Traction
Opera co-creation
for a social
transformation

Data Management Plan

DISSEMINATION LEVEL	PUBLIC (PU)
WORK PACKAGE	WP1 MANAGEMENT
DELIVERABLE NUMBER	D1.4
VERSION	V1.0
SUBMISSION DATE	30/06/2020
DUE DATE	30/06/2020

Authors

Authors in alphabetical order		
Name	**Organisation**	**Email**
Dorleta García	VICOM	dgarcia@vicomtech.org
Esther Novo	VICOM	enovo@vicomtech.org

Control sheet

Version history			
Version	**Date**	**Modified by**	**Summary of changes**
V0.1	22/05/2020	D. García (VICOM)	Initial draft and request for contributions from partners.
V0.2	22/06/2020	D. García (VICOM), ALL	Integration of input from all partners. Draft version ready for review.
V0.3	29/06/2020	D. García (VICOM)	Updated with review comments.
V1.0	30/06/2020	E. Novo (VICOM)	Final version ready for submission.

Peer review		
	Reviewer name	**Date**
Reviewer 1	Pablo César-CWI	26/06/2020
Reviewer 2	Camille Donegan-VRI	26/06/2020

Table of content

Abbreviations

Abbreviation	Definition
AWS	Amazon Web Services
BY-NC-ND 4.0	Attribution - Non-Commercial - No Derivatives 4.0 International
CC	Creative Commons
DDD	Digital Repository of Documents
DMP	Data Management Plan
DoA	Description of Action
ES	Spain
FAIR	Findable, Accessible, Interoperable and Reusable
GB	Gigabyte
GDPR	General Data Protection Regulation
IPR	Intellectual Property Rights
ORDP	Open Research Data Pilot
PPT	PowerPoint
PT	Portugal
VR	Virtual Reality
WP	Work Package

Executive summary

The aim of TRACTION is to conduct experimental projects with people in risk of social exclusion, who will test and share new ideas. Bridging the social and cultural divides involved, it will challenge many existing beliefs, structures and habits. The exceptional resources of the TRACTION partnership will help us meet that challenge through mutual support. The immediate outcomes will be new routes for social and economic integration for the people involved, better relationships between opera producers and society, and cutting-edge technological development. But the long-term prize is the definition of new processes that renew the art's potential to build cohesive societies and imagine a revitalised, common culture in which everyone can feel that they belong

As part of the TRACTION project, WP1 has the objective of organising the management of the project. The project contains contributions from a number of partners and individual activities that will require close coordination to ensure that project milestones are satisfactorily achieved. The activities related to the management of the project will ensure the timely execution of the work plan, the proper communication between participants, the data management plan for the project, the creation of reporting and quality control structures and procedures, the representation and communication with external entities, primarily the European Commission, and all financial-related activities concerning funds and budget allocation. In particular, Task 1.4 is devoted to data, ethics and privacy management, and it will determine how research data will be handled during the research project and after the project is completed.

Therefore, the objective of this document is to specify what data will be open: detailing what data the project will generate, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved. This deliverable will be updated during the project; thus, D1.4 represents the initial plan for Data Management in the project.

At this initial stage, reflections on which data will be generated have been made. Thus, we have been able not only to foresee which data and from which sources we will create, but also to bear in mind what resources we will need, which data will be made available and how. Also, next steps towards a successful Data Management have been taken into account.

The purpose of this deliverable is to give an insight into data management and the life-cycle of the data that will be used. It has been structured according to the European Commission's guidelines[1] as follows:

Chapter 2 - Data summary: this section will analyse what data will be generated.

Chapter 3 - FAIR data: the FAIR principles will be applied to the data mentioned in the previous chapter: how to make data Findable, Accessible, Interoperable and Reusable.

Chapter 4 - Allocation of resources: the amount of resources that are expected to invest in order to make the data FAIR.

Chapter 5 - Data security: means for data recovery and storage are foreseen.

Chapter 6 - Ethical aspects: a relation between data and the ethical guidelines is made.

Chapter 7 – Conclusions: conclusions of the deliverable and next steps are explained.

---

[1]   https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm#A1-template

1 Introduction

1.1 TRACTION concept and approach

Opera uses all the visual and performing arts to create extraordinary worlds of passion and sensibility. It is rightly recognised as a great achievement of European culture. And yet a form that once inspired social and artistic revolutions is often seen as the staid preserve of the elite. With rising inequality and social exclusion, many see opera—if they think of it at all—as symbolic of what is wrong in Europe today. TRACTION aims to change that using opera as a path for social and cultural inclusion, making it once again a force for radical transformation.

We do not want to make opera palatable to those who do not attend. We want to define new forms of artistic creation through which the most marginalised groups (migrants, the rural poor, young offenders and others) can work with artists to tell the stories that matter now. By combining best practice in participatory art with digital technology's innovations of language, form and process, we will define new approaches to co-creation and innovate in three fields: a) Opera creation and production; b) Immersive and interactive digital media; and c) Social integration and community development.

Experimental trials in inner-city Barcelona (ES), a youth prison in Leiria (PT) and diverse communities in Ireland will test and share new ideas. Bridging the social and cultural divides involved will challenge many existing beliefs, structures and habits. The exceptional resources of the TRACTION partnership will help us meet that challenge through mutual support. The immediate outcomes will be new routes for social and economic integration for the people involved, better relationships between opera producers and society, and cutting-edge technological development. But the long-term prize is the definition of new processes that renew the art's potential to build cohesive societies and imagine a revitalised, common culture in which everyone can feel that they belong

1.2 Purpose of the deliverable

Task 1.4 involves the overall management of Data, Ethics and Privacy, in which collected, processed or generated data will be describe, as well as the methodologies and protocols that will be followed and whether data will be made public or not. The result of the task will be D1.4 Data Management Plan, which will provide feedback to WP2. The aim of the Data Management Plan is to be a living document which will be updated during the project, when a significant change occurs.

1.3 Intended audience

The dissemination level of D1.4 is public and will therefore be made available. This document is intended to be an internal guideline for the appropriate data management of the TRACTION project.

2 Data Summary

2.1 Purpose, relation to the objectives and utility

The overall data will be divided into two throughout the deliverable: we will handle creative data, and also documental data.

The purpose of the data collection is also divided into two: on the one hand, creative data, which involves data gathered by technology providers and also data collected by the trial partners during focus groups and discussions, aims to understand the requirements to develop the toolsets that will be used for future opera productions, and which will enable a broader audience to access, co-create and experience new forms of opera. VR artists will also benefit from it in order to create other VR operas in the near future, given that that assets will be stored in the media vault and will therefore be reusable. Other than being used for the toolset, this data will also give us the opportunity to assess the experience of watching opera as well as the impact of different quality levels in audio, video and multisensorial media. Simultaneously, users will obtain an improved toolset and content that works in different network conditions and devices.

On the other hand, documental data will be collected by the partners who are involved in the trials. The purpose of that data is to monitor the progress on cross fertilisation activities and on social impact evaluation and co-creation. Concurrently, they will find it useful to compile research reports and deliverables, as well as assessing the project and its overall impact.

As for the relation with the objectives of the TRACTION project, the data gathering is related to all objectives 1-5:

- O1: Promote, through their empowerment, a transformation of communities at risk of exclusion.

- O2: Establish an effective collaborative and participatory production workflow for the co-creation and codesign of art representations.

- O3: Lay down a community-centric methodology to conduct an efficient and measurable community dialogue that will last in time and be replicable.

- O4: Explore novel audio-visual formats based in European cultural heritage, such as opera.

- O5: Maximise the social and market impact of the TRACTION results.

O1, O3 and O5 are primarily related to the data gathering made by the trials, while O2 and O4 have a direct relation with the technical partners. Relevant data will be collected to maximise the social and market impact of the TRACTION results. The consortium will ensure that the project has a determining impact on the European and global opera houses and theatres, democratising the access to opera and reaching new audiences, as well as exploring the impact in other art representations, in the cultural and creative industries, or across other sectors. The project will also pursue a sustainable social impact, where the promoted transformation will last in time and be expandable to other ecosystems.

2.2 Type of data

Throughout its lifetime, the TRACTION project will generate and collect a variety of data, in different types and formats. Mainly two types of data have been identified:

Description	Expected size	Origin
Creative: audio and video recordings, photographs, drawings, Tilt Brush export files, ambisonic files, 360 video files.	Estimate up to 5 terabytes in total	Data will be provided by the partners during Focus Groups, trials and workshop activities.
Documental: user feedback collected through online questionnaires, consent forms, demographic questionnaires.	Around 500 questionnaires, 100 GB	Questionnaire answers from users and participants (staff, professionals, technicians)

All these types of data in different formats and for different uses will be analysed by the partners case by case to study the potential conflicts coming from commercialization, legal or ethical uses.

However, and as a part of the Open Research Data Pilot (ORDP) from the European Commission, apart from creating this DMP we will provide access to research data to the maximum extent. As described in the "*Guidelines to the Rules on Open Access to Scientific publication and Open Access to Research Data in Horizon 2020*", Open access to research data refers to the right to access and reuse digital research data under the terms and conditions set out in the Grant Agreement. The aim of TRACTION is to improve and maximise access to and re-use of research data generated by the project, taking into account: the need to balance openness and protection of scientific information, commercialisation and IPR, privacy concerns, security, and data management and preservation questions. Some of these aspects will be addressed initially in this deliverable and further developed as part of WP1 and WP6 activities.

3 FAIR data

As mentioned in section 2, two types of data have been distinguished in order to analyse their Findability, Accessibility, Identifiability and Reusability: creative data and documental data. While some principles will be applied to all the data gathered, in some others a distinction will be made based on the type.

3.1 Making data Findable

Aside from justified exceptions, all the research data produced and/or used in the project (both creative and documental) will be discoverable with metadata (name, media presentation description, subtitles, tags, timeline content such as chat messages, multisensorial outputs), identifiable and locatable by means of standard identification mechanisms such as Digital Object Identifiers[2]. Search keywords will be provided to optimize possibilities for re-use and the following naming convention will be followed:

TRACTION_[DATA-TYPE]_[DESCRIPTIVE-NAME]_[VERSION-NUMBER]

To facilitate finding but also handling research data produced by the TRACTION project, clear version numbers and release notes (when applicable) will be provided.

To the maximum extent possible, metadata standards will be followed. When needed, extensions to the existing standards will be used in the project and proposed as a contribution to the bodies responsible for maintaining those standards.

In a nutshell, in order to be able to retrieve and find data, the requirements have been listed below:

- (meta)data will be assigned a globally unique and persistent identifier
- data will be described with rich metadata
- metadata will clearly and explicitly include the identifier of the data it describes
- (meta)data will be registered or indexed in a searchable resource

3.2 Making data openly Accessible

TRACTION has a policy of Open Source for all its assets. To this aim, a collection at UAB DDD has been created. The DDD, the UAB Digital Repository of Documents, is the tool that collects, manages, disseminates and preserves the scientific, educational and institutional production of the University. It also gathers digital documents that are part of or complement UAB library collections, where publications will be hosted. DDD only collects email addresses of users that do not belong to the university community; and under no circumstances can the Library Service staff view personal passwords. Users that have not accessed their personal accounts in the previous 5 years will be discharged unless they have pending baskets. The DDD repository shows an organized, open access and interoperable collection. In accordance with the General Data Protection Regulation

---

[2] http://www.doi.org/

(hereinafter, GDPR), the DDD will not disclose the information of its users regarding deliveries or individual visits to the repository, nor the Email address required to register in case the user is not a member of the university community.

As for the availability of the creative data, some of it will be made accessible on the web (TRACTION webpage, UAB DDD, DCU DORAS); some other only through the toolsets (immersive player, media vault and performance engine), or Matlab files. In compliance with the GDPR, some of the creative data such as photographs, videos and sound recordings will be treated as personal data and will therefore not be openly accessible unless explicit permission is granted.

Code produced within the TRACTION project will be deposited on GitHub and on secure servers.

With regards to documental data, this is, raw data, it is considered internal working material. Interviews, questionnaires, assessments, etc. will be considered confidential for the internal use in reporting, under the Traction ethics protocols[3]. GDPR needs to be applied so as to protect data and privacy, therefore, before making any content available, all data will be anonymised (see WP6 deliverables for further details).

The results of the research study will be made available to the trial participants and to the public in an aggregated and anonymised form, and open access to research data will be granted. Moreover, open access to publications, 'gold' or 'green' open access[4] will be granted.

3.3   Making data Interoperable

Interoperability of data is one of the goals of the TRACTION project, as it fosters collaboration and efficiency. Existing data-exchange standards will be used whenever possible.

In case it is found unavoidable to use uncommon or to generate project specific ontologies or vocabularies, the first approach will be to extend the existing ones. If this option is not feasible, then mappings to more commonly used ontologies will be pursued.

To be able to exchange data, the following requirements will be met:

- (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- (meta)data use vocabularies that follow FAIR principles.
- (meta)data include qualified references to other (meta)data.

The data gathered during the focus groups and interviews will be transcribed. They will then be analysed using open coding methodologies. It is not planned to make the data available for third parties (because of potential privacy issues).

---

[3] https://ddd.uab.cat/pub/presentacions/2020/220756/TRACTION_Ethics_Guidelines.pdf
[4]   https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm

3.4   Making data Re-usable

In order to make data re-usable, a choice for a licence was made. The consortium decided that Creative Commons licences were the best choice, more specifically the CC BY-NC-ND 4.0, as we can see below.

Figure 1 CC BY-NC-ND 4.0 licence terms[5]

In order to apply this licence, a notification must be made in all the assets, either in writing ("*© 2020. This work is licensed under a* CC BY-NC-ND 4.0 license*")* or by embedding the licence terms of the licence chosen. CC licences cannot be revoked; therefore, this will assure the re-usability of the assets indefinitely.

---

[5] https://creativecommons.org/licenses/by-nc-nd/4.0/

Aggregated data and research papers reporting findings based on the collected data will be submitted to DCU DORAS and UAB DDD, after anonymising personal data. The submission to the DDD already involves the CC licence being applied, and thus all assets must have the aforementioned notice. Once approved for submission in conferences and journals, and approved on the DCU DORAS and UAB DDD databases, papers are instantly available.

Creative data affected by personal data classification, such as photographs, videos and sound recordings, will not made publicly available unless explicit permission is collected.

4 Allocation of resources

All the data will be safely stored in each partner's servers or in private AWS instances secured by password and with encrypted. All data flowing across the AWS global network that interconnects AWS datacentres and regions is automatically encrypted at the physical layer before it leaves their secured facilities[6]. Upon project completion, all the data will be moved to Vicomtech servers with open access for researchers upon request. The data will be made available for 5 years after the end of the project. Consent forms will be kept in a safe place by the principal investigators and will be destroyed five years after the investigation has finished.

Each of the partners has appointed a data controller and processor, in compliance with the 2016/679 General Data Protection Regulation.

Regarding the long-term preservation of data, partners have allocated resources related to the purchase of the needed storage. AWS costs are currently being managed by VICOM. Secured servers for keeping text data, forms, questionnaires, metrics from tools are already in place. If necessary, extra storage can be purchased.

On the other hand, data gathered by trials will be stored on hard drives and box files if it is not digital; and therefore, costs of preservation will be minimal. All data will be returned to LICEU, SAMP and INO where possible after use, and they will store it for five years after the project has finished.

With regards to Open Access to scientific publications, all TRACTION partners that intend to publish scientific papers have a minimum budget of 3,000 EUR for conference and other dissemination costs (on top of travel and expenses). These budget allocations should allow covering (at least partially) of the costs of complying with the open data requirements of the project.

---

[6] https://aws.amazon.com/security/?nc1=h_ls

5   Data security

All data stored in AWS servers is automatically backed up. At Vicomtech backups of all the data is performed regularly and, for increased security, data is also stored on a secure private cloud. The data stored in DCU servers will be located in the Performance Engineering Laboratory under strict security protection.

As for the data kept in office formats (.doc, .xls, .ppt, etc.), they are secured in laptops, protected by a password, and backed up onto clouds and portable hard-drives, stored in locked file cabinets. Hard copies of sensitive data will be held in a secure, locked room at each partner's premises. Digital copies will be stored on a secure, private, internal online repository.

6   Ethical aspects

The TRACTION project is aligned with the ethics requirements proposed by the two Universities in the consortium: UAB and DCU, and their internal Research Ethics Committees' requirements regarding sensitive data, storage and personal data security schedule.

An internal webinar on Ethic Requirements for TRACTION members was organised. The PPT was uploaded to the open repository at UAB DDD and can be found here: https://ddd.uab.cat/record/220755. All testing should follow Ethical procedures and Sex and Genre aspects. This includes signing in an information sheet and consent form in an accessible format. Informed consent is understood by the partners as an on-going process rather than a level of legal protection. It is not intended to be a one-time act of having a participant sign a form. Informed consent is designed to inform research subjects about the purpose, risks, potential benefits and alternatives to the research that allows people to make a decision about whether or not to participate based on their own goals and values. This exchange of such information should occur at enrolment and throughout the study. The consent form was created to be accessible to all and signed by participants to give release to images and videos portraying them. It has been submitted to the UAB DDD open access repository and can be found here: https://ddd.uab.cat/record/221188

More details about ethics and data protection can be found in related WP6 deliverables.

7   Conclusions

Throughout this document we have identified the data that will be gathered, how to make it Findable, Accessible, Interoperable and Re-usable, the allocation of resources needed to maintain the data FAIR, and also contemplated the ethical issues that could arise from the data.

This is an initial version of the Data Management Plan, completed with rough information that will be more specifically completed as the project progresses. It will be updated if a significant change occurs regarding data, and also at the end of the project, at a finer level of specificity.

As a prevision to future steps, a register with all the assets will be created, specifying the type of data that it will contain, whether it will be accessible, when and where. In the next versions of the Data Management Plan, that register will be added.