



micado

migrant integration cockpits & dashboards

D7.5 Data Management Plan



This project has received funding from the European Union's Horizon 2020 Innovation Programme under Grant Agreement No 822717.

Project

Acronym: **MICADO**

Title: Migrant Integration Cockpits and Dashboards

Coordinator: HafenCity University Hamburg

Reference: 822717

Type: Innovation Action

Programme: HORIZON 2020

Theme: Addressing the challenge of migrant integration through ICT-enabled solutions (DT-MIGRATION-06-2018)

Start: 1 January 2019

Duration: 42 months

Website: www.micadoproject.eu

Consortium:

- **HAFENCITY UNIVERSITÄT HAMBURG** (HCU), Germany
- **FREIE UND HANSESTADT HAMBURG** (FHH), Germany
- **HAMBURGISCHES WELTWIRTSCHAFTSINSTITUT GEMEINNÜTZIGE GMBH** (HWWI), Germany
- **UNIVERSITEIT ANTWERPEN** (UANTWERPEN), Belgium
- **OPENBAAR CENTRUM VOOR MAATSCHAPPELIJK WELZIJN VAN ANTWERPEN** (OCMW Antwerpen), Belgium
- **INTEGRATIE EN INBURGERING ANTWERPEN** (Atlas Antwerpen), Belgium
- **DIGIPOLIS** (DIGIPOLIS), Belgium
- **ALMA MATER STUDIORUM - UNIVERSITA DI BOLOGNA** (UNIBO), Italy
- **AZIENDA PUBBLICA DI SERVIZI ALLA PERSONA CITTA DI BOLOGNA** (ASP Bologna), Italy
- **CONSORZIO PER IL SISTEMA INFORMATIVO** (CSI PIEMONTE), Italy
- **COLEGIO PROFESIONAL DE POLITÓLOGOS Y SOCIÓLOGOS DE LA COMUNIDAD DE MADRID** (CPS), Spain

- **CONSEJERIA DE POLITICAS SOCIALES Y FAMILIA COMUNIDAD DE MADRID**
(CPSF-DGSSIS-CM), Spain
- **UNIVERSIDAD REY JUAN CARLOS** (URJC), Spain
- **TECHNISCHE UNIVERSITÄT WIEN** (TU WIEN), Austria
- **SYNYO GmbH** (SYNYO), Austria

Deliverable

Number: **D7.5**

Title: **Data Management Plan**

Lead beneficiary: HCU

Work package / Task: WP7

Dissemination level: Public (PU)

Nature: Open Research Data Pilot (ORDP)

Due date: 30/06/2019

Submission date: 28/06/2019

Authors: **Benedikt Seitzer, HCU**

Contributors: **Agota Barabas, Jan Barski, HCU**

Review: **Carmen Munteanu, SYNNO**
Luca Gioppo, CSI

Acknowledgement: This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 822717.

Disclaimer: The content of this publication is the sole responsibility of the authors, and in no way represents the view of the European Commission or its services.

HISTORY OF CHANGES			
version	date	comment	author(s)
V 1.0	27.05.2019	First full draft. Chapters included: <ul style="list-style-type: none">- Introduction- Data Summary- Data Access- Allocation of Resources- Data Security	Benedikt Seitzer (HCU), Agota Barabas (HCU)
V 1.1	04.06.2019	Contribution	Carmen Munteanu (SYN)
V 1.2	10.06.2019	Second draft	Benedikt Seitzer (HCU)
V 1.3	21.06.2019	Review 1	Luca Gioppo (CSI)
V 1.4	24.06.2019	Review 2	Carmen Munteanu (SYN)
V 2.0 final	28.06.2019	Final draft	Jan Barski (HCU)

Executive summary

The Data Management Plan (DMP) is the central document concerning the handling of research data during and after the MICADO project. This includes the content, storage, and accessibility of the data, as well as data security.

Because significant changes in collected data and storage technologies are expected to arise during the project, the DMP will be updated and is considered a living document. Therefore, versioning of this document is crucial. The DMP will be updated on a yearly basis. An updated version of the DMP will accompany each project report submitted to the European Commission.

Table of Contents

Introduction	8
About MICADO	8
The Data Management Plan	8
Data summary	8
Data Collection and Generation	9
Data Survey	9
Data access	13
Accessibility	13
Making Data Findable	13
Interoperability	14
Re-usability	14
Personal Data	14
Allocation of resources	15
Data security	15
References	15

1 Introduction

1.1. About MICADO

The project “MICADO: Migrant Integration Cockpits and Dashboards” aims to create an EU-wide applicable platform solution in support of migrant arrival and participation processes. MICADO creates digital services for three main target groups: public authorities, civic society representatives, and migrants and refugees.

The project relies on existing and new data consensually provided by migrants, helpers and public authorities in order to generate customized information and services for the target groups. The gathered data has a two-fold purpose, on one hand, it creates the basis for the MICADO’s key technical elements such as chatbot language interfaces and automatic translation services for the migrants’ cockpit to facilitate integration services and communication with authorities. On the other hand, the data collection will be used to create a universal service package, e.g. location and activity mapping, personal profiling, or matchmaking for mentors, jobs, and education opportunities.

1.2. The Data Management Plan

This Data Management Plan (DMP) aims to provide and specify overall methods and approaches for research data collection and processing within the MICADO project. The DMP follows the “FAIR” (findable, accessible, interoperable, and re-usable) principles in order to enhance the data search processes and to support data re-use by the consortium (Wilkinson, 2016).

The DMP is a living document and it will be updated along with the project implementation and as changes are occurring. The DMP includes a timetable for updates in relation to periodic evaluations and major milestones of the project implementation.

This report describes data collection and generation methods, and provides a technical description of data types, formats, and specific data accessibility aspects. Furthermore, the DMP details data security, licensing, and ethical aspects related to data management.

A Nextcloud server has been installed to make data findable, accessible, interoperable, and re-usable within the MICADO consortium. The Nextcloud solution empowers collaboration since it also includes the opportunity to write direct and group messages, have bi- and multilateral calls, and includes an issue-tracking system. Joint editing of documents is also possible.

2 Data summary

The data collected and generated by the partners during the MICADO project should always be related to a purpose and to the objectives of the project stated in section 1.1. To assure this and also to prevent data redundancy, the project lead set up a data survey. This survey will be conducted on a yearly basis and filled out by all project partners since the state of the data will change as the

project progresses. If a project partner has not filled any column of the data survey, the partner is expected to have not generated or collected data so far. The data survey can be seen in section 2.2.

The data collected so far is of different formats and types. It is in most cases communication material and deliverables. This data is saved either as docx- or pdf-file. Also each local group (Madrid, Antwerp, Bologna, and Hamburg) holds a list of the local contacts as the Local Experts and Stakeholder Committees (LESCs). Also statistical data are collected by some of the project partners and are most commonly saved as xlsx-file. This includes local and national surveys as the 'Integrationsmonitoring der Länder' for Germany or the MIPEX tables, as well as Europe-wide indicators such as the 'Eurostat indicators on migrants and integration'.

The total size of data generated and collected by the consortium in the first six months is 330 MB.

Data will be reused during the project. Since there has been a lot of research on social and economic aspects of migration, the consortium includes this secondary data as well to avoid double work. For local analysis, the reused statistical data will be saved and shared with the consortium via Nextcloud.

In the following two sections, the used methods on data collection and generation is described, followed by an overview of the tasks within which data is generated and collected.

2.1. Data Collection and Generation

We distinguish between primary data being generated and secondary data being collected by the MICADO project partners.

As stated in the project's Grant Agreement, a common methodology on data collection will be designed in deliverable D1.2 'Migration Challenges for MICADO' due month 8. This includes the data addressing political, social, economic, and cultural challenges, and mainly focuses on secondary data. A systematic literature review based on the Campbell Systematic Review will be conducted to gain organized information and data on the topics stated above.

The Local Exploration Kit as part of Task 2.1 creates a substantial common methodological framework for running local stakeholder workshops in all partner cities/regions. Besides defining a design for the co-creation sessions with the different user groups, the Local Exploration Kit provides data collection templates for qualitative data. Before the local co-creation sessions begin, the interviewees have to sign a letter of consent (D9.2). A decision on data evaluation taken place on anonymized/pseudonymized or not-anonymized data can freely be made by the interviewee. The local co-creation sessions will take place during months 5-11 to specify local demands and needs in regard to migrant integration. In deliverable D9.5, the processing of personal data is explicitly addressed and determined.

2.2. Data Survey

An overview of the data collected and produced so far can be seen in the

following table. Each data will be linked to the partner collecting/generating the data, the task/deliverable the data is needed for, the type of the data (e.g. personal/non-personal), the origin of the data, and its format (e.g. docx, xlsx, etc.). This data survey will be updated regularly. This way, it can be assured that the data saved fulfils a purpose in relation to the objectives of the MICADO project.

Partner	Description	Task/ Deliverable	Type	Origin	Form at
HCU	All Deliverables	General	Non- Persona l	HCU	docx, pdf
HCU	LESC contacts	T7.3.	Persona l	Local Group Hamburg	xlsx
HCU	co-creation contacts	T2.3	Persona l	HCU	xlsx
FHH	Local contacts	T7.3	Persona l	FHH	xlsx
FHH	Communication Plan, Policy Brief	D7.2	Non- Persona l	FHH	docx
FHH	Communication survey	T7.2	Persona l	FHH	xlsx
FHH	Administrative	General	Persona l	FHH	docx, pdf, xlsx
HWI	Integrationsmonit oring der Länder	D2.2	Non- Persona l	National and Regional Statistical Offices	xlsx
HWI	Integrations Barometer Survey	D2.2	Non- Persona l	Expert Council of German Foundations on Integration and Migration	pdf
HWI	Eurostat indicators on migrants and integration	D1.2.2.	Non- Persona l	Eurostat database	xlsx
HWI	City-level indicators on population structure in	D1.2.3	Non- Persona l	Database "Regionaldatenb ank"	xlsx

	Hamburg				
HWWI	Background information on national and local policies/programmes	D1.2.3	Non-Personal	Public authorities/NGOs	xlsx
UANTWERPEN	co-creation workshops	T2.3	Personal	UANTWERPEN	docx
UANTWERPEN	LESC contacts	T7.3	Personal	Local Group Antwerp	xlsx
OCMW Antwerpen	City-level indicators on population structure in Antwerp	T2.2	Non-Personal	rijksregister, OCMW Antwerpen	xlsx
OCMW Antwerpen	registration system clients of social welfare services	T1.2.3	Non-Personal	e -Vita	
OCMW Antwerpen	LESC contacts	D7.3	Personal	Local Group Antwerp	xlsx
Atlas Antwerpen	Local Context	D1.2.3	Non-Personal	Flemish/Antwerp public authorities	docx
Atlas Antwerpen	Policy Map and MIPEX tables	D1.2.3	Non-Personal	Flemish/Antwerp public authorities	xlsx
Atlas Antwerpen	LESC contacts	T7.3	Personal	Local Group Antwerp	xlsx
UNIBO	Co-creative workshops and interviews	D2.3	personal	UNIBO	docx
UNIBO	national and local policies/programmes	D1.2.3	Non-Personal	Public authorities/NGOs	xlsx
UNIBO	City-level indicators on population structure in Bologna	D1.2.3	Non-Personal	National and Regional Statistical Offices	xlsx
UNIBO	MIPEX Update	D1.2.3	Non-Personal	National and local Laws and policies	xlsx

ASP Bologna	City Context and Policy Map reports	T1.2.3	Non-Personal	Reports / Statistical Reports / ASP Bologna	docx, xlsx
ASP Bologna	Contact list Bologna/Italian Partners	General	Personal	ASP Bologna	xlsx
CPS	City context report	T1.2.3	Non-Personal	CPS	docx
CPS	Mipex update	T1.2.3	Non-Personal	MIPEX	xlsx
CPS	Policy map	T1.2.3	Non-Personal	CPS	xlsx
CPS	Secondary data at local level	T1.2.2	Non-Personal	National and Local Statistical Offices	xlsx
CPS	Literature Report	T1.2.1	Non-Personal	CPS	docx
CPS	Deliverable D.1.2	T1.2.4	Non-Personal	OECD, European Commission, Eurostat, National Statistical Offices	docx
CPS	Co-creative sessions and interviews	D2.3	Personal	CPS	docx
CPS	Co-creative sessions	T2.3	Non-Personal	CPS	xlsx
CPS	Co-creative sessions and interviews recordings transcripts	T2.3	Personal anonymized	CPS	mp3, docx
CPSF-DGSSIS-CM	LESC contacts	T7.3	Personal	Local Group Madrid	xlsx, pdf
CPSF-DGSSIS-CM	immigration statistics annual report	D1.2.3	Non-Personal	CPSF-DGSSIS-CM	pdf
CPSF-DGSSIS-CM	Administrative	General	Personal	Consortium and CPSF-DGSSIS-CM	pdf, docx, xlsx
CPSF-DGSSIS-CM	pilot tests		Personal	Consortium and CPSF-DGSSIS-CM	pdf, docx, xlsx

CPSF-DGSSIS-CM	contact list of stakeholders and networks	General	Non-Personal	CPSF-DGSSIS-CM	xlsx, docs
TU WIEN	Deliverable	D1.3	Non-Personal	Consortium	doc, pdf
TU WIEN	Figures for D1.3	D1.3	Non-Personal	TU Wien	ai, png, pdf, psd
TU WIEN	Administrative	General	Personal	TU Wien, Consortium	pdf, docx, xlsx

3 Data access

In the following sections, specifications on making the data FAIR within the consortium are made. Also a section on personal data is included, since the MICADO consortium processes personal data of representatives of vulnerable social groups.

3.1. Accessibility

Access to the project-related data will be ensured within the consortium via Nextcloud. The servers on which Nextcloud is hosted are administered by the HCU and no cloud solution by external actors is hosting any data collected or generated by the MICADO consortium.

Nextcloud is available as a website (<https://micado.hcu-hamburg.de>) and compatible with the following browsers (https://docs.nextcloud.com/server/13/admin_manual/installation/system_requirements.html):

- Microsoft Internet Explorer 11 (latest version)
- Microsoft Edge
- Mozilla Firefox
- Google Chrome/Chromium
- Apple Safari

The Nextcloud server can also be embedded using a desktop client and mobile devices. Desktop clients are available for the following operating systems:

- Windows 7+
- macOS Lion (10.7)+ (64-bit only)
- Linux (CentOS 6.5+, Ubuntu 14.04+, Fedora 21+, openSUSE 13, SUSE Linux Enterprise 11 SP3+, Debian 8 (Jessie)+, Red Hat Enterprise Linux 7)

Applications for mobile devices can be downloaded in the Google Play Store and the AppStore for at least the following versions:

- iOS 9.x+
- Android 4.x+

Data can also be synchronised locally using the desktop clients to work on the data even if there is no internet access. This makes the data highly accessible for all project partners no matter which operating system is used or if there is internet connection.

Each collaborator has been equipped with an account and has access to work package-related data and administrative data concerning the project. Also, each Local Group (Bologna, Antwerp, Madrid, Hamburg, and Vienna despite not being a pilot city) will have an own folder, which only the members of the Local Group can access (read and write). These folders have the purpose of storing personal data such as contact lists for the LESCAs and even attendees lists for the co-creation sessions.

3.2. Making Data Findable

To make all kinds of data findable for the project partners, the folder structure on Nextcloud is adjusted to the work packages. Each project partner is informed about conventions by a 'readme' and a short introduction to Nextcloud.

Each file uploaded to Nextcloud will be tagged with keywords, making it easier to find it applying the internal search provided in the platform.

3.3. Interoperability

MICADO is an interdisciplinary project with sociologists, economists, computer scientist,s and urban planners working together. To achieve a common image of the project and facilitate convergence, a glossary has been set up under the public github-project of MICADO (<https://github.com/micado-eu/MICADO/wiki/Glossary>). Each project partner is invited to contribute to the glossary.

The platform independence of the Nextcloud solution (see section 3.1) ensures interoperability concerning the data access and also enables users to work collaboratively on one shared document at the same time. The latter is made possible by installing the add-on Collabora.

3.4. Re-usability

Article 26 of the Grant Agreement (Chapter 4, Section 2, Subsection 3) discusses the ownership of results. Results being generated during the MICADO project are owned by the beneficiary that generates them. This includes any (tangible or intangible) output of the action such as data, knowledge, or information — whatever their form or nature, whether it can be protected or not — that is generated in the action, as well as any rights attached to it, including intellectual property rights. As long as no personal- or copyrights are violated, the consortium aims to publish all results produced.

IP and data exploitation and dissemination strategies are defined further in deliverable D7.1. The Plan for the Exploitation and Dissemination of Results (PEDR) will be periodically updated, including a General as well as Local Exploitation Plan. Also, operational and business models will be schemed.

Long-term storage and curating of the data after the project duration has ended will be discussed in more detail after the PEDR has been published. For example, the Zenodo Repository funded by the European Commission could be a feasible solution.

3.5. Personal Data

In context of the MICADO project, personal data will be collected. The collected data will be saved on Nextcloud on several levels. Anonymized data can be saved centralized in the work folders to which every collaborator has access. This includes, for example, anonymized lists of interviewees taking part in the co-creation sessions. The non-anonymized lists of participants will be saved in the Local Group folders so that only the collaborators who actually need to work with the data (invitation to further workshops, etc.) have access to them.

As already stated, all participants of co-creation sessions will need to sign the letter of consent for the collection and processing of personal data for research processes (D9.2).

As stated in D9.4, the MICADO project will not include processing previously collected personal data. In case the consortium discovers the necessity to apply this method, a relevant partner will provide explanation how the data subjects will be informed of how previously collected data will be processed. The project activities will include, however, processing previously collected aggregated (anonymised) data.

4 Allocation of resources

Setting up a github-project does not produce any costs. The publishing of the MICADO source code and documentation is therefore ensured. Also versioning and a changelog are integrated within github.

For internal project communication and data sharing, the Nextcloud solution with the add-on Collabora has been set up. Both Nextcloud and Collabora are open source products and licensing is done via the GNU Affero General Public License (AGPLv3, <http://www.gnu.org/licenses/agpl-3.0.de.html>). Even though the software is free of charge, costs are generated by the two servers used. These costs will be paid as an in-kind contribution of 250€/year and server by the HCU. The physical server specifications are listed below:

Implemented Solution	Number of cores	RAM	Storage
Nextcloud Webserver	4	32 GB	550 GB
Collabora Server	8	16GB	25 GB

5 Data security

This chapter points out the main steps followed by the consortium, to make

project related data secure.

In Deliverable D9.1, a Data Protection Officer (DPO) is named to ensure GDPR compliance of the processing and storage of data. The data protection officer consults the consortium and is asked for advice each time decisions and steps on data collection, data storage, and processing have to be made. For example, the DPO voted for Nextcloud as the data storage solution and also influenced methods on data collection such as the Local Exploration Kit.

The Nextcloud solution is open source and has been audited by the NCC Group. The review against the control Clause 14 of ISO27001 'Security in Development and Support processes' was done in 2016 and can be downloaded from the Nextcloud site (<https://nextcloud.com/secure/>).

Since Nextcloud is an on-premise service, Nextcloud GmbH has at no time access to the data stored on the server. Nextcloud also provides event logging, backup tools, fine-grained access control to files via groups, and the secure sharing of links by using passwords and expiration dates. Nextcloud uses industry-standard SSL/TLS encryption for data in transfer.

6 References

Literature:

Wilkinson, D. M. (2016) *The FAIR Guiding Principles for scientific data management and stewardship*. Scientific Data. 3:160018, DOI: 10.1038/sdata.2016.18

Websites:

<http://www.gnu.org/licenses/agpl-3.0.de.html>, retrieved on 07.06.2019

<https://github.com/micado-eu/MICADO/wiki/Glossary>, retrieved on 07.06.2019

<https://micado.hcu-hamburg.de>, retrieved on 07.06.2019

<https://docs.nextcloud.com>, retrieved on 07.06.2019

<https://nextcloud.com/secure/>, retrieved on 10.06.2019