MeMAD

Methods for Managing Audiovisual Data

memad.eu
info@memad.eu

Twitter - @memadproject
Linkedin - MeMAD Project

MeMAD Deliverable

D1.6 Data Management Plan, update 2

Version 1.0

Grant Agreement number	780069
Action Acronym	MeMAD
Action Title	Methods for Managing Audiovisual Data: Combining Automatic Efficiency with Human Accuracy
Funding Scheme	H2020-ICT-2016-2017/H2020-ICT-2017-1
Version date of the Annex I against which the assessment will be made	23.6.2020
Start date of the project	1.1.2018
Due date of the deliverable	30.3.2021
Actual date of submission	30.3.2021
Lead beneficiary for the deliverable	Yle
Dissemination level of the deliverable	Public

Action coordinator's scientific representative
Prof. Mikko Kurimo
AALTO – KORKEAKOULUSÄÄTIÖ, Aalto University School of Electrical Engineering, Department of Signal Processing and Acoustics
mikko.kurimo@aalto.fi

MeMAD project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 780069. This document has been produced by theMeMAD project. The content in this document represents the views of the authors, and the European Commission has no liability in respect of the content.

Authors in alphabetical order		
Name	Beneficiary	e-mail
Sebastian Andersson	Lingsoft	sebastian.andersson@lingsoft.fi
Sabine Braun	SURREY	s.braun@surrey.ac.uk
Jean Carrive	INA	jcarrive@ina.fr
Maija Hirvonen	UH	maija.hirvonen@helsinki.fi
Harri Kiiskinen	YLE	harri.kiiskinen@yle.fi
Tiina Lindh-Knuutila	LLS	tiina.lindh-knuutila@lingsoft.fi
Mikko Kurimo	AALTO	mikko.kurimo@aalto.fi
Jorma Laaksonen	AALTO	jorma.laaksonen@aalto.fi
Lauri Saarikoski	YLE	lauri.saarikoski@yle.fi
Liisa Tiittula	UH	liisa.tiittula@helsinki.fi
Tiina Tuominen	YLE	tiina.tuominen@yle.fi
Raphaël Troncy	EURECOM	raphael.troncy@eurecom.fr
Dieter Van Rijsselbergen	Limecraft	dieter.vanrijsselbergen@limecraft.com

Document reviewers		
Name	Beneficiary	e-mail
Jorma Laaksonen	AALTO	jorma.laaksonen@aalto.fi
Jörg Tiedemann	UH	jorg.tiedemann@helsinki.fi

Document revisions			
Version	Date	Authors	Changes

Abstract

This document describes the data management processes in the project Methods for Managing Audiovisual Data: Combining Automatic Efficiency with Human Accuracy (MeMAD).

This is the third and final, updated version of the data management plan. Practices defined during the project have been described, and the document has been extended to provide an overview to the data curation of the project result. This work is described in more detail in the project deliverable D1.7.

Contents

1 Introduction

This document, final of three iterations, describes the MeMAD project's data management and best practices as they have evolved during the project. Development of this plan has been an iterative process which has continued throughout the project.

The main changes compared to the previous version of the Data Management Plan are the following:

- Updates reflecting the internal data management of the project, reported in detail in the deliverable D1.5.
- Updates and additions covering the life cycle, storage and licences of project results and post-project data, reported in detail in deliverable D1.7.

Parallel to this document the project has produced also another data-related deliverable, D1.7 *Collection of annotations and/or video data resulting from the project*, which describes the handling of project results data and possible datasets licensed by the project for post-MeMAD use. This document describes the main principles of the project data management, and most of the details on the storage and licensing of project results have been described in deliverable D1.7.

2 Data Summary

The purpose of data collection and generation within the project has been to facilitate the development and evaluation of methods for multimodal analysis of audiovisual (AV) content. A very large part of the data that has been gathered and used by the project is either already publicly available research data or strictly licensed audiovisual data from industry partners.

The main data resources produced by the project is in the form of computer program code and algorithms, trained machine learning models, metadata for media produced by the machine learning (ML) systems, and processed AV content. In addition, interview, observation and test data has been collected in user studies and experiments.

The research and evaluation data the project has used is in three main formats:

a) audiovisual digital data

b) general metadata, subtitles and captioning aligned to audiovisual content

c) specific metadata describing the content of the audiovisual material

The project has created data of the following type:

a) annotated datasets of audiovisual data

b) program code and algorithms

c) trained models using neural networks and machine learning algorithms

d) survey, interview, observation and test data

e) dissemination data such as publications, demonstration videos and web content

In addition, there have been intermediate data types used within the project that are not necessarily preserved:

a) AV content processed to a format more suitable for further analysis (resampling, transcoding, etc.)

b) intermediate data types for metadata and AV aligned data (subtitles, content descriptions, etc.)

c) datasets resulting from program code development

d) user experience data relevant only for intermediate purposes

MeMAD has mostly used previously created audiovisual content for research purposes. For testing and development purposes, project partners and external partners have provided additional audiovisual content. Several freely available datasets have been used by the different work packages for their own specific needs. Industry partners within the project have provided datasets consisting of their own media for the project. In addition to the data managed by MeMAD, external partners have also used their own data in testing the systems and methods developed by the MeMAD consortium. A detailed description of the research datasets has been provided in deliverable D1.2, and the deliverables from work packages WP6 and WP7 describe in detail the data used in project prototype evaluations and proof of concept activities with external parties.

The research and evaluation data has been obtained from two major sources:

a) research data that can be collected from public sources and open data providers

b) published or archived media from industry partners and external partners

State of the art research datasets have been obtained and compiled by each partner and work package individually, according to their own research needs. Additional datasets have been provided by project partners, mainly Yle and INA. Deliverable D1.2 *Collection of Annotated Video Data* describes these datasets in detail. Within the project, a summary of the datasets has been kept centrally, and the partners/work packages (WP) have been invited to mark the datasets they have been using. The total size of the research and evaluation datasets within the MeMAD project has been large: the largest research oriented datasets in MeMAD have been tens of terabytes in size.

These datasets would be of immense use for any other parties working with automatic analyses of audiovisual data, including general AI research, media studies, translation studies etc, as well as commercial parties developing methods for e.g. media content management. Possibilities to re-use this data after the project are reported in detail in deliverable D1.7.

3 FAIR Data

Data management in MeMAD has been guided by the set of principles labelled FAIR[1]. The purpose of these principles is to make data Findable, Accessible, Interoperable and Reusable. Characteristic for MeMAD data management has been that while most of the project outputs have been made openly accessible, access to many of the media collections used during the project has been restricted because of the business and IPR needs of the media producers and rights holders.

3.1 Making Data Findable, Including Provisions for Metadata

Requirements of the common prototypes have provided the framework within which the project data has been managed, and deliverables D6.1, D6.4 and D6.7 (*Specifications of Data Interchange Format*, versions 1 to 3) provide guidelines on how to document the data. Respectively, project collaboration and data exchange have provided guidelines for internal project data management, described in more detail in deliverables D1.3 and D1.5 (*Data Collection and Distribution Platform*, initial and final versions).

The project has extensively used the prototype platform developed in the project work package WP6 also for the project's internal data management. The platform provides possibilities to annotate the uploaded data and it also logs the media uploads and workflows run on these media items, making it easy to find and track different media items and their versions. The platform also provides user control and data security features needed by the project.

One part of the research data produced by the project has been a knowledge graph representing in RDF the legacy metadata associated to the audiovisual programs as well as some of the automatic analysis results produced by the project. This knowledge graph follows the Linked Data principles, which means that every object is identified by a dereferencable URI. The project has established a policy to mint those URIs following some existing best practices from the (Semantic Web) community. First, the MeMAD ontology has for namespace URI <http://data.memad.eu/ontology#> with the recommended prefix to be "memad". Second, the general pattern for identifying metadata object is http://data.memad.eu/[source|channel]/[collection|timeslot|series]/[UUID] where:

- source | channel (in lower case)
    - channel codes for INA: ['fcr', 'fif', 'fit', 'f24', 'fr2', 'fr5']
    - channel codes for Yle: ['tvfinland', 'yle24', 'yleareena', 'yletv1', 'yletv2', 'yleteema', 'ylefem', 'yleteemafem']
    - 'surrey' for the material used by the University of Surrey
- collection | timeslot | series (in lower case and in ASCII and slugified)
    - we replace: white space (' '), semicolon (':'), comma (','), slash ('/'), quote ('"'), brackets ('(' or ')' or '[' or ']'), exclamation mark ('!'), interrogative mark ('?') and hash sign ('#') by a hyphen '-'.

[1] https://www.go-fair.org/fair-principles/

    - we delete consecutive hyphens to only have one, at most; we do not end with a hyphen; we do not start with a hyphen.
- UUID = MeMAD custom hashing function using a seed where:
    - seed for INA is “record ID” (of a program OR a subject)
    - seed for Yle is “guid” OR “contentID”

Finally, media objects have been identified using the pattern http://data.memad.eu/media/[UUID]

Each project WP has used their own naming schemes according to internal conventions of the research groups, typically following systematic structure that states e.g. data origin, version numbers etc. The aim has been to make individual files findable and identifiable even when no additional metadata has been provided.

Parts of the project data have been stored into open repositories and for the licence-restricted datasets, metadata entries have been created into relevant data catalogues such as META-SHARE[2], which improves their findability.

All code produced by the project has been stored on GitHub (https://github.com/memad-project), while most other data has been stored in the MeMAD project community on Zenodo (https://zenodo.org/communities/memad/). The data stored on Zenodo includes models, annotations, research data from user evaluations, all public deliverables created by the project, selected materials from dissemination events, and a comprehensive list of publications.

Publications themselves have been stored in institutional or other repositories, and the list of publications on Zenodo contains links to these for ease of access. Items that are related to each other contain two-way references between and within repositories, included in the metadata of these items.

Confidential deliverables and confidential research data, such as recordings of interviews and focus groups, have not been stored publicly. Furthermore, intermediate versions of code have not been stored, only the final versions or versions that were current and stable at the end of the project. The project has tracked all data that is to be destroyed to ensure data security and confidentiality. The project website has not been archived by the project consortium, but it will remain available until November 2024. A version of the MeMAD website, archived by the Wayback Machine, can be found at https://web.archive.org/web/*/www.memad.eu .

The platforms described above support the findability of the project data, but they also contribute to making the data openly accessible. The next section describes in more detail how the project data has been distributed.

[2] https://metashare.csc.fi/repository/search/?q=memad

3.2 Making Data Openly Accessible

The data used and produced within the project can be divided into five groups according to differences in licensing and reusability:

1. research-oriented data obtained from public repositories
2. research and evaluation data obtained from project industry partners
3. annotated media data produced from groups 1 and 2 during the project
4. algorithms and program code produced by academic research groups
5. proprietary technologies developed by project industry partners.

Of these data types, the data in groups 3 “annotated media” and 4 “algorithms and program code produced by academic research groups” is the easiest to open for public re-use and has been made available as widely as possible.

Data in group 1 “research-oriented data obtained from public repositories” often comes with a licence that does not allow re-distribution even though use for research purposes is free. However, since this data typically is already openly available for research purposes, there has been no evident need to re-distribute additional copies of it by the MeMAD project..

Data in group 2 “research and evaluation data obtained from project industry partners” is typically published media data which has strict licences concerning re-use and distribution, for example, tv-shows produced by broadcasting companies. This group also includes the user data collected during prototype testing. An open access publication of this kind of media does pose legal challenges and may require investments in licensing fees. In the context of MeMAD, the aim has not been to make this data set publicly available to parties outside the project - the focus has been on enabling at least the project’s internal use. However, subsets of this data have been selected based on their estimated potential value, and the project has worked on securing a wider access to these subsets and clarifying licences for their use. This work has been described in deliverable D1.7.

Data in group 5 “proprietary technologies developed by project industry partners” concerns tools and methods that the industry partners have contributed to the research project in order to facilitate and evaluate certain phases of the research. They reflect a considerable economic investment on the part of the industry partners, and have been aimed at developing further technologies and solutions with commercial purposes, thus not suitable for open distribution.

Research Partners and Their Data and Source Code

The MeMAD project has strived to publish all its research in as open a way as possible. This principle has applied to data and source code produced by the research partners within the project. The source codes have been made available in GitHub and the created models and scientific publications in Zenodo.

Industry and Commercial Partners and Their Data

Commercial partners in the MeMAD project have shared the output from automatic analyses generated during the MeMAD project if sharing them has not been prohibited due to business needs or the copyright restrictions of the original media they have been based on.

Concerning the data of group 5, most of the technologies Limecraft, Lingsoft and LLS have developed as part of MeMAD have not been made openly available by default, but mentioned parties have worked to provide access and visibility to these technologies also - if not as source code, at least as services . Lingsoft's APIs for speech recognition in Finnish and Swedish as well as the NER in Finnish and Swedish will be made available through the European Language Grid platform (ELG). On the other hand, Limecraft, Lingsoft and LLS have evaluated the open distribution of components developed during the project if those are parts that form an extension to a sizable, existing open source component, or in cases where the open distribution of a component makes sense economically, e.g., to enforce the commercial ecosystem that Limecraft, Lingsoft and LLS intend to build around MeMAD technologies. At the same time, the industry partners have contributed to group 4 open data initiatives initiated by the academic partners. Contributions for improvements, bug fixes and new features were made to open algorithms and program code whenever this contribution was in the spirit of the original data repository.

Yle Dataset

Yle has provided the project with a selection of AV material and related metadata from its broadcasting media archives. The AV material, altogether ca. 500 hours, consists of in-house produced TV programs. The rights of Yle are limited to typical business use, such as broadcasting, and specifically do not include open distribution. A licence agreement with the national copyright societies was established early in the project, allowing Yle archive material to be used freely within the MeMAD project and also the distribution of the material to researchers for project purposes. Subsets of this dataset have been selected as potentially valuable for post-MeMAD use as described in deliverable D1.7.

The selection of program metadata includes the times of transmission, content descriptions, classifications and the producing personnel for the TV programs. This data is not limited by copyright, but as the data has originated from in-house production processes for a specific use, opening it may be limited by issues related to e.g. personal or journalistic data. Yle metadata has been included in the open access project results, if no limitations to do this have been identified.

INA Dataset

Since 1995, INA has been the legal depository of French television and radio. Legal deposit is an exception to copyright and INA has no intellectual property rights over the content deposited. The cataloging data (title, broadcast date, producer, header, etc.) are accessible for free, in accordance with the rules in force, by a search engine located on the site http://inatheque.ina.fr. INA also markets a collection, mainly made of content produced by public television and radio stations, for which INA holds the production rights. INA thus

offers broadcasters and producers excerpts and full programs, and pays back a contribution to the rights holders.

To promote research, INA provides for strictly research purposes (academic or commercial), various collections available on accreditation through the INA Dataset web site (http://dataset.ina.fr). INA has provided to MeMAD's partners, in relation to the conditions of use described on the INA Dataset web site, a specific corpus of television and radio programs related to the European elections in 2014.

INA also offers an open data collection of metadata on the thematic classification of the reports broadcast on the evening news of six channels (TF1, France 2, France 3, Canal +, Arte, M6) for the period January 2005 -June 2015, available at https://www.data.gouv.fr/fr/organizations/institut-national-de-laudiovisuel/.

Accessible storage of Project Results

While the primary data from the AV sets is not openly accessible, the project has created metadata entries of these datasets into CLARIN and META-SHARE, accompanied with contact information needed for licensing and accessing these datasets.

During the project, research data created has been stored to the project's internal platforms and selection of this data has been included in the project result dataset as deliverable D1.7. Zenodo has been used as the final depository for research data, as it provides a good variety of both human and machine readable methods for discovering and accessing the data, with no additional costs to the project.

The program data ("code") has been stored as a GitHub repository, and can be accessed thus by both via the www-interface to the repository as well as with Git directly. Documentation for the Git system is available on the internet for free, and the use of the program is discussed on several open forums worldwide. Program code used to analyse and process the datasets that is based on algorithms and techniques discussed and presented in scientific publications, is open source by default, and the released datasets contain information on the relevant program code for their use. However, in the case of products intended for commercialization by the project industry partners, the release of the program code has not been possible by default.

Research and evaluation data has been distributed via suitable tools. As most of the previously prepared research datasets have been available either as open access or via specific agreements, the partners using them have acquired the data directly from the providers. Regarding research data from MeMAD project industry partners (Yle, INA), the partners have their own systems for distributing large datasets. INA data has been available on the INA ftp server, and the Yle data has been distributed via a high-speed file transfer service suitable for distributing large datasets.

Such project result datasets that contain neither licenced nor sensitive information have been made by default open for access to all interested parties. This does not apply to the proprietary media or data provided by project industry partners.

No need for a specific data access committee within the project has been recognized. When needed, this work has been covered by the ethics committees of respective project partners. The research data provided, while under a restrictive research licence, has contained neither sensitive information on persons, nor institutions. The user data collected during the project is sensitive by nature, and person-related details have not been quoted or published. The data has been used only for research purposes, and recordings in which persons may be identified have not been shown in public and will by default be destroyed when the project work has been finished.

3.3 Making Data Interoperable

One of the project goals has been to create a set of interoperable research and evaluation data. The following have been selected as the interoperable data formats:

Data Type	Data Format	Explanation
video	video/mp4[3]	TV programs and other video data
subtitles	SubRip Text[4] Advanced SubStation Alpha[5]	subtitles/captions for videos
video annotations	ELAN Annotation Format[6]	segmentation and textual annotation for videos
various metadata structured according to well-defined ontologies	text/turtle[7]	ontology encoded in OWL/RDF, including many interchange formats also discussed in D6.7.
knowledge graph	text/turtle	RDF triples and named graphs, following a number of well-known ontologies such as EBU Core, NIF, Web Annotations, etc.
raw media analysis results	text/csv[8] application/json[9]	media content annotations
structured data	application/xml[10]	multiple uses

Table 1. MeMAD's interoperable data formats.

[3] https://www.iso.org/
[4] https://www.matroska.org/technical/subtitles.html#srt-subtitles
[5] http://moodub.free.fr/video/ass-specs.doc
[6] https://www.mpi.nl/tools/elan/EAF_Annotation_Format_3.0_and_ELAN.pdf
[7] https://www.w3.org/TR/turtle/
[8] https://tools.ietf.org/html/rfc4180
[9] https://tools.ietf.org/html/rfc8259
[10] https://www.w3.org/TR/xml/

In general, known best practices have been followed. As much as possible of the produced and used data has been stored in formats that are well known and preferably open; structured text formats have been preferred when suitable.

Project deliverables (D6.1, D6.4, D6.7), describing the project prototype specifications, include definitions for data exchange within the prototype system. These deliverables concern mostly the interoperability of the prototypes and frameworks within the project in a demonstrator (and later also production-ready) context. Hence, these formats return here to some extent (e.g., for the exchange of ontology-based data), but many other formats were used during the research and development phase of the project that was often supported by less elaborate or standardised formats.

The project result datasets have been described using well known ontologies including EBU Core, DCMI, and Web Annotations. In the case it has been necessary to create project-specific vocabularies/ontologies, mappings to commonly used ontologies have been provided when judged useful for the project or dissemination purposes.

3.4 Increase Data Re-use

Data collected specifically for the project by its industry partners as proprietary datasets have been strictly licensed, and in many cases the partners do not hold all the rights for the data or media. Therefore, it is highly difficult to license these datasets for open ended further use, especially under any kind of an open access licence. Copyright societies granting licences typically wish to limit the duration and scope of licences in unambiguous terms, which does not favour open ended licences that would be optimal for data re-use. The MeMAD project approach has been to acquire licences which are as open as possible, and include in the agreement negotiations mechanisms for other parties to licence the same dataset for similar purposes in the future.

Project deliverable D1.7 describes the work done on clarifying licences for this data. As a summary of that work, licences for research use are easier and cheaper to secure than those permitting commercial use. Also, since the European legislation on data mining and copyright is undergoing major changes at the time of writing, interested parties have been slightly cautious about making long term licence agreements at this time.

Licensing challenges affect mainly the primary data – video, audio and most ancillary data such as subtitles – but parts of this data, e.g. neutral metadata elements, could and should be shared. Secondary data, such as annotations created during the project, are more straightforward to share for re-using. The project has shared these as a part of the project results data described in deliverable D1.7.

Data produced by the project itself can and should be open for re-use in accordance with the commercialization interests of the project industry partners; this work has taken place during the project and will continue after the end of the project. Additional data curation and refinement has been performed as necessary for the purpose of the project and the resulting collections have been published by the individual partners according to the principles of open research if possible.

4 Allocation of Resources

Since research data has been made FAIR partially as part of other project work, exact total costs have been hard to calculate. Many of the datasets used already carry rich metadata, are already searchable and indexed, are accessible and presented in broadly applicable means and forms, are associated with their provenance and meet domain-relevant community standards.

Explicit costs for increasing the FAIRness of the data have been related to, as a minimum, acquiring licences for proprietary datasets in the form of licence fees, but also in these cases part of the costs come from work associated with drafting licence agreements and promoting FAIR principles among data and media rights holders and their representatives.

Direct licence fee costs have been covered from WP1 budget. Work hours dedicated to licence negotiations and data preparation have been covered from each partner's personnel budget respectively, as they have allocated work months to WP1 each.

Each consortium partner has appointed a data contact person, and the overall responsibilities concerning data management have been organized through work done in WP1, dedicated to data topics.

For long-term preservation of open research data the project has chosen Zenodo as a platform, and since it is free to use, no resources need to be allocated to cover these costs from the project.

5 Data Security

Each project partner has their policies and means to keep the data safe on their sides with secure methods of storing and transferring the data and access control on shared data.

Project internal data platforms have been provided by INA and Limecraft, and this work follows their security policies. This has been described in more detail in the project deliverables D1.3 and D1.5.

The project prototype uses Limecraft Flow[11] as a platform. Limecraft follows the guidelines from ISO/IEC 27001[12] for best practices in securing data. Limecraft is also a participant in the UK Digital Production Partnership and its "Committed to Security Programme"[13].

- Data stored as part of the Limecraft Flow infrastructure is hosted in data centers within the EU, and all conform to the ISO/IEC 27001 standard for data security. In addition to infrastructure security provided by Limecraft's data center partners (physical access controls, network access limitations), Limecraft's application

[11] https://www.limecraft.com/
[12] https://www.iso.org/
[13] https://www.digitalproductionpartnership.co.uk/what-we-do/committed-to-security-programme/

platform also enforces internal firewalling and is only accessible for administration using dedicated per-environment SSH keys.
- Any exchange of data is subject to user authentication and subsequent authorization (either from Limecraft employees, which requires special access rights, or from clients whose access is strictly confined to the data from their own organisations). Additionally, any exchanges occur exclusively over encrypted data connections.

Long term preservation of the data that has been opened for further use has been centralized into the Zenodo platform and relies on the best data security practices of this service[14]. Media datasets provided by Yle and INA are parts of their archive collections, and will be preserved and curated through their core business of media archiving.

6 Ethical Aspects

The Project has followed the guidelines for responsible conduct of research[15]. Project research groups have sought and received ethical approval from respective research organisations.

Part of the research data may contain personal information and it has been handled following guidelines and regulation such as GDPR. For the relevant project tasks such as WP6 evaluation runs, a Data Contact has been nominated and a contact point on personal data related issues has been set up to answer queries and requests for personal data related issues.

Metadata provided by industry partners may have issues related to the journalistic nature of the original datasets. Some of these datasets, such as the metadata provided by Yle, have been designed and intended for in-house production use of a broadcaster, and opening this data to outside users may result in needs to protect sensitive or confidential information stored within the data. These issues have been resolved by removing and/or overwriting sensitive and confidential information in the research data set before delivering it to the project.

The user data (interview, observation and test data) collected during the project from experiments, user studies and authentic workplace interactions between human beings are sensitive data and have been protected and handled with proper care and measures (see MeMAD DoA, Chapter 5).

Interviews and user experience studies conducted in connection with the MeMAD prototypes may contain aspects which describe internal processes at the industry partners' organisations or sensitive personal information about the interviewees. Disclosing information that has commercial interest or sensitive personal information may preclude these datasets from open distribution. The confidentiality of such data is determined by the

[14] https://about.zenodo.org/infrastructure/
[15] See for example the Finnish National Board on Research Integrity https://www.tenk.fi/en and University of Surrey Code on Good Research Practice https://www.surrey.ac.uk/sites/default/files/2018-05/code-on-good-research-practice.pdf

informed consent statements provided for the study participants and the data use principles described in them.

The video recordings from interviews, focus groups and other research sessions have been transcribed and anonymised, and all potentially identifying information have been removed from publications. The video recordings and all personal information about study participants have been destroyed at the end of the project. Furthermore, the anonymised transcripts have only been shared with other project participants. Researchers who have participated in the project are permitted to use the data in publications even after the end of the project. The transcripts have been stored anonymously and securely, and they will not be made public at any time.