ICT-687655

Delivering Single and Multi-Screen Content Services for Immersive, Customised and Shared Experiences in Homes and Social Spaces

D1.4 Data Management Plan

Due date of deliverable: September 2018

Actual submission date: December 2018

Start date of project: 1 December 2015   Duration: 36 months

Lead contractor for this deliverable: BBC

Version 1.4, 30 November 2018

Confidentiality status: Public

Consortium partners

#	Party	Description
1	BBC	Coordinator
2	BT (British Telecommunications Public Limited Company)	British multinational telecommunications holding company headquartered in London, United Kingdom.
3	CWI (Stichting Centrum Voor Wiskunde En Informatica)	The national research institute for mathematics and computer science of the Dutch National Science Foundation
4	IRT (Insitut fur Rundfunktechnik)	German non-profit company dedicated to technical research and development of broadcast and multimedia technology for the benefit of the public.
5	Illuminations	A British innovative media production company that specialises in performance and the arts
6	Cisco	Designs, manufactures and sells Internet protocol (IP) and other products related to the communications and information technology industry worldwide
7	ChyronHego	Swedish company that offers media products and services for live television, news, sports, corporate and government video production

Project website

https://2immerse.eu/

Abstract

This Data Management Plan (DMP) describes the data management processes and all of the data processed as part of the 2-IMMERSE H2020 project. It describes the data management life cycle and is part of the 2-IMMERSE Consortium's effort to make research data findable, accessible, interoperable and re-usable (FAIR). Good research data management leads to knowledge discovery, innovation and extraction of new valuable information.

This document provides a final check on whether initial data management plans (as described in this integrated document) have been adhered to, and measures that are implemented to identify processing of personal data and to ensure that each consortium partner is aware of its obligation to comply with new data protection legislation such as GDPR. This plan also includes details of some measures adopted by the consortium as a whole to minimise personal data and to mitigate data protection related risks.

Target audience

2-IMMERSE team members, the Project Officer from the European Commission, 2-IMMERSE reviewers and auditors. This is document is intended to be published on the 2-IMMERSE website[1], however parts that are commercially or otherwise sensitive may be removed from the public version.

---

[1] http://2immerse.eu

Definitions

Abbreviation	Definition
DMP	Data Management Plan
EC	European Commission
GDPR	The General Data Protection Regulation 2016/679/EU
IP	Intellectual Property
IPR	Intellectual Property Rights
OB	Outside Broadcast
DMApp	Distributed Media Application
GA	Google Analytics
AWS	Amazon Web Services
BSCW	Groupware for efficient teamwork and document management hosted by Eurescom. (https://www.bscw.de/en/)

Externally Contracted Services

Supplier	Service	Contracting partner
Acumen Field Ltd	Research management and recruitment	Supplier was contracted by the BBC
Eurescom	Project coordination support; Provision and management of an EU hosted instance of the BSCW team management software, including safeguards for secure storage and deletion of sensitive project data	Supplier was contracted by the BBC. The same supplier was contracted separately by BT.
TakeNote	Audio transcription – provision of anonymous transcripts from the Theatre-In-Schools service trial (see Section 4)	Supplier was contracted by the BBC
The Donmar Warehouse	Theatre production services - provided Shakespeare (Julius Caesar) IPR AV content for playout in the Theatre-In-Schools service trial and facilitated the service trial (see Section 4).	Supplier was contracted by Illuminations
Dorna	Dorna Sports, S.L. is the commercial rights holder for the motorcycling sport of MotoGP. They provided MotoGP race content for use in the MotoGP-At-Home service trial.	The supplier was contracted by BT.

Each consortium partner is acting as a data controller in respect of the processing of the personal data on its own behalf, or while personal data is in its possession or control. All partners acknowledge that in relation to the processing of the personal data, they must comply with their respective obligations under all relevant data protection legislation, including GDPR.

Summary

This Data Management Plan (DMP) describes the data management processes and all of the data processed within the 2-IMMERSE project, including technical results from the deployment of multi-screen content in the service trials. Datasets are described in accordance with the Horizon 2020 guidelines on the Data Management Plan (DMP)[2]. This facilitates exploitation of those datasets within and beyond the lifetime of the project and consideration of data subjects' freedoms and rights.

This document identifies processing of personal data, including sources of data and situations where data was shared (within the consortium or with third parties such as in the case of academic/research publications). The project's Dissemination Plan in work package 6 - *D6.3 Final Report on Dissemination, Standardisation and Exploitation Planning*, discusses the means of promoting the use of data by the research community and across the industry (e.g., academic publications, the open-source release of 2-IMMERSE software and its reference architecture).

This document is an updated version of an older data management plan submitted as work package 1 deliverable - *D1.3 Data Management Plan*. It also includes a section that maps all the data processed as part of the 2-IMMERSE project, which is based on Data Management Questionnaires completed by all consortium partners (see section 9). Documentation of is part of the coordinated effort of the consortium to ensure compliance with the new General Data Protection Regulation (GDPR) and implement measures to mitigate risks. This is in addition to Consortium partners' specific obligations to comply with any relevant data protection legislation and to coordinate their activities with the data protection officers in their organisation.

---

[2] http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

List of Authors

Mark Lomas (BBC), James Walker (CISCO), Maxine Glancy (BBC), Ian Kegel (BT), Martin Trimby (BT), Ido Azil (BBC)

Reviewers

All consortium partners

Table of Contents

1 Introduction

The purpose of the 2-IMMERSE Data Management Plan (DMP) is to document the datasets used and derived during the course of the project. This facilitates exploitation of those datasets within and beyond the lifetime of the project and includes consideration of data subjects' freedoms and rights. The structure of this document will follow the guidelines[3] for data management in projects carried out within the Horizon 2020 scheme as set out by the European Commission (EC).

The 2-IMMERSE project has captured several types of resulting data including: 1. User data in the form of questionnaires and audio transcripts, 2. Ethnographic data in the form of videos and questionnaires, 3. System logs and user analytics data captured as text files and tabulated into spreadsheets, and 4. Software. The type of data collected for each service trial is described in detail in the data management questionnaire in section 9, but the table below provides a high-level summary of the data collected for each service trial, together with a short description of each trial

Service Trial	Description	Num. participants	Data collected
Theatre-At-Home	**Watching live theatre in the home:** Supporting friends coming together in a virtual theatre box to enjoy a live filmed production of a play whilst gaining additional related content and communications capabilities on their personal screens.	23	Participant and parent/guardian contact details, video & audio recordings of interviews with producers and trial participants. Anonymous Online survey, questionnaire and phone interviews with requited trial participants. System logs.
MotoGP-At-Home	**Watching motor sports at home:** Exploring how coverage of MotoGP can be enhanced using synchronised video and data across multiple screens that adapt responsively to the display characteristics of the experience level of individual viewers.	100	Participant names and contact details. Demographic screening questionnaire answers. Consent forms. Video recordings of participants in lab-based evaluation. Names and contact details of facilitators from external research management company. Anonymous user experience questionnaires. System logs and data analytics.
Live Football-At-Home	**Watching live Football-At-Home:** Technical trial to explore multi-screen production tooling	Approximately 6 consortium members and	Raw video and photos of pre-broadcast preparation and live broadcast activity within a BT Sport

---

[3] http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

	for live football events, directing synchronised content across multiple screens to the home.	a dozen BT sport staff.	outside broadcast (OB) truck. Surveys completed by BT Sport staff. System logs.
Football Fanzone	**Watching football on multiple screens:** Enhancing the experience of watching key sporting events, by providing additional content to the viewer's mobile devices to increase engagement with, and enjoyment of, the event.	Shown to public at IBC	System logs.
Theatre-In-Schools	**Watching live Theatre-In-Schools:** Enabling students to augment the main filmed presentations of a play with access to related supporting content and experiences to help them deepen their understanding of the play.	51 students 4 teachers	Participant and parent/guardian contact details and consent forms. Video & audio recordings of interviews with producers and trial participants. Paper-based questionnaires. System logs.
Production Tools Trials	**Evaluation of multi-screen authoring tools for pre-production and live triggering**	13	Consent forms. Photographs of participants in experiments captured by CWI in the Netherlands. Audio and video recordings of sessions. Conversation transcripts. System logs.

2 Data Summary

During the project consortium partners carried out experiments and service trials exploring the provision and value of multi-screen content experiences of drama and sport in private and public venues. User feedback and behavioural data has been collected during these service trails.

All data has now been collected and Table 1 Data Management Questionnaire shows which of the storage and management strategies have been adopted, informed by the Horizon 2020 FAIR Data Management Plan (DMP) Template [2].

A mixture of anonymous data and personal data was captured during the project. All personal data was subsequently anonymised, summarised and destroyed as confirmed in the Data Management Questionnaires completed by all consortium partners (see section 9).

2.1 Data Management Methodology

All consortium partners completed a questionnaire to allow mapping and documenting all the data being processed as part of the project. This data mapping was in particular aimed at identifying all personal data processing activities, including types of personal data / data subjects and measures to minimise and control its processing.

2.2 Purpose of user data collection

As part of the project (multi-screen content) trials, user feedback and behavioural data has been collected to inform experience design and technology development: from setting up and configuring multi-screen environments, through signing up for services (on-boarding), enjoying the service and closing down. Such data helps to communicate guidelines that can be extrapolated beyond the service trial scenarios – see work package 3 and 4 deliverables for more information.

Data may also include information collected by micro-services and client devices that supported the multi-screen content during 2-IMMERSE sessions to troubleshoot and support live service trials and to generate insights into user behaviour.

All trial participants were properly informed how their data will be used and for what purpose, before they agreed to participate.

2.3 Reuse of anonymised user data

The (anonymised) user feedback and behavioural data generated by each pilot and its preceding experiments has been used to research the value of multi-device experiences. This includes evaluation of prior assumptions and designs. The insights drawn have been published along with (anonymised) summary data. Beyond the end of the project, such (anonymised) data may help other researchers and organisations draw further conclusions through re-analysis, comparison or combination with other data.

2.4 Origin of the Data and Software

Anonymised data has been derived from subjects in lab experiments, workshops and the service trials. This data is in the form of quantitative measurements made through: data analytics, Likert scales and logging transcripts of participant video footage over the course of the trials, feedback questionnaires, structured interviews, focus groups before and after the experiments (e.g. Shakespeare play, motor race). Data has been provided by members of the public voluntarily

agreeing to take part in each experiment and service trial, after signing participation agreements. See section 7.2 for details about the terms and context of the participation agreements.

Software code is another form of data that has been developed over the course of the project both to support the trials and the interactions within them and to support flexible and customisable multi-screen environments.

2.5 Expected Size and Utility

User data that was used for project's research activities included anonymised data extracted form of transcribed interviews, questionnaire responses, data analytics of screen usage and device interactions, video recordings of trial participants and observations. This anonymised behavioural data has been useful to understand the potential of multiscreen environments to enhance the experience of drama and sport. In addition, researchers or broadcast practitioners may be interested in considering the (anonymised) data to learn general lessons across other genres of entertainment.

3   FAIR Data

This section covers the steps 2-IMMERSE has taken to make the research data findable, accessible, interoperable and reusable (FAIR).

3.1   Making data findable, including provisions for metadata

Summaries of both the qualitative and quantitative anonymised data have been made available through academic/research publications. Software code specifications have been documented in written work package 2 deliverables D2.1 through to D2.6 and have been published as open source and made available via GitHub[4].

3.2   Making data openly accessible

Results of the 2-IMMERSE project have been published as written work package deliverables on the 2-IMMERSE website[5]. In addition, a core set of software repositories corresponding to the service platform have been published under the Apache Licence v2.0 on GitHub[4].

3.3   Increase data re-use (developed software)

In order to permit the widest possible re-use, core components of the software platform developed by the 2-IMMERSE consortium will be published on GitHub under Apache Licence v2.0 at the end of the project. The Apache Licence v2.0 licence is summarised as:

> *"A permissive license whose main conditions require preservation of copyright and license notices. Contributors provide an express grant of patent rights. Licensed works, modifications, and larger works may be distributed under different terms and without source code."* [6]

Apache Licence 2.0 is officially recognised by both the Open Source Initiative (OSI) and the Free Software Foundation (FSF) and is widely used. Thus, consortium partners consider it the most suitable for release of software developed as part of the project.

This makes it easier for other researchers to reproduce a working instance of the 2-IMMERSE platform with which to conduct further research and reproduce certain results.

Since the data collected was very specific to the 2-IMMERSE service trials, it is of very limited use in a wider context. Service trial data has already been processed and published via the work package deliverables and is available to the public through the 2-IMMERSE website[5].

For the avoidance of doubt, this DMP does not include any licence to use software or any promise to release project results as open source software in the future. Software developed as part of the 2-IMMERSE project may only be used in accordance with the specific licence attached thereto.

---

[4] https://github.com/2-IMMERSE

[5] https://2immerse.eu/

[6] https://choosealicense.com/licenses/

4   Data Security Summary

This is a summary of the data security measures taken by the 2-IMMERSE consortium. More details can be found in section 9 – Data Management Questionnaire.

Quantitative and qualitative data (for example, ratings and transcribed comments) obtained from users participating in the 2-IMMERSE were stored in secure encrypted repositories within the partners' organisations. Anonymisation of personal data took place soon after collection and before being processed for the purposes of research. Anonymous results derived from the original data have been archived in BSCW; a groupware for teamwork and document management hosted in Germany by our supplier [Eurescom], a European provider of project management services. BSCW is backed-up regularly and is hosted in the EU. Only 2-IMMERSE consortium members and BSCW administrators have access to the data. All personal data was deleted after anonymisation and was not retained for any longer than necessary. Personal data was stored separately from evaluation data and has been not been made available to any parties outside the consortium, other than the third parties contracted to recruit participants and facilitate the service trials, who provided the personal data in the first place.

5 Ethical aspects

This section summarises how 2-IMMERSE has addressed potential ethical issues that could have had an impact on data sharing, such as informed consent for data sharing, processing and long-term preservation.

Research participants gave informed consent for how their data would be used by the 2-IMMERSE consortium and its contracted recruiters and facilitators, by signing a participation agreement prior to involvement. The participation agreement was in the form of a consent form that was provided by the BBC and processed by the research management and recruitment supplier [Acumen Field Ltd] for the MotoGP service trials and the second production tools trial, and by the BBC and the production services supplier [Donmar] for the Theatre-At-Home and Theatre-In-Schools service trial. The participation agreement requested permission for the long-term storage and publication of relevant data.

Data such as the video recordings made of users participating in the 2-IMMERSE experiments and service trials were transcribed and anonymised so that participants could not be identified and the original recordings were deleted.

The Theatre-In-Schools service trial involved collected paper-based questionnaires and conducting informal interviews with students between the ages 14-15yrs who engaged in an activity where they interacted with a group device, viewed guidance and materials for a group activity and undertook activity in a timed session. The students' contributions were provided after parents or guardians signed the participation agreement. All service trials were carried out in UK schools and run by the teacher and the teaching assistant from that school. Only the teachers had access to the student's personal information, the consortium partners had no personal details of the students taking part (e.g. names) but did interact with them to conduct the informal interviews accompanied by the teaching assistant. Only anonymous usage data was collected by the consortium. Specifics of the data flow for the Theatre-In-Schools service trial are documented in Section 9 – Data Management Questionnaire.

The interviewers comprised organisers from the production services supplier [Donmar] contracted by Illuminations and project team members from Illuminations, BT and BBC. All interviewers were accompanied by a member of staff from the school at all times. Organisers from the production services supplier [Donmar] who helped design the lesson plan in conjunction with the teachers, were also Disclosure and Barring Service (DBS) checked for criminal record as they work with schools on a regular basis.

Guidance from teachers was provided about the length of time the students should be working with the technology to ensure health and safety regulations in terms of using screens were adhered to. Guidance was also provided on the age appropriateness of the Theatre-In-Schools content for the students' age (14-15yrs old) and for viewing in the classroom. Children had watched the Julius Caesar play as part of their school curriculum with the teachers prior to the trial.

6 Data Analytics

6.1 Overview

This is a summary of how system logs and data analytics were captured and managed by the 2-IMMERSE software during the service trials.

System logs were captured for all service trials. User analytics data was also captured during the MotoGP-At-Home service trial using a Google Analytics account belonging to BT. The system logs recorded pseudonymised data that was later anonymised (see comments below), such as the random device IDs and session IDs generated by the system and the anonymous user IDs randomly generated by BT and issued to the research management company [called Acumen Field Ltd[7]] contracted by the BBC. The external research management company dealt with recruitment of participants for the MotoGP-At-Home service trial and the second production tools trial.

Consortium partners kept records of which randomly generated user IDs were given to each of the external trial facilitators employed by the research management company. The facilitators internally documented, but never shared how the randomly generated user IDs were allocated to participants. The user IDs were given to facilitators in batches. Generally, different user IDs were given for each participant, although there was some re-use. Consequently, system log messages and user analytics data captured by the software during the MotoGP-At-Home service trial cannot be related back to participants by either the 2-IMMERSE consortium or the research management company alone. In addition, the external research management company contractually committed to follow a detailed data management policy and to ensure that personal data would only be retained for a 3-month period during the project, and then deleted.

It would have been possible to correlate the system logs and user analytics data to a participant if one was in possession of both the analytics data, how the randomly generated user IDs were assigned to participants and knowledge of the participants themselves. The data captured was therefore technically in a temporary pseudonymised state. The research management company subsequently deleted the mapping it kept, therefore fully anonymising the logging and analytics data collected. Had a correlation been made the information obtained would not have been valuable because the data captured by the analytics/logging did not include 'sensitive data', as defined by GDPR.

Facilitators interviewed participants to obtain answers to a series of survey questions about the MotoGP-At-Home service trial and recorded the answers in an spreadsheet. The survey was anonymous and each completed survey was assigned a randomly generated survey ID. The system log messages were correlated to the randomly generated user IDs in Kibana, a secure data analysis and visualisation tool hosted on the 2-IMMERSE software platform. This allowed the service trial to be monitored and helped to identify session IDs which corresponded to proper service trial sessions. However, since the user IDs were randomly generated by the consortium, the data being analysed didn't bear any relationship to the identity of those participating.

Only BT had access to its Google Analytics account which contains the analytics data. BT ensured that no personal data was stored on its account and was responsible for compliance with the Google Analytics terms of service[8]. The UserID and ClientID features of Google Analytics were not used; only events, custom dimensions and custom metrics were used (i.e., only anonymous data/identifiers were captured). The data will be kept by Google for approximately 18 months in accordance with

---

[7] https://www.acumenfieldwork.com/

their terms[8]. In order to analyse the data captured during the MotoGP trials, custom reports were created and custom segments within Google Analytics. Raw data was exported to Excel and Pivot tables and charts were generated. This processed data remains stored on an encrypted BT computer consistent with BT's InfoSec standards and also on [BSCW,] a groupware for teamwork and document management hosted in Germany by [Eurescom,] a European provider of project management services. All consortium partners have password protected access to data stored in BSCW. Summary results generated from this data have been published in the work package 4 deliverable - *D4.5 MotoGP Trial Evaluation Results*, and may also be included in future publications.

Devices (phones and tablets) with Google accounts were loaned to participants for the duration of the service trial by consortium partners (each partner purchased devices in order to be able to conduct service trials without relying on participants' equipment). There was no requirement for trial participants or the external research management company team to interact with Google features on the devices or enter their own credentials (and no evidence that they did has been seen after devices were returned). For all devices used during the trials, all accounts used were anonymous test accounts created by BT, BBC, CISCO and CWI.

6.2   Summary

- No personally identifiable information was captured by the logging system or via Google Analytics
- Pseudonymised device IDs, context IDs and consortium-assigned user IDs, that were later anonymised, were captured alongside the usage data from the service trials.
- UserID and ClientID features were not used in Google Analytics
- Only anonymous identifiers were captured in custom dimensions and metrics
- Sharing was disabled for all analytics data captured in Google Analytics
- Temporary Google accounts, created by the consortium partners, were used to sign into tablets and phones during the trials
- 2-IMMERSE consortium equipment was used in all service trials as opposed to the participant's own devices.
- Only BT has access to the Google Analytics password protected account and data will be kept for approximately 18 months.

---

[8] https://www.google.com/analytics/terms/us.html

7 GDPR

GDPR became applicable in all EU member states from 25th May 2018 onwards. The Consortium partners acknowledge the requirements of GDPR and have been coordinating their activities with their respective data protection officers.

The 2-IMMERSE consortium partners have completed a Data Management Questionnaire (see section 9) to allow mapping of all the data being processed as part of the project. This data mapping provides descriptions of the information held, how long it would be kept for, for what purpose and descriptions of technical security measures in place.

All personal data processed was stored in an encrypted format and subsequently anonymised or otherwise deleted soon after it was collected. The 2-IMMERSE consortium did not make data available in a form which identifies a data subject or otherwise processed in a manner that may adversely affect any individual.

Each consortium partner is acting as a data controller in respect of the processing of the personal data on its own behalf, or while personal data is in its possession or control. All partners acknowledge that in relation to the processing of the personal data, they must to comply with their respective obligations under all relevant data protection legislation, including GDPR.

7.1 Data protection by design and delivery of multi-screen services beyond the trials

The 2-IMMERSE software service platform developed as part of the project to power multi-device experiences doesn't capture, store or process data associated with any user identity. However, it does include a placeholder authentication system for the purpose of limiting access to 2-IMMERSE platform functionality and content. Parties wishing to utilise the software in future to collect personal data would have to further develop data protection measures and to integrate a legally compliant customer handling systems and data collection functionality.

7.2 Data Subjects Consent and Data Access

The 2-IMMERSE consortium deleted all personal data collected as part of the trials after initial processing and therefore only retained anonymous data. All personal data recorded during the 2-IMMERSE project was obtained with the explicit consent of participants. This consent also covered processing and publication of the anonymised data.

Consent was obtained using positive opt-in and it was clearly explained what consent was being given for by trial organisers and facilitators. Additional measures were used to ensure that consent was informed where it was obtained for the capture and anonymisation of audio, video and photographs. Consent forms are stored securely by BBC and the external research management and production services suppliers [Acumen Field Ltd and Donmar], and were deleted where potential participants have eventually not taken part in the trials.

Measures were implemented to ensure that a Subject Access Request (SAR) made by a former participant would result in confirmation that no personal data is being held. However, it would be impossible for the consortium to determine which datasets are associated with the participant in question.

7.3 Data protection risk assessment

No risks have been identified. As no data that can be used by third parties to identify trial participants was actually processed by the 2-Immerse platform, no significant impact to their rights and freedoms was likely in case of a security breach.

7.4 Authentication service

Although the consequence of a security breach would have been a nuisance at worst (as no data that can be used by third parties to identify trial participants was actually processed by the 2-IMMERSE platform), technical and organisational measures were implemented to demonstrate basic security measures and to protect the video content provided by the production services supplier [Donmar], [Dorna] and BT (content that was produced for viewers).

Experts at CISCO implemented measures to safeguard the smooth running of service trials (MotoGP-At-Home, Theatre-In-Schools, Football Fanzone and Football-At-Home), to secure the trial data collected and prevent harm to the platform. A proprietary 2-IMMERSE authentication service was deployed to Amazon Web Services (AWS) alongside the other 2-IMMERSE services to validate login credentials and guard against misuse of the platform. These measures were also aimed at illustrating how an organisation adopting this type of multi-device service architecture might go about integrating security features such as user accounts.

The authentication service worked in conjunction with an application gateway service (Tyk) to validate access tokens on behalf of service end-points and provide TLS/SSL termination. This allowed encrypting all client-server network transactions. A permissions model was implemented by the authentication service to limit user privilege and therefore further safeguard the platform and the running of the trials.

Login credentials were manually generated by consortium partners for development use and generated to give to the research management company for use in the service trials. The login credentials used by trial participants were randomly generated in bulk and didn't bear any correlation to the identity of specific individuals.

8 Progress in the third year

The following table is an update to the table originally published in the work package 1 deliverable - *D1.3 Data Management Plan*, summarising the table of promises and achievements in the third year of the 2-Immerse project

Section	Objective	Did we do it?
1.0	Data will be collected during experiments and pilots (field trails) exploring the provision and value of multi-screen experiences of drama and sport in private and public venues	Participants took part in four 2-Immerse trials (Theatre-At-Home, MotoGP-At-Home, Football-At-Home, Football Fanzone & Theatre-In-School). The number of participants is listed in section 1. This includes: - Data from an event exploring young people's responses to different presentations of football. This was a STEM career event hosted by BT at Adastral Park, UK. i.e. no names or contact details were collected. The school year was logged. The data was stored securely and only BT has access. - Video recordings and audio transcripts from interviews with producers and participants for production tool trials. - Paper-based questions from participants in Theatre-In-Schools and extracted conversations and quotes. - Participants have also completed online surveys about their experiences in the Theatre-At-Home service trial. - Google Analytics data and system logging data from MotoGP, and Football has been collected and analysed for the service trials and production tool trials. Logging data was also collected for Theatre-at-home. - Ethnographic studies were undertaken for the football service trials to understand the director's role in creating the broadcast mix of video, graphics and commentary narrative for the match. Video was captured of pre-broadcast preparation and live broadcast activity within the outside broadcast (OB) truck. - Surveys completed by internal participants for the FA Cup Final live football trial have also been collected. In all cases consent was given through completion

Section	Objective	Did we do it?
		of participant agreements and data collected was anonymised as detailed in this DMP. BT and Illuminations were responsible to clear the rights in the content that was used and viewed by participants during the trials (rights to use the content were cleared for R&D purposes).
1.1	The data generated by each pilot and its preceding experiments will be evaluated to assess the success of the assumptions and designs supported by the implementation of the technology and the delivery of the experience.	We have completed analysis of the Theatre-At-Home questionnaires and of the young peoples' perceptions of different presentations of football on TV. We have collected all data for the MotoGP-At-Home service trial which included questionnaires and system logging data and it has been analysed and written up in the work package 4 deliverable - D4.5 MotoGP Trial Evaluation Results. Analysis of the Theatre-In-Schools data has been reported in the work package 3 deliverable - *D3.5 (User experience results: interactions for Theatre-In-School)* and in the work package 4 deliverable - *D4.7 (Theatre-In-School evaluation results)*. Evaluation of the live football service trials and the use of live production tools has been completed and reported in *D4.6 (Football trial evaluation results)*.
1.2	The insights drawn will be published along with summary data. Anonymised raw data may help other researchers draw further conclusions through re-analysis, comparison or combination with other data.	Project deliverables and industry/academic conference papers referenced data summaries. All data was anonymised in these documents. The consortium has no intention to publish raw, anonymised data at the end of the project due to its limited utility.
1.4	The data will be in the form of transcribed interviews and questionnaire responses, data analytics of screen usage and device interactions and video recording of video chat and observation	Theatre-At-Home and Theatre-In-Schools service trials used interview data (stored in password protected files on encrypted project hard disks and in BSCW); online survey data (collected via Survey Monkey, an online survey tool, and password protected); and behaviour data (password protected files). All data is anonymised using randomly generated participant IDs, not real names. Young peoples' perceptions of different presentations of football on TV has been written up for an NEM conference and presented. The data

Section	Objective	Did we do it?
		and work have been submitted in the work package 3 deliverable D3.3 – User Testing Results: Interactions for MotoGP – where we discuss in more detail the design thinking work that went into the Football prototype service.
		For the MotoGP-At-Home service trial - survey & interview data is stored in password protected files. All data is anonymised using randomised participant IDs.
		Data analytics of screen usage and device interactions were collected for MotoGP-At-Home, whilst system logs offering further insight were collected for all service trials.
		Video observations were captured of pre-broadcast preparation and live broadcast activity within the outside broadcast (OB) truck to inform the live Football technical trial.
		Surveys completed by internal participants for the live football trial have also been collected.
1.4	This behavioural data may be useful to anyone wishing to understand the potential for multiscreen environments to enhance the experience of drama and sport. In addition, researchers or broadcast practitioners may be interested in considering the data to learn general lessons across other genres of entertainment.	To date, none of the behavioural data from Theatre-At-Home or MotoGP-At-Home service trials have been shared with anyone outside the project team to re-analyse. Summaries of behaviour data have been shared via industry/academic conference papers. MotoGP data has been shared with BT Sport and will be shared with BBC Sport.
2.0	Quantitative and qualitative data (for example, ratings and transcribed comments) obtained from participants in user research within 2-IMMERSE, will be stored safely and PD will be retained for no longer than is necessary for the purposes of the research.	All temporary data sets were erased after use. Video and audio recordings were transcribed and anonymised after they were captured, before being deleted. During the trials data was stored in BSCW and locally on encrypted project hard disks.
2.0	Any data which could identify an individual participant (including name, age or demographic class) will be encrypted, stored separately from evaluation data,	Only anonymised data was retained by the project. Personal data was captured and used temporarily to facilitate user testing and lab trials. This data was erased after it was initially processed.

Section	Objective	Did we do it?
	and retained for no longer than is necessary for the purposes of the research.	
2.0	All data will be handled in accordance with both EU regulations around data protection, and national government regulation in the country where the study takes place.	All partners acknowledge that in relation to the processing of the personal data, they must comply with their respective obligations under all relevant data protection legislation, including GDPR. This DMP is part of the measures implemented to identify processing of personal data and to increase compliance and transparency. It also includes details of some measures adopted by the consortium as a whole to minimise personal data and to mitigate data protection related risks.
2.1	Transcripts of interviews will be created using an existing speech to text system	We took audio recordings and transcribed the data manually and also used an external agency [TakeNote]. Audio files and transcriptions are stored on BSCW. No personally identifying data was used, except for first names which weren't included in the resulting transcripts. (No surnames, addresses or phone numbers, etc. were recorded.) For MotoGP, interviewers from the external research management company [Acumen Field Ltd] transcribed the data during the interviews. No audio recordings were taken. Participants who completed the trials in BT/BBC were video recorded, but random participant IDs and/or first names were used, and files were password protected. First names were only used to conduct trials safely and reliably on BBC/BT premises. Personal details were not included in transcripts of the recordings. Video recorded has eventually been deleted after anonymised data was extracted? For the assessment of different presentations of football presentations by young people this work did not involve interviews, only scores.
2.1	Computer Assisted Qualitative Data Analysis Software [CAQDAS] will be used to manage this data. Both qualitative and quantitative data will be made available through publications and through metadata tagged on-line datasets where possible.	CAQDAS software was not ultimately used to manage data. For the Theatre-At-Home service trial data analytics, the logging and monitoring infrastructure used a secured Elastic Stack instance provided within Mantl, a privately hosted open source platform developed by CISCO. This infrastructure enabled logs to be generated by all 2-IMMERSE services, as well as each Client Application (running on a TV emulator or

Section	Objective	Did we do it?
		companion device), to be time-stamped and aggregated using a single consistent logging format. Logs were viewed, analysed and interpreted using the Kibana web application (a data analysis and visualisation tool).
		For all subsequent service trials, a privately hosted instance of Rancher[9] was used with a secured ElasticSearch instance for service and client log data aggregation and analytics as previously. We also used Google Analytics as a complementary solution for logging of user interactions with DMApp Components. Unfortunately, it was not possible to achieve this for the Theatre-At-Home trial within the time available, but was successfully used for the MotoGP service trial.
		We used Survey Monkey analytics for Theatre-At-Home online surveys.
		For the MotoGP-At-Home trials, the external research management company [Acumen Field Ltd] used a bespoke survey app which allows for data to be collected offline (in participants homes) and then uploaded later. The survey software they use to power this is LimeSurvey.
2.1	Code specifications will be released through project deliverables. The 2-IMMERSE system architecture is built from a number of defined services, each scoped with specific roles and responsibilities. These services will be designed to scale elastically, running the required number of instances to meet dynamic load requirements.	Source code for platform services, DMapps and DMApp components, production tools, infrastructure configuration etc. are securely stored and backed up using a private Gitlab instance hosted by consortium partner IRT. A core set of repositories will be published under the Apache Licence v2.0 on GitHub at the end of the project.
		Community management responsibilities for each repository have been divided up amongst the principle authors or contributors within the consortium.
		Attribution of all other consortium partner contributions will be included in the software released.
2.1	As described in work package 2 - *D2.1 System Architecture*, the project will use Mantl, http://docs.mantl.io/en/latest/ a modern platform for rapidly	The Theatre-At-Home service trial was run on a Mantl cluster hosted by CISCO, with anonymised usage logs stored in a secured ElasticSearch Cluster and logs archived to a separately secured

---

[9] https://rancher.com/

Section	Objective	Did we do it?
	deploying globally distributed services. Mantl provides an integrated set of industry-standard open-source components. It is cloud infrastructure provider agnostic, and can be deployed on AWS, OpenStack, Vagrant, Bare Metal etc. Mantl is licensed by 2-Immerse partner Cisco under the Apache Version 2 License.	ElasticSearch instance. The MotoGP-At-Home service trial was run on a privately hosted instance of a Rancher managed cluster on AWS in the EU. As previously, anonymised usage logs are stored in a secured ElasticSearch Cluster, with anonymised user interaction data for DMApp components using Google Analytics. Although there is the capability of storing user preferences into a persistent user profile, this capability was not used in the MotoGP trial. The Rancher managed cluster was also used for the Live Football technical trials and Theatre-In-Schools user trials.
2.2	Our default intention is to make the results of the project available through Open Source repositories. We are exploring the use of OpenAIRE https://www.openaire.eu/ repository https://zenodo.org/ Zenodo.	We are delivering a functional subset of the 2-IMMERSE software repositories as an open-source release on GitHub at the end of the project as opposed to publishing via OpenAIRE.
2.3	One of the goals of the project is to provide a functioning implementation of HbbTV 2.0 and inform the next generation of that standard. Data management of the code created will be the responsibility of the technical team under Cisco's leadership and the user data management will be the responsibility of the social science team led by BBC.	The BBC social science team led by the BBC has taken responsibility for data management within the project. Nonetheless all partners acknowledge that in relation to the processing of the personal data, they must comply with their respective obligations under all relevant data protection legislation, including GDPR. Each consortium partner is acting as a data controller in respect of the processing of the personal data on its own behalf, or while personal data is in its possession or control. As noted previously, source code for platform services, DMApps and DMApp components, production tools, infrastructure configuration etc. are securely stored and backed up using a private Gitlab source control system instance hosted by project partner IRT. The MotoGP-At-Home service trial was ported to run on real HbbTV2.0 television in addition to an emulated implementation.
2.4	Data when generated will be kept in secure encrypted repositories within the partners'	Data has been kept in secure encrypted repositories within the partner organisations and centrally within BSCW. The 2-IMMERSE consortium

Section	Objective	Did we do it?
	organisations. when it is released for wider availability it will be through secure stores such as OpenAIRE on the CERN sponsored Zenodo facility.	does not intend to release raw data sets where these are of limited wider value. Therefore, summaries have been published in papers and work package deliverables.

9 Data Management Questionnaire

The following data management questionnaire has been completed by all consortium partners to map all datasets captured and processed as part of the 2-IMMERSE project. This was done to identify risks and any associated impact associated with use of personal data. An illustration of the data mapping technique has been included in Appendix A - Data Mapping Diagrams.

#	Please provide a full description of the data	Please indicate Data Category(ies): **Letter A** for personal data (any data that can be used to identify individuals); **Letter B** for resulting data that has been created as part of the Project and may be used as part of Dissemination activity; **Letter C** for other sensitive data	Please indicate whether this is an existing data or data that is expected to be generated in later stages of the Project?	Please provide details about consents that have been obtained from data subjects and rights owners	Where is the data stored and what measures are implemented to keep it safe?	Who has or received access to the data?	Please provide details about the purpose(s) the data is being processed	Where data includes personal data, please provide details about the need to process PD and whether measures to minimise PD have been considered (e.g., pseudonymisation* or anonymisation) *"Pseudonymisation" refers to the technique of processing personal data in such a way that it can no longer be attributed to a specific "data subject" without use of additional information which must be kept separately and be subject to technical and organisational measures to ensure non-attribution
1	**MotoGP service trial & production tool trial 2:** A "Profiles & Schedules" password protected Excel	A (Contact details, including names, email	Existing	Consent forms for data usage with positive opt-in were	Password protected Excel and Word documents stored and on [Acumen's] internal	BBC (UK) and CWI (the Netherlands) have received limited versions of the data. Participant names and	To recruit participants for service trials. The 'profiles &	At the end of the trials all participant's personal data was anonymised by BBC and CWI. Personal data is only

	spreadsheet containing participant names, phone numbers, email addresses and demographic screening questionnaire answers collected by an external UK research management company [Acumen] on behalf of the consortium. For MotoGP-At-Home, this spreadsheet was edited to only include participant IDs alongside participant's screener responses and the feedback they gave in the trials. No other personal data, names, identifiable numbers or addresses were in this copy of the spreadsheet given to the consortium. For the lab-based Production Tools trial (held September 2018 which [Acumen] recruited for, but BBC/CWI facilitated in the BBC lab), the participant's name, phone number and email address, were captured in the 'Profiles & Schedules' spreadsheet given to the consortium by the external company.	addresses and phone numbers of potential trial participants)		collected by [Acumen] on behalf of the consortium and were deleted where potential participants have eventually not taken part in the trials.	server hosted in the UK accessible only by the company staff.	contact details were replaced by anonymous/random participant IDs for the MotoGP-At-Home service trial and were retained only for the second production tools trial.	schedules' password protected Excel spreadsheet informed who to expect at the BBC reception for lab-based trials. It also meant BBC/CWI could ring/email the participant directly for time related coordination purposes.	processed on [Acumen's] password protected, internally hosted server. All the 'profiles & schedules' spreadsheet containing this personal data were deleted once the trial was completed.
2	**MotoGP service trial:** Names and contact details of facilitators from external research management	A (Contact details, including names, email	Existing	Facilitators are hired by [Acumen] under a contract.	Password protected Excel and Word documents stored on [Acumen's] internal server hosted in the UK	Contact details of facilitators were shared with all consortium partners by password protected spreadsheet and exchanged	Facilitate recruitment and participation in trials.	Data only included contact details necessary for the coordination of the trials. The spreadsheet shared with

	company [Acumen].	addresses and phone numbers of facilitators. There were 6 facilitators.)			accessible only by company staff.	between all persons taking part in the trials, including (i.e., facilitators and the participants they recruited).		all consortium partners was deleted after completion of the service trials.
3	**MotoGP service trial & production tool trial 2:** Anonymous user experience questionnaires	B		Existing	Consent forms described in row 1 of this table covered collection and processing of anonymous questionnaire results. The consent forms were used to recruit candidates.	Captured using Limesurvey online survey tool and stored in an [Acumen] administrated AWS EU west region 1 data centre and secured by company login credentials. This server has been decommissioned and the data securely erased.	BBC (UK) and CWI (the Netherlands) have received copies of the data, anonymised by removal of participant names and contact details and replaced by anonymous/random participant IDs. This data was deleted after it had been summarised and published.	(i) Provide qualitative data on benefits of multi-screen object-based broadcasts; (ii) publication of research papers, work package deliverables and presentations.	Participant's personal data was anonymised such that the data could no longer be attributed to the participant by BBC and CWI.
4	**MotoGP service trial & production tool trial 2:** Video recordings of participants in lab-based evaluation.	A		Deleted	Consent forms granting permission based on positive opt-in to film participants were collected for the consortium by the external research management company [Acumen].	Videos were only associated with anonymised participant IDs and stored on encrypted external hard disks kept in locked cupboards at the BBC in the UK. These videos have been securely erased.	BBC (UK) gave CWI (the Netherlands) a copy of the videos. These videos, including all copies were deleted by the BBC and CWI.	(i) Provide qualitative data on benefits of multi-screen object-based broadcasts through transcription of notes and quotes; (ii) publication of research papers, work package deliverables and presentations.	The videos were processed to obtain quotes and notes, which were anonymised such that the data could no longer be attributed to the participant. The videos were securely erased once they were no longer needed.
5	**Production tool trial 2:** Photographs of participants	B		Existing	Consent forms granting permission to	Photographs were captured and stored on an encrypted removable	No one outside the partner's (CWI's) organisation has received access to the	Publication of research papers, work package	

	in experiments captured by CWI in the Netherlands.			use photographs in research publications under the condition that faces would be blurred to hide the identity of participants.	hard disk by CWI in the Netherlands. All faces were blurred to anonymise participants in the pictures and the original photos were deleted.	photographs.	deliverables and presentations.	
6	**Theatre-in-schools and Theatre-at-home:** Participant and parent/guardian contact details	A	Deleted	Consent forms for data usage with positive opt-in were collected by BBC and the external UK based production services supplier contracted by Illuminations [Donmar Warehouse]. This was used to recruit participants.	Participants in Theatre-at-home were recruited from friends and colleagues of the 2-IMMERSE consortium. Contact details were collected via email and consent forms were also exchanged by email.	Except the production services supplier, no one outside the consortium has received access to the data.	To recruit participants for service trials.	Children's details were not collected. All details were erased after the service trial was completed by both the BBC and production services supplier [Donmar].
7	**Theatre-in-schools & Theatre-at-home:** Video & audio recordings of interviews with producers and trial participants	A B	Deleted Existing	Consent forms for data usage with positive opt-in were collected by the external UK-based production services supplier contracted by Illuminations	Recordings were stored on encrypted removable hard disks by Illuminations (UK), and on dictaphones (audio only) by BBC (UK). Audio was extracted, anonymised and given to an external UK based supplier contracted by the BBC [TakeNote] or	Theatre-In-Schools audio recordings were given to [TakeNote], a third-party audio transcription service. No one has received access to the raw data.	(i) Provide qualitative data on benefits of multi-screen object-based broadcasts; (ii) publication of research papers, work package deliverables and	Anonymous audio transcriptions were derived from audio recordings of the interviews with producers and trial participants. The recordings were erased once transcribed.

					Consent forms for data usage with positive opt-in were collected by the BBC and used to recruit participants.	the BBC for transcription.		presentations.	
8	**Theatre-In-home:** Anonymous Online survey, questionnaire and phone interviews with requited trial participants.	B		Existing	Consent forms for data usage with positive opt-in were collected by the BBC and used to recruit participants as part of the online survey.	Data was stored securely in a BBC Survey Monkey account. The account and data have subsequently been deleted. Anonymous quotes from the phone interviews were transcribed and collated in a password protected Word document and stored on an encrypted hard disk while being analysed, then were erased.	Anonymous data has been exported and uploaded to BSCW, a document management system hosted by Eurescom in the EEA, where all consortium partners have access.	(i) Provide qualitative data on benefits of multi-screen object-based experiences; (ii) publication of research papers, work package deliverables and presentations.	
9	**Theatre-In-schools:** Paper-based questionnaires	B		Existing	Consent forms for data usage with positive opt-in were collected by The external UK based production services supplier contracted by Illuminations [Donmar] and the BBC and used to screen participants.	Anonymous questionnaire data was collated in a password protected Excel spreadsheet and stored on a secured BBC (UK) laptop, and in BSCW, a document management system hosted by Eurescom in the EEA.	No one outside the 2-IMMERSE consortium has received access to the raw data within the document management system. Results will be published in deliverables and papers.	(i) Provide qualitative data on benefits of multi-screen object-based experiences; (ii) publication of research papers, work package deliverables and presentations.	For theatre-in-school, no personal data was captured. Students (between the ages of 14-15yrs) used 'Group Number and a letter -A/B/C' (e.g., Group1A; Group1B; etc) on their paper-based questionnaires.
11	**Football:**	A		Deleted	Formal consent forms were not	Video content, photos and action spreadsheets	A quad view video was created and made available	i) Ethnographic studies were	No names were used to identify production staff in the

	Ethnographic Study Raw video and photos of pre-broadcast preparation and live broadcast activity within a BT Sport outside broadcast (OB) truck.			used. Rights to use sport related content were cleared for R&D purposes by BT and Illuminations.	are stored on an encrypted BT laptop.	to all consortium partners as well as a time-stamped action spreadsheet via BSCW. Only BT has access to the raw video.	undertaken for the football service trials to understand the director's role in creating the broadcast mix of video, graphics and commentary narrative for the match.	video. BT Sport made people aware that filming of staff would be taking place. This data will be securely deleted at the end of the project.
12	**Football:** Surveys completed by BT Sport staff for the live football service trial to gauge interest and feedback on the value of the experience offered.	A	Deleted	Formal consent forms were not used. Questionnaire was limited to internal BT Sport staff.	The questionnaire data is stored on a BT (UK) laptop, which is encrypted and password protected. Original responses were received via email and will be disposed of under the normal email retention policy (2 years).	No one outside BT's organisation has received access to the raw data.	i) Provide qualitative data on benefits of multi-screen object-based experiences; (ii) publication of research papers, work package deliverables and presentations. This is a very small/thin dataset and consequently, the consortium doesn't have any plans to publish an analysis.	The data was partially anonymised (first names only) for storage. If the 2-IMMERSE consortium decides to publish a summary of the results, the survey data will be fully anonymised, otherwise it will be deleted by BT at the end of the project. This data will be securely deleted at the end of the project.
13	**Theatre-in-schools, Theatre-at-home, MotoGP and**	B (micro service &	Existing	Consent forms with positive opt-in for data	Stored in a secured Elastic Search instance in AWS EU West Region	All consortium partners have password protected access to the logging data in Elastic	(i) Provide quantitive data on benefits of	We did not collect client IP addresses in ElasticSearch logs. All client log messages

				collection and processing were completed by all trial participants and managed by external research management company [Acumen], BBC and an external UK based production services supplier [Donmar]. See row 1 of this table for further details. This covered the collection, storage and processing of system logs. The consent forms were used to screen candidates.	1.	Search. No one outside the consortium has received access to the raw data. A small number of logs will be kept for reference after the project ends.	multi-screen object-based broadcasts; (ii) publication of research papers, work package deliverables and presentations; (iii) support diagnosing software faults	were sent via the 2-IMMERSE Logging Service and conform to the message format defined in D2.3 and D2.4. No IP address fields have been used.
	Football service trials: Anonymised client/server system logs. No personally identifiable information was captured.	component specific log messages)						
14	**MotoGP service trials:** Anonymous user interaction analytics data.	B	Existing	Consent forms with positive opt-in for data collection and processing were completed by all participants. See row 1 of this table for further details. The consent forms	Stored by Google and accessible only via a single password protected user account. UserID and ClientID features were not used in Google Analytics (GA). Only anonymous identifiers were captured in custom	No one outside BT Sport has received access to the raw data. Login credentials were not shared with any other parties. Processed data is available to all consortium parties via BSCW.	(i) provide quantitative data on benefits of multi-screen object-based broadcasts; (ii) publication of research papers, work package deliverables and presentations.	No personally identifiable information was captured because all IDs were anonymous.

| | | | | were used to screen candidates. | dimensions and metrics.

Sharing was disabled for all analytics data captured in GA.

Temporary Google accounts, created by the consortium, were used to sign into loaned consortium devices (tablets and phones) during the trials

2-IMMERSE consortium equipment was used in all service trials as opposed to the participant's own devices.

GA doesn't expose IP addresses for Analytics reporting and we haven't created any custom dimensions which carry them. We did not enable IP Anonymisation in our GA config at first, but subsequently enabled it to mask all IP addresses which could have been used by Google to estimate the location of users.

Only anonymous data was collected. Data is kept by Google for approximately 18 | | | |

2 IMMERSE

				months before being deleted. Advertising features of Google Analytics wasn't enabled. Raw analytics data was exported into Excel where Pivot tables and charts were created. The resulting data was stored on an encrypted BT laptop and imported into BSCW where all consortium partners can access it.			

Table 1 Data Management Questionnaire

10   Appendix A - Data Mapping Diagrams

A data mapping process was used to help answer the **Data Management Questionnaire** and identify all the datasets captured or generated by the project. Illustrations captured during this exercise are included here to show the data planning process undertaken. They do not accurately represent or imply how data was actually captured, processed or stored during the project. They are snap-shots of an incomplete and unverified data map that were kept as illustrations only.
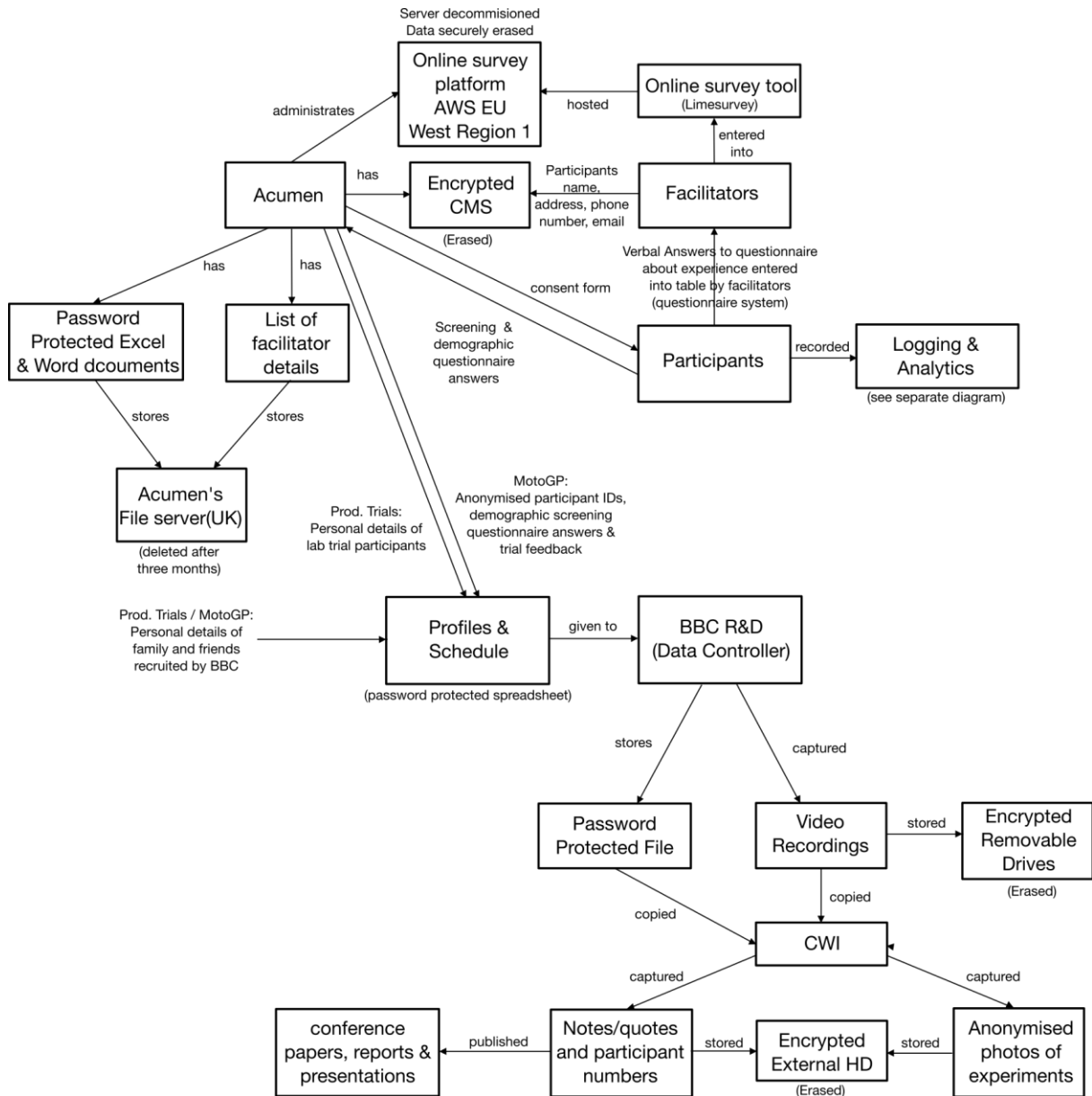
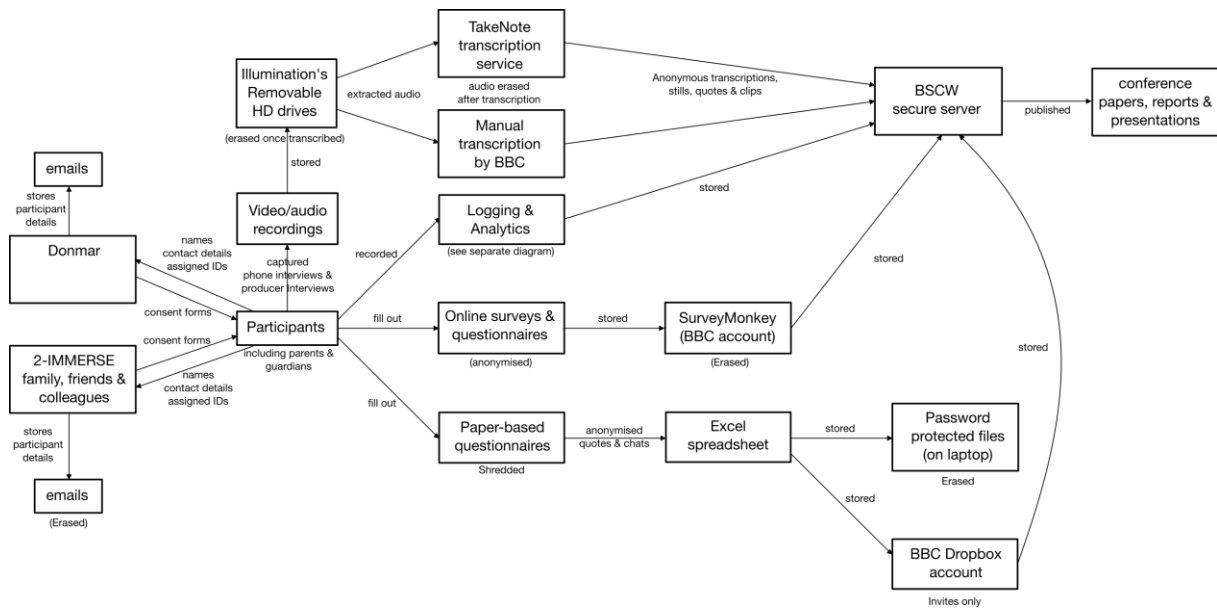**Figure 1 Example data map for MotoGP service trial and Production Tools trial #2**

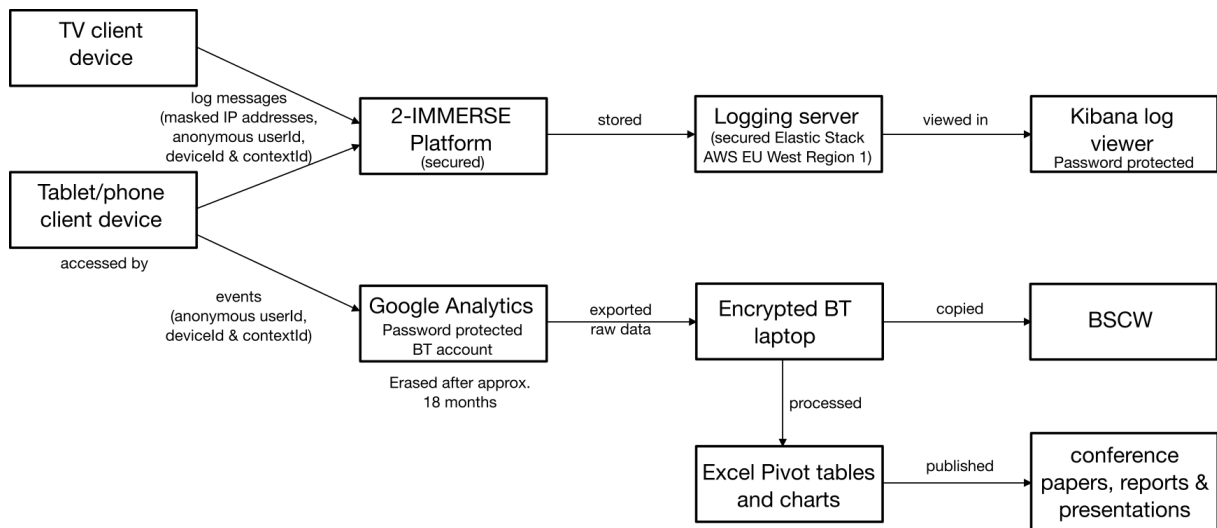**Figure 2 Example data map for Theatre-At-Home and Theatre-In-Schools service trials**

Figure 3 Example data map for analytics and logging data used in service trials