Creating Opportunities by Multi-indicator Measures towards Unity (COMMUNITY) project

A Data Management Plan created using DMPonline.be

Creator: Leen d'Haenens

Affiliation: KU Leuven (KUL)

Funder: IOF

Template: KU Leuven BOF-IOF

Grant number / URL: C24M/22/007

ID: 209299

Start date: 01-10-2022

End date: 30-09-2026

Project abstract:

COMMunity (Creating Opportunities by Multi-indicator Measures towards Unity) proposes unique and timely interdisciplinary strategic research on the intertwining relationship between (changes in) anti-discrimination policy, legislation, media coverage on policy, cases and equality bodies, and public perceptions of anti-discrimination policy in Belgium. The societal and scientific urgency for such an approach has strongly increased due to 1) the UN 2030 Agenda on SDG's (Sustainable Development Goals) including 1, 5, 10 and 16 focusing on poverty, inequality and strong institutions to guarantee justice, as a clear plea to researchers to address one of society's greatest grand challenges; 2) the withdrawal of Flanders from inter-federal equality body Unia and the establishment of a new Flemish Human Rights Institute (Vlaams Mensenrechteninstituut, 2022); 3) increasing news media attention on 'contested' equality bodies' practices in Belgium (see e.g., the role of the Institute on the Equality of Men and Women in the lawsuit against media figure Bart De Pauw in 2021); and 4) the impact of the enduring Covid situation on the number of discrimination notification reports (see Unia 2020 report, p. 9: 16% of discrimination notifications are Covid-related). All this emphasises the urgent need for a thorough interdisciplinary scientific analysis of policy changes and impact, (news) media attention, and public perceptions of anti-discrimination policy, equality bodies and actions to protect citizen rights. The COMMunity proposal is timely, embracing a broad citizen protection paradigm, insisting that the scientific and societal debates go beyond the mere knowledge and attitudinal dimensions into the direction of empirically visible behavioural public actions as fostered by its clear valorisation objectives and co-creation initiatives with relevant societal stakeholders (equality bodies, policy makers, civil rights organisations, anti-discrimination interest groups, media organisations, and education partners).

Last modified: 27-07-2024

Creating Opportunities by Multi-indicator Measures towards Unity (COMMUNITY) project

Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset name / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
		Indicate: ***N****(ew data) or* ***E****(xisting data)*	Indicate: **D**(igital) or **P**(hysical)	Indicate: **A**udiovisual **I**mages **S**ound **N**umerical **T**extual **M**odel **SO**ftware Other (specify)		Indicate: <1GB <100GB <1TB <5TB >5TB NA	
WP1	Policy docs	E	D	T	.pdf	<100GB	
WP2	Media analysis	E	D	AV	.pdf	<100GB	
WP3	Mapping attitudes	N	D	survey data	.spss	<100MB	
WP4	Campaign analysis	N	D	interview transcripts	.word	<100MB	
WP5	Valorisation research	N	D	Survey data/experiments	.spss	<100MB	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

In the context of WP1 and WP2, we will make use of the iCANDID infrastructure as the main data collection and analysis platform. iCANDID can integrate access to different types of textual and audio-visual news data harvested from freely accessible online media resources (e.g., digital media content) and proprietary media repositories (e.g., Belga.press, ENA). https://icandid.libis.be/#/

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.

- Yes, human subject data (Provide SMEC or EC approval number below)

Ethical approval for the project, including all work packages, has been obtained from the KU Leuven Social and Societal Ethics Committee: G-2022-6002-R2.

Will you process personal data? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).

- No

All survey and interview data will be analysed at the aggregated level.

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or Data transfer agreements, Research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, codebook.tsv etc. where this information is recorded).

A codebook will be developed for each data set, which will be uploaded with the corresponding dataset when these are published in repositories. The codebook acts as a comprehensive reference guide for every variable in the dataset. It includes detailed information such as variable names, descriptions, data types, and units of measurement. This detailed documentation ensures that users can interpret and analyze each variable accurately. Beyond the technical details, the codebook will also provide context and meaning to the data. It explains the rationale behind each variable, its relevance to the research, and any assumptions or considerations made during data collection. This contextual information is invaluable for users seeking to understand the dataset in its entirety.

Will a metadata standard be used to make it easier to find and reuse the data?
If so, please specify which metadata standard will be used.

If not, please specify which metadata will be created to make the data easier to find and reuse.

- Yes

For the interview data, which will be deposited on KU Leuven RDR, DataCite will be used as a metadata standard. KU Leuven RDR is the institutional repository that helps KU Leuven researchers publish, share, cite, and preserve their research data in a findable, accessible, interoperable and reusable way.
For the online survey data, a data article will be written in Data in Brief (https://www.sciencedirect.com/journal/data-in-brief), in which all procedures and relevant measures in the survey will be outlined. This article will also contain a link to the dataset on Mendeley Data.

Data Storage & Back-up during the Research Project

Where will the data be stored?

- Shared network drive (J-drive)

The data will be stored on a shared shared network drive at KU Leuven. For KU Leuven staff with a PC or laptop managed by ICTS or a local IT organisation, a shared network drive is provided by default. KU Leuven offers a robust platform with impressive features. With no hard

limit on storage size, we can store and access extensive datasets critical for scientific work. The flexibility of collaboration both within and outside KU Leuven enhances the potential for interdisciplinary partnerships.
For external collaborations, the minimal registration requirement ensures a balance between accessibility and security. All data is housed in KU Leuven datacenters, guaranteeing a secure and controlled environment. The high availability, coupled with storage mirroring in a second datacenter, reinforces data resilience and uninterrupted access. In essence, KU Leuven's data infrastructure aligns with the rigorous standards necessary for the scientific community, providing a secure and collaborative environment for researchers.

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution

KU Leuven employs cutting-edge "snapshot" technology for its backup system, ensuring the meticulous preservation of incremental data changes. The standard backup regime, outlined below, reflects a comprehensive approach to data protection. There will be daily backups at midnight, capturing the state of the data at the end of each day. The system preserves the last 6 daily backups, ensuring a comprehensive daily history.
This backup strategy not only emphasizes the frequency of data capture but also ensures that a varied time span of backups is available. The reserved storage capacity guarantees that these backups are readily accessible, contributing to a robust and reliable data management system at KU Leuven.

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

The newly generated datasets that will be generated are relatively small (all around or well below 1GB), which presents no problem at all for the KU Leuven storage system. The existing media data sets are part of the iCANDID data hub (operated by LIBIS).

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Access is only possible through KU Leuven's highly secure system. The data is stored encrypted on the central storage infrastructure, ensuring that only KU Leuven staff has access to the data.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

No estimated costs for the newly generated research data. The existing media data are part of the iCANDID data hub (a HERCULES mid-size infrastructure covered by FWO subsidy until April 2028).

Data Preservation after the end of the Research Project

Which data will be retained for 10 years (or longer, in agreement with other retention policies that are applicable) after the end of the project?

In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

- All data will be preserved for 10 years according to KU Leuven RDM policy

In accordance with the KU Leuven RDM policy, the data will be retained on the KU Leuven network system, for no additional cost, for 10 years. After 10 years, we will evaluate whether it is still necessary to store the data. If not, they will be deleted. If they are still needed, a new date will be set for reevaluation. Only researchers who have worked on this study will have access to the data after the end of the project.

Where will these data be archived (stored and curated for the long-term)?

- KU Leuven RDR

The survey data will be permanently archived on Mendeley Data, while the policy data, media content data will be kept in the iCANDID repository. The interview transcripts will be deposited on KU Leuven RDR.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

No cost.

Data Sharing and Reuse

Will the data (or part of the data) be made available for reuse after/during the project?
Please explain per dataset or data type which data will be made available.

- Yes, as restricted data (upon approval, or institutional access only)
- Yes, as open data

The online survey data will be uploaded for free use on Mendeley Data. The policy document data, media content data will be part of the iCANDID data hub, and the interview transcripts will be uploaded to KU Leuven RDR.

If access is restricted, please specify who will be able to access the data and under what conditions.

Only the researchers who were involved in the current project will have access to the data.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

Please explain per dataset or data type where appropriate.

- No

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- KU Leuven RDR (Research Data Repository)
- Other data repository (specify below)

Mendeley Data for online survey data;
KU Leuven RDR for the interview transcripts;
Media content data and policy texts on iCANDID.

When will the data be made available?

- Upon publication of research results

Data will be made available upon publication of research results.

Which data usage licenses are you going to provide?

If none, please explain why.

- CC-BY 4.0 (data)

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

- Yes, a PID will be added upon deposit in a data repository

A DOI will be developed for each data set.

What are the expected costs for data sharing? How will these costs be covered?

None.

Responsibilities

Who will manage data documentation and metadata during the research project?

Leen d'Haenens

Who will manage data storage and backup during the research project?

Leen d'Haenens

Who will manage data preservation and sharing?

Leen d'Haenens

Who will update and implement this DMP?

Leen d'Haenens