed consents A.2

A.2.1 Informed consent for people with learning disabilities





Faculty of Wellbeing, Education and **Language Studies** Department of Education The Open University Walton Hall Milton Keynes **United Kingdom** MK7 6AA Tel +44 (0) 1908 652901

ARCHES CONSENT FORM

Name_	
Please circle your choice:	
YES NO OR YES NO	
1. I want to take part in the project	ARCHES
YES NO 2. The project was explained to me	
YES NO	
3. I was informed that I can say I don't want to participate at any point. YES NO	No
4. All my information will be kept safe. YES NO	files

Figure 7: ARCHES Consent Form for people with learning disabilities (1/2).



5. I can be photographed and/or filmed







ES NO

6. My ARCHES research can be used for education or research and publication





YES



7. My picture can be used in Newspapers or on TV or on the INTERNET







(You can say NO at any time.)

I give permission to the Open University and Bath University to share and use my words and thoughts. This is my copyright. I agree to give my copyright to the Open University and Bath University. They will not use my copyright for profit.

 \mathbb{C}

Signed:

D	u	E.
1.400		

Signature of Supporter
(if appropriate)

Name

Open University Research Leader Jonty Rix Jonathan.rix@open.ac.uk Tel: 01908 655903



ARCHES Project Leader in Spain

Javier Gutiérrez Meana javier.gutierrez@treelogic.com Tel: 0034 985 966 136



.



Figure 8: ARCHES Consent Form for people with learning disabilities (2/2).



Informed consent for people with sensory disabilities A.2.2





Faculty of Wellbeing, Education and **Language Studies** Department of Education The Open University Walton Hall Milton Keynes United Kingdom MK7 6AA Tel +44 (0) 1908 652901

ARCHES Consent Form

agree to take part in this research project, 'ARCHES'. I have had the purposes of the research project explained to me. I have been informed that I may refuse to participate at any point by simply saying so. I have been assured that my confidentiality will be protected as specified in the project information sheet and agreed within the participatory research meetings. I agree that the information from this participatory research can be used for educational or research purposes, including publication.					
I understand that if I have any concer Either:	rns or difficulties I can contact the leaders of the research				
Open University Jonty Rix Jonathan.rix@open.ac.uk Tel: 01908 655903 Level 3 Stuart Hall Building Walton Hall Milton Keynes MK7 6AA	Treelogic ARCHES Project Leader Javier Gutierrez Maena javier.gutierrez@treelogic.com Tel: (0034) 985 966 136 Parque Tecnológico de Asturias · Parcela 30 E33428 Llanera · ASTURIAS [Spain]				
representatives of the ARCHES parti	tion to the Open University and Bath University as cipatory research groups for non-profitable use in the design of the accessible technology.				
Signed:	Date:				

Figure 9: ARCHES Consent Form for people with sensory disabilities.

The project leading to this application has received funding from the European Union's Horizon 2020 research and

innovation programme under grant agreement No 693229