Horizon 2020 Programme
DG CNECT
Next-generation Internet

Project acronym: **DECTS**

Project title: **Deaf Emergency Chat and Training System**

Third Party: **OwnYourData**

Deliverable 2.3

Data Management Plan

Deliverables leader:	**OwnYourData**
Authors:	Christoph Fabianek
Due date:	2020-10-15
Actual submission date:	2020-07-16
Dissemination level:	Public

Abstract: This report is part of a third-party project DECTS that has received funding from the NGI_Trust, the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 825618.

Document Revision History

Date	Version	Author/Editor/Contributor	Summary of main changes / Status
2020-03-24	0.1	Christoph Fabianek	Initial document
2020-05-15	0.9	Christoph Fabianek	Version ready for review
2020-05-31	1.0	Christoph Fabianek	First release
2020-07-16	1.1	Christoph Fabianek, Eduard Gringinger	Second release and final release

Disclaimer

The sole responsibility for the content of this publication lies with the authors. It does not necessarily reflect the opinion of the European Commission. The European Commission is not responsible for any use that may be made of the information contained therein.

Copyright

This document may change without notice.

NGI_Trust project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825618.

Table of content

NGI_Trust project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825618.

Executive Summary

The deliverable provides a Data Management Plan (DMP) for the DECTS project. This document is the first version of the DMP. It is a living document and will be updated whenever relevant changes occur to the project.

In the DECTS project we implement and evaluate an end-to-end workflow for deaf and hard of hearing persons for consent management, secure and privacy preserving personal data provisioning in case of an emergency and rolling out the existing Austrian solution at the European level.

The current DMP version describes two datasets with personal identifiable information that are used in the course of the project.

1 Introduction

This version of the Data Management Plan (DMP) is the second iteration released in project month three and is based on the DMP from the Data Market Austria (https://datamarket.at, public delivery 1.3) project that funded the original development of Semantic Containers. The document has been created by OwnYourData, the project partner in charge of the project data management task, in consultation with the other project partner MeeCode by Mario Murent. The DMP describes the data that the project collects and generates, how it will be exploited and how it will be curated and preserved.

The consortium partner OwnYourData is responsible for implementing the Data Management Plan and ensures that it is reviewed and revised during the project runtime. New versions of the DMP will be created whenever important changes to the project occur due to inclusion of new datasets, changes in consortium policies or external factors. At the project end, a final version of the DMP will be released.

2 Data Summary

2.1 Data types and origin

The following data types are included in the Data Vault:

- Owner information: account information of a user
- Emergency data: additional information provided in the course of an emergency chat
- Any other data added through installed plugins

The following data types are included in a Semantic Container (used as Chat Bot in the Deaf Emergency Training System):

- Owner information (i.e., operator of the Semantic Container)
- Semantic description of the data in the container
- Additional Metadata about the container (incl. processing capabilities)
- Payload (actual data stored in the container, i.e., chat protocols)
- Logging information about accessing the Semantic Container

The current version of the DMP reflects processing of the following datasets:

- Emergency data in the Data Vault
- Chat Protocol in the Semantic Container

3 FAIR Data

3.1 Making data findable, including provisions for metadata

Semantic Container provides a platform for sharing commercial and non-commercial data. To make data discoverable, metadata is provided through API endpoints. A detailed and up-to-data description of the API service endpoints is available here:
https://api-docs.ownyourdata.eu/semcon/

A detailed description of the metadata attributes can be found in the current version of the Semantic Container Design Document available here:
https://www.ownyourdata.eu/semcon/design

The OwnYourData Data Vault also provides API endpoints documented here:
https://api-docs.ownyourdata.eu/datavault/

A complete and immutable audit trail for data access is available within the OwnYourData Data Vault in the User menu > Access Log. (Each create / read / update / delete operation is linked through the hash value of the previous operation and the complete record is stored in the Ethereum Blockchain. With this approach it is not possible to add, edit, nor delete any entries in the access log.)

3.2 Making data openly accessible

Semantic Container can store and provide open, closed, and semi-closed data. It is up to the data providing container operator (data controller) to decide if and how data is made accessible to defined recipients. In case data is shared between two parties the following mechanisms exist within a Semantic Container to provide access control:

1. OAuth 2.0: OAuth is an open standard for access delegation and Semantic Container use the Authorization Framework based on Bearer Token Usage as described in RFC 6750.

2. Usage Policy matching: two Semantic Container exchange data if and only if the receiving container has a Usage Policy that is equal or a subset of the providing container. A semantic reasoner is used to evaluate compliance between the Usage Policies.
   Note: Usage Policies are based on the policy language as defined by the SPECIAL project (https://www.specialprivacy.eu) in the following document: https://www.specialprivacy.eu/images/documents/SPECIAL_D21_M12_V10.pdf

3. Digital Watermarking: optionally a unique digital fingerprint is applied to data provided by a Semantic Container, i.e., any data request results in a dataset with insignificant errors that uniquely identifies the recipient of the data set; in case such a dataset is leaked and appears in an unintended location, the person who originally requested and leaked the dataset can be identified

Of course, it is recommended to provide additional access control mechanisms to control access to a Semantic Container within a network.

The Data Vault provides a plugin system based on OAuth 2.0 to access and share data. Since data in the Data Vault is encrypted key management is an important issue to be clarified between sharing parties.

3.3 Making data interoperable

To make data interoperable the Semantic Container platform as well as the OwnYourData Data Vault provide a standardized API documented here:

- Semantic Container: https://api-docs.ownyourdata.eu/semcon/
- Data Vault: https://api-docs.ownyourdata.eu/datavault/

Additionally, inter-container data operations for Semantic Container generate a well-defined provenance trail based on the PROV-O ontology:
https://www.w3.org/TR/prov-o/

3.4 Data re-use & licensing of data

If data providers provide data that is licensed by third parties, they are responsible for disclosing and specifying the licensing terms.

All data, data operations, and the container itself are hashed and the information is written into the public Ethereum blockchain for auditability using the OwnYourData Notary service:
https://notary.ownyourdata.eu

4 Allocation of Resources

4.1 Estimated Costs

At the current state of the project, only a first estimation of costs is possible that occur for the data management task during the project runtime.

The Data Vault is hosted on a Kubernetes Cluster running on 3 root servers at netcup.de with a cost of € 5,99 per server and month. The cost for hosting the Data Vault is provided by OwnYourData at no charge.

The cost for hosting a Semantic Container (acting as a Chat Bot) on a root server at netcup.de is € 13,10. Note that such a server can easily host up to 20 containers and probably more.

Other services for Semantic Containers like notary, validation, billing, and digital watermarking are foreseen to be provided by OwnYourData at no charge.

4.2 Responsibilities

The consortium partner OwnYourData is responsible for implementing the DMP and ensures that it is reviewed and revised during the project runtime.

Name and contact details of the person responsible during the project runtime:

Christoph Fabianek
Michael Scherz Strasse 14
2540 Bad Vöslau
christoph@ownyourdata.eu
+43 677 617 53 112

The project will only be responsible for storing, preserving, and backing up the datasets mentioned in section 2 of this document. Any other data – either open, closed, or semi-closed – stored in Semantic Containers will be the responsibility of the respective container operator (data controller).

4.3 Long Term Preservation

For long-term preservation of a dataset it is recommended to "commit" a container and store the resulting Docker image either on a public repository (e.g., https://dockerhub.com) or in a private repository. See the Semantic Container Design document for a description on how to commit a container.

5 Data Security

Currently, Semantic Containers provide only a very basic level of security through Authorization. It is up to the Semantic Container operator (data controller) to perform regular backup and setup a recovery strategy. Risks related to data security (including illegal procurement or manipulation of data) especially for closed or semi-closed datasets need to be covered based on the concrete nature of the data by the person or organization operating a Semantic Container (the data controller).

6 Legal and ethical Aspects

This section addresses questions related to ethics and legal compliance of the included datasets and defines how ethical issues and IPR are managed in the project.

6.1 Data Protection

Whenever personal data is processed, the compliance with the principles of data protection are to be proven by the controller. These principles encompass, for instance, data minimisation, meaning to only process the data necessary for the pursued purpose. Privacy by design indicates to create data processing technically already in favour of strong protection of personal data. The technical designs of Semantic Containers and the OwnYourData Data Vault are in line with the underlying tenor of avoiding or reducing data processing to the extent necessary.

6.2 Measures to ensure ethical and legal standards

Measures to ensure compliance to ethical and legal standards are currently not part of the Semantic Container framework and the OwnYourData Data Vault. However, all measures are taken to ensure the use of legal and ethical unquestionable datasets in the course of the project.

6.3 Privacy and trust

Issues of privacy and trust amongst data trading participants need to be identified by the Semantic Container operator (data controller) as well as the Data Vault operator (Verein zur Förderung der selbstständigen Nutzung von Daten / Public charity to foster personal use of data, ZVR[1]: 789007092). The following list of indicators leading to negative levels of privacy and trust are used as guideline:

- an inconsistent level of protection for natural persons and private data;
- divergences in the handling and storage of data hampering the free movement of personal data;
- a lack of knowledge regarding data sharing;
- difficulties in determining the trustworthiness of data suppliers;
- lack of knowledge of the law leading to potential violations;
- inconsistent levels of protection for members across participating organisations.

[1] use this link to query the ZVR number in the register of associations from the Austrian interior ministry: https://citizen.bmi.gv.at/at.gv.bmi.fnsweb-p/zvn/public/Registerauszug

7 References

1. Data Market Austria: https://datamarket.at/
2. OwnYourData Notary Service: https://notary.ownyourdata.eu
   API for Notary service: https://blockchain.ownyourdata.eu
   API documentation for Notary service: https://api-docs.ownyourdata.eu/notary/
3. PROV-O: The PROV Ontology - https://www.w3.org/TR/prov-o/
4. Semantic Container API reference: https://api-docs.ownyourdata.eu/semcon/
5. Semantic Container Design document: https://www.ownyourdata.eu/semcon/design
6. Tidepool: https://www.tidepool.org, https://github.com/tidepool-org
7. Novo Nordisk Innovation Challenge: https://novonordisk.innovationchallenge.com
8. Usage Policy: policy language as defined by the SPECIAL project (https://www.specialprivacy.eu) in https://www.specialprivacy.eu/images/documents/SPECIAL_D21_M12_V10.pdf