Multimedia **A**uthoring and **M**anagement using your **E**yes and **M**ind

H2020-ICT-2014-644780

D1.4

Data Management Plan

Dissemination level:	Public (PU)
Contractual date of delivery:	Month 6, 31/10/2015
Actual date of delivery:	Month 6, 31/10/2015
Workpackage:	WP1– Management
Task:	T1.2- Technical and quality management T1.3 - Management of knowledge and Intellectual Property Rights
Type:	Report
Approval Status:	Final
Version:	0.4
Number of pages:	27
Filename:	D1 4_DataManagementPlan_Final.docx

Abstract: D1.4 determines the strategy by which the research data generated by the project will be made open for maximizing their re-use. In particular, we initially describe the context where data will be collected by MAMEM, as well as their nature. Then, we discuss about the data management and sharing practices and lay out our ethical and confidentiality considerations.

co-funded by the European Union

Copyright

History

Version	Date	Reason	Revised by
V0.1 (alpha)	04/09/2015	Table of Contents – Checked for completeness by the consortium	Spiros Nikolopoulos, Elisavet Chatzilari and Ioannis Kompatsiaris
V0.2 (beta)	09/10/2015	Ready for internal review version	Elisavet Chatzilari and Spiros Nikolopoulos
V0.3 (beta)	27/10/2015	Address comments from internal review	Elisavet Chatzilari and Spiros Nikolopoulos
V0.4 (final	31/10/2015	Address final remarks from the consortium and finalize the document	Elisavet Chatzilari, Spiros Nikolopoulos and Ioannis Kompatsiaris

Author list

Organization	Name	Contact Information
CERTH	Spiros Nikolopoulos	nikolopo@iti.gr
CERTH	Elisavet Chatzilari	ehatzi@iti.gr
CERTH	Ioannis Kompatsiaris	ikom@iti.gr

Executive Summary

The consortium will support the Open Data Initiative and will ensure that part of the collected data will be made available to the public, provided that privacy and security issues are addressed (see Section 5). Our goal in this document is to: a) specify the data that will be collected during the lifetime of MAMEM; b) investigate the best practices and guidelines for sharing the project outcomes and facilitating open access to research data, while ensuring compliance with the established ethical and privacy rules; and c) define how the data collected in the project will be made available to third parties in contexts such as scientific scrutiny, peer review and use for research purposes;

Abbreviations and Acronyms

CDISC	Clinical Data Interchange Standards Consortium
GDF	General Data Format
DMP	Data Management Plan
EEG	Electroencephalogram
GSR	Galvanic Skin Response
HCI	Human Computer Interface
NMD	Neuro-muscular disorders
PD	Parkinson Disease
SCI	Spinal Cord Injuries
XDF	Extensible Data Format

Table of Contents

List of Tables

1 Introduction

During the lifetime of MAMEM, data of different nature will be generated and collected. These data are user-related and thus require a clear plan on how they are to be managed, i.e., stored, accessed, protected against unauthorized or improper use, etc.. Thus, the main goals of MAMEM's Data Management Plan (DMP) are:

1. Outline of the types of data already generated and/or foreseen for generation at this stage of the project, including the context and procedures of this generation, as well as the degree of privacy and confidentiality of the data.
2. Outline of the protocols that will be followed to assess the generated/collected data with respect to their sensitiveness.
3. Outline of the data acquisition plan for the duration of the project.
4. Outline of the measures that are foreseen for the adequate management of the data from the ethical and security points of view.

In accordance with the guideline on data management in Horizon 2020 [1], the following aspects are discussed in the DMP: a) dataset reference and name, b) dataset description, c) standards and metadata, d) data sharing and archiving and preservation (including storage and backup).

The remainder of the deliverable is structured as follows. Section 2 describes the settings in which data is and will be generated. In Section 3, the data collection processes are outlined. In Section 4, the data are handled according to the different categories regarding confidentiality. In this respect, there are data which are confidential and need special protection, data which are not confidential and can be shared, as well as data which depend on the informed consent of the participant. Section 5 outlines the ethical and confidentiality considerations with respect to the MAMEM data and Section 6 summarizes the deliverable.

2 MAMEM context and data description

The objectives of data gathering in MAMEM is to: a) define the user requirements for the MAMEM system including the usage scenarios, b) develop the user profiles, c) develop and assess the machine learning techniques, and d) evaluate the system for its ability to integrate our end-users back into the society. Specifically, MAMEM is meant to provide the necessary technology so that the disabled can interact with multimedia content and the computer in general, while providing motivating mechanisms opting to ensure their inclusion into the society. Thus, MAMEM should:

- Find out about the situation of the users, their needs and requirements from the system;
- Capture, record and make available at the necessary scale, real-time and accurate information about eye-movements, brain electric signals and bio-measurements.
- Understand and translate the raw signals from the eye tracker, the EEG recorder and the GSR sensor into meaningful mouse/keyboard actions by developing novel signal processing techniques robust to noisy environments.
- Specify user models and profiles based on their (dis-)abilities, emotions, sensitivity to specific persuasive strategies, social abilities, extent to which training is needed, and follow a set of persuasive design principles to develop interfaces that will effectively stimulate users to use them and encourage their behavioural change (e.g. increase adherence to therapy/exercises, increase trust in the interface).
- Have impact on future research on natural computer interfaces and assistive technology for the disabled, developing a user-friendly communication interface between the disabled and the computer.
- Assess the degree of success in bringing disabled people back to the society as a result of their newly acquired ability to manage and author multimedia content.

All technologies to be developed in order to achieve the above goals are data-driven. This means that a variety of data will be collected during the lifetime of the project. We can distinguish between three types of data that will be collected at three different stages of the project. First, data will be collected from a detailed literature review on the three use cases of the project (i.e., subjects with Spinal Cord Injuries, subjects with Parkinson Disease and subjects with Neuro-muscular disorders). In addition, focus groups with the participation of health professionals will be conducted, and questionnaires that will be filled by both the patients and their caregivers. The objective of these data is to define the user based system requirements and guide the development of the user profiles and the persuasive design principles. Second, in order to develop the algorithms that will facilitate the interaction through eyes and mind, raw signal data from the eye tracker, the EEG recorder and the GSR sensor will be collected from able-bodied users that will be willing to facilitate the development process. Finally, during the clinical trials data will be collected from the end users in order to measure their social inclusion prior and after using the MAMEM technology and adapt the user profiles and the persuasive design principles based on the received feedback. This will include data originating both from monitoring their interaction with the computer, their activity in social media, as well as from properly structured questionnaires.

2.1 Data to be collected for requirements

The first goal of the data collection at this stage of the project is to provide specific requirements for the MAMEM platform derived from the fact that the end users will be subjects with disabilities (i.e., '*clinical requirements*') for the three examined cohorts: Spinal Cord Injury (SCI), Parkinson's disease (PD) and Neuromuscular Disorder (NMD). In addition, the data collected at this stage will serve as input to WP5 and will be utilized towards building effective profiles of the end users and defining the persuasive design protocols.

There will be three different data collection protocols during this stage, each of which will be applied to each of the three cohorts. More specifically, for each patient cohort data will be gathered with the following strategies:

1. **Literature review:** The objective is to summarize the clinical requirements including the special physical requirements for wearing and operating the platform and additional requirements derived from mental and regular clinical treatment aspects.
2. **Focus group**: The objective is to record and summarize the requirements of health professionals by interviewing three different, medical condition-specific, focus groups.
3. **Questionnaires:** The objective is to identify the usage scenarios of the MAMEM system (e.g. transfer photos/videos to the computer and upload to their social account).

2.2 Data to be collected for training signal processing algorithms

The objective of MAMEM is to provide a natural Human Computer Interaction (HCI) interface, which will translate the raw signals captured with the three utilized sensors (i.e. eye tracking, EEG and GSR) into meaningful commands. Thus, data will be collected with these sensors at all stages of the project where MAMEM technology will be used and evaluated. In addition, data will be collected with the three sensors during the development of the algorithms for human-computer interaction through eyes and mind. This is essential in order to successfully analyse, understand and translate the raw signals to meaningful commands, as well as train the machine learning algorithms.

2.3 Data to be collected during the clinical trials

The goal of the data collection at this stage of the project is to assess the degree of success in bringing disabled people back to the society as a result of their newly acquired ability to manage and author multimedia content. Given that social integration is a multi-dimensional concept that cannot be easily assessed, the main objective is to research and adopt a methodology for monitoring social inclusion in a controlled group of subjects with disabilities (e.g. indicators for quantifying social integration, observing social activity and digital productivity). This will be achieved by collecting data in two ways:

1. **Objective indicators**: Measure the activity of a person in social networks by establishing a number of indicators that can be automatically measured (e.g. number of on-line friends/connections/followers, number of posts, number of likes/comments/shares per post, overall number of distinct users that liked/shared/commented on a post, expansion of the user's social graph, etc.).
2. **Subjective indicators**: implicitly quantify social integration by monitoring any significant change in the overall throughput of the patient's labour. In this case the information

source will be the subjective feedback provided by the patients in the form of questionnaires.

3 Data collection

3.1 Data collected for requirements

In order to obtain clinically driven requirements for MAMEM, taking into account the particularities that arise from the clinical condition of each patient cohort of users, we adopted the following approach by collecting data from three different sources

1. We conducted a thorough **literature survey** to extract the needs for each patient cohort based on general knowledge for each disability (e.g. clinical background, psychology and quality of life, computer use, factors limiting the use of MAMEM technology by each patient type, etc.).
2. We conducted three **focus groups**, each one centred on each patients' cohort, in which clinicians and para-clinicians from different disciplines, 'brain stormed' about the current difficulties, current solutions and the requirements to the add-ons that MAMEM can provide.
3. We constructed **questionnaires** aimed at both users and care givers in order to extract clinically driven requirements from the point of view of the end user. Further we obtained **ethical approvals** in order to approach and interview patients and care givers with these questionnaires.

The data collected by this process is summarized in two deliverables (D6.1 [2] and D6.2) and is the basis on which we will define the usage scenarios, the user requirements and limitations and modify the inclusion/exclusion criteria for the end users.

3.1.1 Data from focus groups

In each of the clinical sites (i.e. SHEBA, MDA and AUTH), a focus group of professionals gathered in order to discuss the particularities of each use case patient type. Table 1 lists the experts that were included in the focus group of the corresponding use case. Each focus group assembled a report including the following data:

a. Description of the procedure.
b. Address to specific items at the focus of discussion.
c. Quotes and citations.
d. Integrative summary of the outcome.

Table 1: Composition of the conducted focus groups

Use case	SCI	PD	NMD
Focus group participants	Medical Doctor	Neurologists - 2	Doctor
	Rehabilitation nurse	Psychiatrist	Nurse
	Social worker	Physiotherapists – 2	Social worker
	Health psychologist	Social worker	Health psychologist
	Occupational therapist	Occupational therapist	Occupational therapist
	Physiotherapist		Physiotherapist

Speech therapist

3.1.2 Data from questionnaires

Table 2 provides indicative examples for the different types of data that are collected through questionnaires.

Table 2: Type of data collected through the questionnaires

Topics	Examples	Use Case
Demographic and clinical information	Demographic information, clinical information	All
Computer working environment, habits and difficulties	Computer use habits, Difficulties, Current working environment, Assistive devices	All
Needs, missing functions and demands of improvements	Needs / ability to control assistive devices, computer usage tools, etc.	All
Disease specific Clinical information	Diagnosis, wheelchair, bedridden, parts of immobility / tremor, etc.	All

3.2 Data to be collected for training signal processing algorithms

Except from the benchmarking datasets that already exist in the literature (Section 3.4), we will collect raw signal data during the project lifetime in order to provide state-of-the-art signal processing algorithms that will effectively translate the raw signals of the EEG, eye-tracking and GSR sensors to commands. The context within which the data collection will occur will be defined according to the specific needs of each experiment. Researchers and engineers will be used as subjects for this process. The specific datasets that will be collected during this phase of the project depend mainly on the interaction paradigms that will be employed in our project. Table 3 provides indicative examples for the different types of data that will be collected from the three sensors.

Table 3: Type of data collected from the sensors

Sensor	Data Types	Possible control
Eye-tracker	trajectories of eye-movements, fixation points (coordinates)	Detect the area in the screen where a control should happen
EEG sensor	electrical activity (measured in micro-volts) from 64 channels, event markers, sampling frequency	Execute specific commands (e.g. back, left, etc.) or add context to the commands.
GSR sensor	Skin conductance (in microSiemens), sampling frequency	Provide input to the persuasive design (e.g. select a specific assistive dialogue if the patient is frustrated), system evaluation

3.3 Data to be collected during the clinical trials

During the clinical trials the following data will be collected:

1. **Raw signal data**, which will be used by the MAMEM platform to allow the end user to interact with the computer. This includes signals from the EEG, eye-tracking and GSR sensors, as mentioned in Table 3.
2. **Interaction logs** of the end user with the computer. This will consist of activity logs (e.g. which buttons where pressed, in what order, what were the main sources of delay), as well as logs in social media activity (e.g. the number of on-line friends/connections/followers, number of posts, number of likes/comments/shares per post, overall number of distinct users that liked/shared/commented on a post, etc.). The exact type of data to be collected will be defined in D7.1.
3. **Questionnaires** filled by the end users and their care givers, similar to the ones constructed during the clinical requirements definition. The purpose of these questionnaires will be to evaluate the satisfaction of the end users with respect to the system.

The raw data signals of the users participating in the pilot trials will be wrapped up in a research dataset consisting of synchronized signals from an eye tracker, an EEG recorder and a GSR sensor. The value of this dataset will reside on the protocol followed during the pilot trials, as well as the variety of the target groups participating in these trials (i.e. different cause for their disabilities, different age group, different levels of mental health and interaction potential). This dataset will become available through our website (following the guidelines for sharing data as detailed in Section 4.4) and a number of actions will be undertaken so as to make it available through centralized repositories for bio-signals, such as [3] and [4].

3.4 External sources of data

The fields of extracting meaningful information from EEG and eye tracking signals have been well established for many years. During these years, there have been many benchmarking datasets released [3], [4], which can also be used within MAMEM for research purposes. The datasets available at [3] are open access datasets of bio-signals that each of them has been collected for different tasks (e.g. spelling, motor imagery tasks). In addition, links to dataset from BCI competitions are included. Similarly, there are three datasets available at [4], which consist of combined data from EEG and eye-tracking sensors. The three different settings include involuntary eye movements during face perception, visual search within natural images and natural reading. These external sources of data will most probably become useful in evaluating the efficiency of our interaction algorithms, or for testing new interaction paradigms. In the case where there is a need to publish our methods side-by-side with these external datasets, we will make sure to acknowledge and link to the source where these data have been originally obtained.

4 Data management and sharing

The data collected in MAMEM can be categorized in two types based on their source. There is data retrieved from external sources (e.g. the web, medical experts, etc.), which typically do not require any special care regarding sensitivity, privacy and confidentiality. The second type of data is generated within the MAMEM project. These are generated by recording data in the form of questionnaires for requirements, raw signals for training the signal processing algorithms, as well as by monitoring the activity of our end-users in social networks and in interacting with their computer. Such types of data are user-related and thus require a clear plan on how they are to be managed (i.e., stored, accessed, protected against unauthorized or improper use, etc.).

4.1 Guidelines for generated data

The pilot trials of MAMEM will take place in Greece (Cities of Athens & Thessaloniki) and in Israel (City of Tel-Aviv). During these pilot trials we will collect information about our research participants that have to do with their clinical profile, their habits and needs in what refers to human-computer interaction. We will also collect information about the signals that have to do with the brain-electrical activity, the eye movements and the bio-measurements of our research participants. In addition, the data may include, but is not limited to, personal information about the user such as: name, date of birth, health state, interests, location, images, or relations to other users. Thus, the data controller in our case will be our end-user partners (i.e. MDA Hellas, SHEBA and AUTH) and the data processor will be all other involved partners (i.e. CERTH, SMI, EBNeuro, UNI KO-LD). In handling these data, we will make sure to comply with national and EU legislation, as well as follow the best practice for ethics in Human-Computer Interaction (Ethics in HCI and Usability). In Section 5 of the DoA [5], detailed guidelines are provided for the activities of: a) data collection, b) storage and transmission, c) retention, d) treating sensitive data, and e) term of usage. In managing MAMEM's data we will make sure to comply with all guidelines specified in the DoA [5]. An example for the informed consent that has been prepared for the researchers participating in the first EEG signal capturing experiments can be seen in Appendix A.1.

4.2 Categories of data based on their confidentiality level

We can categorize the collected data based on their confidentiality level in three categories; i) "open" data, which can be shared openly, ii) "protected" data, which can be shared but the participants have to provide their consent, and iii) "confidential" data, which cannot be shared outside the project. Note that in all cases, the data collection and sharing processes will follow the strict guidelines detailed in the DoA [5] (e.g. anonymization).

Table 4 categorizes the data generated in MAMEM to one of the aforementioned categories. We can see that apart from the sensitive data that have to do with the medical status of our participants, all other types of data can be made public provided that the necessary consents are obtained from the participants.

Table 4: MAMEM data categorized based on their confidentiality level.

Topic	Objective	Data Type	Source	Category

Clinical background, psychology, computer use, limitations	Requirements, Inclusion/Exclusion Criteria	Text	Focus Groups, Literature Review, Patients & Caregivers	Open
Biography aspects (Medical information, psychological disorders, etc.)	Requirements, social inclusion quantification	Text	Patients & Caregivers	Confidential
EEG, eye-tracking and GSR signals	Training the signal processing algorithms	Signals	Researchers	Protected
Interaction logs, and EEG, eye-tracking and GSR signals	Evaluation of MAMEM	Signals, Text	Patients	Protected
Activity in social networks (e.g. likes, followers, posts, etc.).	Assessing the status of social integration	Numbers (summaries)	Patients	Protected

4.3 Data cataloguing

Below we provide details on how we intent to document a dataset before making it public, so as to facilitate smooth identification and access.

4.3.1 Dataset reference and name

Each dataset will be given a unique identifier. At this stage in the project individual data sets have not been formally identified, as they cannot be fixed until the user based requirements are finalised on M7 (D6.2). Once this is done we will be able to develop knowledge of the target datasets required and so can commence populating the data catalogue.

4.3.2 Dataset description

For each dataset in the catalogue there will be a description outlining; the nature and scale of the data, to whom it could be useful, and whether it underpins a scientific publication. Information on the existence (or otherwise) of similar data and the possibilities for integration and reuse would also be included.

4.3.3 Metadata standards for bio-signals

For the biomedical signals we plan to use the General Data Format (GDF) [6], which is a scientific and medical data file format. The aim of GDF is to combine and integrate the best features of all bio-signal file formats into a single file format. The GDF format is often used in brain–computer interface research. However, since GDF provides a superset of features from many different file formats, it could be also used for many other domains. There are free and open source libraries, which provide implementations for reading and writing of

GDF (e.g. BioSig [7], libGDF [8]). For metadata storing, the Extensible Data Format (XDF) [9] and the CDISC (Clinical Data Interchange Standards Consortium) suite are the prevailing standards. Finally, given the widespread use of certain toolboxes for signal processing (e.g. Matlab [10]) we also plan to make our data available in formats convenient for easy processing (e.g. as mat files).

4.4 Data sharing

Depending on the confidentiality level of the data different sharing approaches will be employed.

4.4.1 **Private sharing**

This approach consists in setting up an FTP server by the partners acting as data controllers (see DoA – Section 5 [5]), so as to share within the consortium all data generated in the project. The FTP server will be used to host all different types of data (i.e. “open”, “protected” and “confidential”) and will apply strict accessibility rules, such as anonymization, password protection, and transmission in a 128-bit encrypted form through a secure communication channel (SSL). The task of archiving the data on the server will be carried out by the partner that collected the data (i.e. the data controller), who is also responsible for obtaining the informed consent of the participants from his recordings. The Private sharing approach will be put in place to facilitate the smooth communication of the generated datasets, while ensuring that only the members of the consortium can access them.

4.4.2 **Controlled sharing**

This approach consists in sharing the dataset through the projects web-site and allowing externals to download them, after requesting to provide their personal and affiliation details and to agree with the terms of a data usage agreement (for a template see Appendix A.2). In this case, we will include the types of data belonging to “open” and “protected” categories, whereas for the latter we will only share the part of data where informed consents have been obtained. More specifically, based on the cataloguing information described in Section 4.3 we will generate the appropriate web-forms (under MAMEM’s web-site) to describe and link the datasets. All technical details for accessing and processing the data will be also included as part of these web-forms. Subsequently, we will seek opportunities to list our datasets in sites that serve as aggregators of bio-signal datasets, such as [3] and [4]. This will boost the visibility of the generated datasets and ensure their widespread use.

4.4.3 **Free sharing**

This approach consists in sharing the dataset through the projects web-site or other sites and allowing externals to freely download them, without requesting to provide their personal details. In this case, we will include the types of data belonging to the “open” category. Data such as deliverables and research papers will be included in this case.

4.5 Archiving and preservation

As already mentioned the datasets (in their raw form) will reside on the institutions acting as data controllers. Thus, in terms of archiving and preservation we will rely on the services offered by major institutions like universities and medical centres. These institutions typically offer services like redundancy, back-up and migration that are consider sufficient for the preservation of the generated data. In addition, these institutions usually maintain a large digital library that is used to index and archive the content generated by their activities. Finally, it is important to mention that the services offered by these institutions are free of charge and may sustain for many year after the end of the project.

5 Ethical and confidentiality considerations

MAMEM will give specific attention to any ethical issues that will arise and will address them in a professional way following very closely established EU regulations and corresponding national laws about user privacy, confidentiality and consent. The adopted ethical practices are described in the DoA (Section 5) and related to: a) the criteria used for the inclusion of participants in the project's clinical trials, b) the inform consents that should be signed by the participants, c) the potential harms and benefits, and d) the conventions, declarations, directives and regulations that are applicable in our case. All the above have been sufficiently detailed in the DoA [5] and will be taken into consideration in executing our data management plan.

6 Conclusions

In this deliverable, we have presented the data that will be generated in the context of MAMEM including data from questionnaires (clinical information, needs, computer use, etc.), raw signals (EEG, eye-tracker, GSR sensors) and interaction logs. Our data management plan is built upon analysing the generated data with respect to their confidentiality level and employing a different sharing approach depending on this level. More specifically, three confidentiality levels are envisaged (i.e. open, protected and confidential) and two sharing practices are described (i.e. private and controlled sharing). In this respect, confidential data and data rated as not shareable according to ethical considerations will not be shared. Finally, the DMP strictly builds on ensuring the necessary informed consents, as well as respecting the sphere of privacy of each participant. This document will be further developed, as we gain more information about the specific requirements of our subjects and the usage scenarios implemented by the MAMEM system.

7 References

[1] Guidelines on Data Management in Horizon 2020: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[2] D6.1 - Clinical requirements for the MAMEM platform for each of the patient cohort, MAMEM Consortium, August 2015. url: http://www.mamem.eu/wp-content/uploads/2015/09/D6.1_Clinical_requirements_Final.pdf

[3] http://bnci-horizon-2020.eu/database/data-sets

[4] http://www2.hu-berlin.de/eyetracking-eeg/testdata.html

[5] Description of Actions - Readable form (requires authentication): http://mklab.iti.gr/mamem/images/3/32/PartB_MAMEM.pdf

[6] Alois Schlögl, "GDF - A general data format for BIOSIGNALS". CoRR abs/cs/0608052 (2006)

[7] "BioSig". sourceforge.net. Retrieved 3 August 2015.

[8] "libGDF download - SourceForge.net". Sourceforge.net. Retrieved 3 August 2015.

[9] "Specifications - xdf - This page contains the format specifications of XDF 1.0 (currently in beta). - XDF (Extensible Data Format) - Google Project Hosting". Retrieved 3 August 2015.

[10] MATLAB and Statistics Toolbox, The MathWorks, Inc., Natick, Massachusetts, United States

A Appendix

A.1. Informed consent for capturing EEG signals

Information Note for Participants and Consent Form

Centre for Research and Technology Hellas - ITI - CERTH
E.K.E.T.A.

Research Number: H2020-ICT

Research Title: Multimedia Authoring and Management using your Eyes and Mind

Research Centre: Centre for Research and Technology Hellas - ITI - CERTH

Please read carefully the following document

You are invited to participate in a research study that employs Electroencephalography as a means of recording brain activity. Before you make the decision to participate in the research, it is imperative that you understand the purpose of the study and what it includes. Should you need any further information, please do not hesitate to contact us.

Purpose of this study

Multimedia **A**uthoring and **M**anagement using your **E**yes and **M**ind (MAMEM) main objective is to enhance the research for the use and connectivity of the brain for individuals with limited or no moving ability. Recording the brain activity while presenting a visual stimulus aims to examining the physical response of the brain.

MAMEM's main objective is to assess the impact of this technology in making these people more socially integrated.

Describing the experimental procedure

During the experiment a colored box is presented on the screen, flickering in 5 different frequencies for 5 seconds followed by 5 seconds without visual stimulation. The process includes 5 sessions and between every session there is a resting time.

Each session starts with 100 seconds and during this the participant looks at the monitor, without being involved in any activity, followed by approximate one minute during which

all five frequencies are presented in a random way (adaptation) and the first 30-second resting period follows. Adaptation is followed by the presentation of the first frequency for 3 times and with a 30-second break the second frequency is presented. In a similar way the remaining 3 frequencies are presented. After the first session is completed, there is a resting period of 10 minutes and then the next session follows. The experiment ends when a total number of 5 sessions is completed.

Potential hazards or inconvenience during the experiment

None. Electroencephalography is a non-invasive method that records electrical activity of the brain. Radiation is not emitted during the experiment and no paramagnetic substance, radioisotopes or intravenous contrast are used.

The only inconvenience you may have is in the movement limitation during both the preparation and the experimental procedures.

Bare in mind

Before the experiment

Your responsibility is limited to timely inform the research staff in cases of clinical history of epilepsy and avoid doing the following for at least 2 hours before the experiment takes place:

- Drinking coffee, tea or coca cola
- Taking medications that affect the Central Neural System
- Having a large meal
- Using make-up

During the Experiment

During the presentation of the visual stimuli, you are strongly encouraged to avoid the following:

- Moving in any way
- Blinking
- Swallowing
- Touching the electrodes

During the experiment there will be sufficient resting time (i.e. the screen is black) and any of the previously described activities can be performed during these time lapses.

Consent for data sharing for research purposes

By signing this document you provide your permission to the members of the research team to have access to your data. Furthermore, you consent to making your data public for research purposes (i.e. non-commercial use). Access to the data will be provided through the dissemination channels of MAMEM. All identification information of the participant will be removed before making them publicly available.

Consent Form for Research Participation

This consent form includes significant information and should aid you deciding whether to participate or not in the research study. If you still have any questions, please feel free to address to the research stuff, before signing this form.

You have read all the provided information that is presented to you in a language that you can read and comprehend. There has been satisfactory explanation about this research and all your questions about it have been answered. Based on this information you volunteer to participate in this research and consent to make your data public for research purposes.

_____

Participant Name

_____

Participant Signature Date (dd-mm-yyyy)

_____

Researcher Name

_____

Researcher Signature Date (dd-mm-yyyy)

A.2. Template for Dataset "Terms of Use"

We will provide restricted access to the MAMEM Datasets to individuals who agree to the following Terms of Use, which allows use for non-commercial purposes, including teaching, academic research, public demonstrations, personal experimentation, and research publications.

In order to provide you this dataset you have to follow the following steps:

1. Read the datasets Terms of Use below which oblige you, among other things, to:
    a. not share this data with anyone else
    b. not attempt to identify any of the individuals
2. Choose which datasets you are going to need in your research.
3. Provide us with a short description of your planned research/project (optional)
4. Fill-in a form with your name, your institutional e-mail and your affiliation.

Datasets Terms of Use

This is an agreement between you and MAMEM project.

These Terms of Use ("TOU") are an agreement between you and MAMEM project ("MAMEM" or "we"). This TOU governs your use of any of the MAMEM Datasets ("Dataset"). The Dataset, which is provided by MAMEM, enables you to obtain certain data from MAMEM experiments to use in connection with your research or applications, as described below.

You represent that you are at least 18 years old. Your use of the Dataset constitutes your acceptance of this TOU, the Privacy Statement and registration guidelines, without modification.

MAMEM may update the TOU by posting new terms and conditions on the site from which you downloaded the Dataset. If you continue using the Dataset, your use is governed by the new terms and conditions.

Please note that we do not provide warranties for Dataset. These terms are in the "Disclaimer of Warranties" section below, and we ask you to read it carefully.

Confidentiality & usage

You cannot send the dataset to any other party (even if they have access to it themselves), nor disclose to anyone else the information contained within it as well as its structure. It is allowed to share data with students in your research group whom you agree to supervise and take full responsibility that their use of the data also meets these TOU.

You may not distribute or publish this Dataset. You may not use this Dataset in any form for commercial purposes. You may, however, use the Dataset to create tools or visualizations that are independently developed by you for the purpose of analyzing or visualizing the Dataset. Examples of commercial purposes would be running business operations, licensing, leasing, or selling content obtained from the Dataset, distributing the Dataset for use with

commercial products, using the Dataset in the creation or use of commercial products or any other activity which purpose is to procure a commercial gain to you or others.

You, and any third party working on behalf of you, will not engage in any unlawful practices in connection with your use of the Dataset.

Anonymisation

You, and any third party working on behalf of you, will not use any information you obtain through the Dataset to try to identify, contact, advertise to or otherwise target any individual;

Non-Commercial License

We grant you a non-commercial license to use the data. You can only use it for academic research that does not earn revenue, and your research also cannot be in collaboration with any commercial entities.

Scope of Research

If the scope of your research changes, then you should contact us for us to agree with the change.

Publication acknowledgement

All uses of this dataset including but not limited to research papers, at conferences, on websites and in press releases, should include a reference to <INSERT HERE THE TECH REPORT OR THE PUBLICATION DESCRIBING THE DATA>. In addition, you should include a reference to MAMEM project as the data source. However, it should not imply that MAMEM endorses the research. It should be clear that MAMEM is an external data supplier

Access to data

MAMEM cannot guarantee continued access to the data, nor that will our service not be interrupted from time to time. The license to use and store our data is recoverable, which means that MAMEM may ask you to cease use of it and to delete it from any storage you have at our sole discretion.

Disclaimer of warranties

You use the Dataset and any related products at your own risk. The Dataset and any related products are provided “AS IS” and with all defects. Except as provided herein, Microsoft expressly disclaims all other express, implied, or statutory warranties. This includes the warranties of merchantability, fitness for a particular purpose, title, non-infringement, lack

of viruses, quiet enjoyment, scope of license, lack of errors, and satisfactory condition or quality.

Termination

1. Termination by you. You may stop using the Dataset if you are dissatisfied with any aspect of the Dataset.
2. Automatic termination. If you breach these TOU or if you sue anyone over patents that you think may apply to or read on the Dataset or anyone's use of the Dataset, this Agreement (and your license and rights obtained herein) terminates automatically.

Feedback

If you provide any feedback to MAMEM regarding the Dataset, you agree that such feedback is voluntarily given, and MAMEM shall be free to use the feedback as it sees fit without obligation or restriction of any kind.

Reservation rights

MAMEM reserves all rights in the Dataset and associated services and technologies, subject to the licenses granted in this Agreement. Nothing in this TOU grants any additional rights (including any implied licenses, covenants, releases or other rights) whether by implication, estoppel, or otherwise.

Copyright notice; Trademarks