Ref. Ares(2022)7673148 - 07/11/2022

Horizon 2020 Grant nº 770309
Religious Toleration and Peace
*Leuven/ Mainz 31 October 2022*

# D 4.2 Data Management Plan

| | |
|---|---|
| **Contributor(s):** | Ilenia Laudito and Peter Heyrman |
| **Institution:** | IEG Mainz/ KU Leuven |
| **Contact:** | Peter.Heyrman@kuleuven.be |

| | |
|---|---|
| **Related project task:** | T 1.1/ T 4.2 (T 4.3) |
| **Type of document:** | report |
| **Document reference:** | […] |
| **Version number:** | […] |
| **Status:** | Final |
| **Last revision date:** | 31 October 2022 |
| **Verification date & verifying subject:** | 31 October 2022 |
| **License:** | NA |
| **Other, if applicable:** | […] |

This project has received funding from the European Union's Horizon 2020 Research & Innovation programme under Grant Agreement no. 770309.

# Table of contents

1. CONTEXT AND OBJECTIVES 3

**RETOPEA** 3

**Premises on data management and open data** 3

**Support** 4

**Development of Retopea's DMP** 4

**European collaboration: OpenAire** 4

2. DATA SUMMARY 5

3. OVERVIEW OF PAST AND SCHEDULED DATA MANAGEMENT ACTIONS 5

**3.1. Retaining and depositing materials** 5

**3.2. Publications** 6

**3.3. Clippings and Docutubes** 6

**3.4. Website** 7

4. DATA HANDLING AND WORKFLOW PLAN FOR CLIPPINGS 7

**4.1. Stage 1: Creating the clipping** 8

**4.2. Stage 2: Testing the clipping for readability** 8

**4.3. Stage 3: Translating the clipping & metadata/keywords** 9

**4.4. Stage 4: Sending the clipping to the partners for translation review** 9

**4.5. Stage 5: Sending the clipping to the Research Data Manager** 10

5. METADATA AND TRANSLATION OF CLIPPINGS 10

6. INGESTION OF THE DOCUTUBES 11

7. COSTS AND ALLOCATIONS OF RESOURCES 11

# 1. Context and objectives

## RETOPEA

RETOPEA brings together eleven partners across Europe to promote religious peace and toleration, particularly amongst young people. It started from the premise that Europe's many historical experiences with establishing peace and toleration after a religious conflict offer a source for inspiration for the present and that this reflection could inspire both students (age 12-18) and policymakers. The project (1 May 2018 to 31 October 2022) produced so-called 'clippings' of historical and contemporary source material on the topic, that can accessed in the online OMEKA-VRE of the project. These clippings can be used in different educational settings and help teenagers to reflect on their own dealings with religious differences. Retopea particularly encourages the production of short videos on the topic, so-called docutubes, as a creative, dynamic approach to helping young people think with history about religious diversity, past and present in both formal and informal settings. The videos created during the project and available on it website will inspire future productions of docutubes, e.g. linked to a Batched Open Course (BOC) on the topic, offered by Retopea-partner Open University (UK). The Retopea-team also examined diverse ways in which religious diversity past and present is represented in diverse media contexts: school textbooks, museums, TV-series, YouTube videos, as well as in political discourse and places of interconvictional interaction. The main research results were published in a collective volume published in 2022 by Bloomsbury Academic. The team also wrote several reports, for instance policy recommendations for interconvictional dialogue and education.

The Data Management Plan (DMP) of RETOPEA outlines and explains how the team handled, collected, processed or generated research data during the project. It describes the data management life cycle for all relevant datasets, highlights the implemented methodology and standards, outlines how these data are and will be shared, curated and preserved.

## Premises on data management and open data

RETOPEA expressed its firm commitment to the open access of its (scientific and public-oriented) output and the research data and materials produced, collected and used during the project. The consortium conforms to this policy: regarding the digital research data generated during the project ('data'), the beneficiaries must deposit them in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate – free of charge for any user – all the research data produced during the project. These include

- The data, including associated metadata, necessary to validate the research results presented in scientific publications and reports.
- Other data, including associated metadata, as specified, and within the agreed deadlines.

Given the diversity of the source materials and data-packages used and produced by Retopea (see above), their accessibility needed to be evaluated bearing in mind privacy, ownership and copy rights or 3rd party agreements explicitly preventing us from sharing the data involved. The RETOPEA data manager (WP 4) coordinated and publicly documented this process in close consultation with the research teams involved and the PI/general supervisor.

RETOPEA provides ample public information and documentation on its digital infrastructure and the tools and instruments used by its researchers. Building and managing the digital infrastructure of RETOPEA was the main objective of the project's WP 4. Deliverable 4.5. offers a good overview of this

infrastructure, D 4.4. specifically reports on the VRE, D 4.5. provides an online manual. D 4.9, 4.10 en 4.12 report on the integration of the data with important external instruments, e.g. the *Digital Edition of Early Modern Religious Peace Treaties* and EGO – European History Online, both hosted by IEG.

## Support

The KU Leuven RDM team supported Retopea and its research team in the area of RDM (or Research Data Management, including guidance to legal requirements and policy guidelines, based on the templates provided by the Digital Curation Centre (UK), and a dedicated helpdesk advising on data storage, metadata and preparing data for sharing. DMPs are reviewed by a team of RDM support staff, who provide feedback to the research teams. DMPs with feedback are kept as a record, so researchers can retrieve DMPs of past projects at all times.

KU Leuven's data policy demands retention of data for a period of at least 5 years after the end of a project, PhD or publication (barring 3rd party agreements). The university also continues to invest in affordable long-term storage and curation.

KU Leuven digital assets management system Teneo is the main instrument of RETOPEA for digital preservation. It is an e-depot with a secured storage system that centrally gathers and stores files and metadata. Teneo guarantees the safe storage and preservation of these data with back-ups in a redundant and geographically distributed environment following the rules of disaster recovery. Teneo archives and provides persistent access to born-digital or digitized data objects and datasets and includes metadata, usage policies, upload and access tools. Its scalable and extensible architecture makes it a sustainable and trusted preservation infrastructure. Teneo complies with all international standards for digital preservation (such as OAIS and Premis). From the start of the data entry, its automatic processes immediately support the validation of the content. In the use of Teneo Retopea was supported by the curation services of LIBIS-KU Leuven.

KU Leuven experts assisted RETOPEA in checking and complying with the regulations on intellectual property rights, research ethics, privacy (incl. the definition of personal data), and the ICT code of conduct for staff on security and safe storage of data.

## Development of Retopea's DMP

The project team developed a full but initial data management plan in October 2018, month 6 of the project (deliverable 4.1). This plan was implemented and continuously updated during the project in accordance with the project's reflection on and experiences with data management.

A revised/updated version was presented at the half-way meeting, scheduled in Granada but held online on 16-17 September 2020, due to corona-restrictions.

The present document is the final DMP, as scheduled in the workplan of the project.

## European collaboration: OpenAire

At the end of the project RETOPEA will upload its main data-packages, in particular the clippings that were produced, towards the Open Access Infrastructure for Research in Europe (OpenAIRE). The project's data archive was made compatible with the OpenAIRE infrastructure by implementing DataCite Metadata Schema v3.1 and the OAI-PMH v2.0 protocol for harvesting dataset metadata.

Standard and commonly accepted formats (PDF/A, ODF, JSON, JPEG2000) will be used to archive and share the data.

## 2. Data Summary

RETOPEA worked with the following types of data:

| Origin | types | Extensions |
|---|---|---|
| Clippings | Text files | .docx, .pdf |
| | Images | .jpeg |
| | Videos | .avi, .wmv, .mp4, … |
| Docutubes | Videos | .avi, .wmv, .mp4, … |
| Research files | Reports | .docx, .pdf |
| | Deliverables | .docx, .pdf |

In quantitative terms, work packages 2 and 3 produced 400 clippings. A limited but highly relevant selection was preferred over a more extensive but more disparate one.

Work package 5 produced 26 docutubes, a number slightly higher than originally previsioned.

Within the context of the project 60 reports were written, mainly linked to the deliverables of Retopea.

The metadata accompanying these files were structured in accordance with the Dublin Core standards, which were deemed refined and standardized enough. For further information on the harmonized meta data scheme, see D 4.8.

## 3. Overview of past and scheduled data management actions

### 3.1. Retaining and depositing materials

The RETOPEA partners and their researchers retained all the research data they produced and used during the course of their personal research work. They were strongly advised to deposit those materials in the research data repositories of their different institutions. For KU Leuven this is RDR (pronounced "Radar") built on Dataverse - open source repository software developed by Harvard University. RDR gives KU Leuven researchers a one-stop platform to upload, describe, and share their research data, conveniently and with support from university staff.

The main deliverables of the project (WP 2, 3 and 5) and the research datasets and materials used to produce them, together with their (administrative, technical, structural,,…) metadata, were centralized by the RETOPEA data manager (IEG Mainz). To this end a joint and centrally managed document and file repository was created in SeaFile, a reliable and open source enterprise file sync and share platform, commonly used by Retopea-partner IEG Mainz and easily accessible for all the partners. A transparent folder structure was implemented. D 4.5. provides further information and a manual.

The files collected by the data manager are destined to be uploaded to the long-term digital preservation repository of KU Leuven and its digital assets management system Teneo, supported by the curation services of LIBIS-KU Leuven. This ingest will be executed early November 2022.

A formal submission agreement for the Retopea data packages uploaded to Teneo, will subsequently be concluded specifying the intellectual property rights. Possible issues on ownership issues will be handled by the PI in accordance to EC guidelines and rules.

## 3.2. Publications

All scientific publications resulting from RETOPEA are made open access i.e. free of charge online access for third users.

The collective volume Riho Altnurme, Elena Arigita and Patrick Pasture (eds), *Religious Diversity in Europe: Mediating the Past to the Young* (London: Bloomsbury Academic, 2022) is published in hardcover; a paperback is scheduled for 2023. All individual chapters, however, are accessible in Green Open Access via the public depository of each chapter by one of the authors from November 2022 onwards. We did not dispose of the budget to publish in gold open access.

The scientific publications of the different members of the research team are made accessible via the academic institutional repositories of the different partner universities. KU Leuven researchers for instance deposited their publications in the Leuven Institutional Repository and Information Archiving System (LIRIAS). It is a local Elements instance ((c) Symplectic UK), with a DSpace repository underneath for the storage of the full-texts.

## 3.3. Clippings and Docutubes

The 400 clippings composed by the Retopea-team can be accessed through the RETOPEA web platform and its public access tools and instruments (OMEKA). They are grouped according to twelve themes: gender and sexuality; migration, trade and travel; propaganda, stereotyping and communication; discrimination and otherness; ideas about tolerance; peace and conflict management; memory and heritage; law, police and public order; places and buildings; religious practice; clothing; family life. Each clipping is provided with contextual information and with one or a few stimulating questions to trigger reflection among students.

The 26 docutubes produced by the consortium as examples and the main reports (deliverables) of the project are accessible on the website.

As mentioned earlier, these clippings, docutubes and documents are also uploaded to Teneo, thus ensuring their long-term preservation and accessibility. Each object in the e-depot is provided with a unique code or Persistent Identifier (PID). These URI's are used to keep objects accessible beyond the end date of the project.

## 3.4. Website

In order to foster the accessibility and use of the website and its underlying VRE in OMEKA as an ongoing source of information for learning with history on religious toleration and peace, this online environment was thoroughly reworked in the last months of the project. Several pages were removed, others redesigned. The website was somewhat transformed from a project website into an online resource on the topic. This make-over particularly fostered the retrievability of the clippings (in the OMEKA VRE) and the model-docutubes produced by the consortium. These contents will support the further valorisation of the project's results in formal and informal educational contexts. The website particularly hopes to foster derivative initiatives to create docutubes, in particular linked to the free online Badged Open Course, 'Young people and religion: creative learning with history', developed by the RETOPEA project members in the UK.

KADOC-KU Leuven will ensure that the project website will be kept online for at least five years after the end of the project, offering durable technological support for its main components, the OMEKA-VRE in particular.

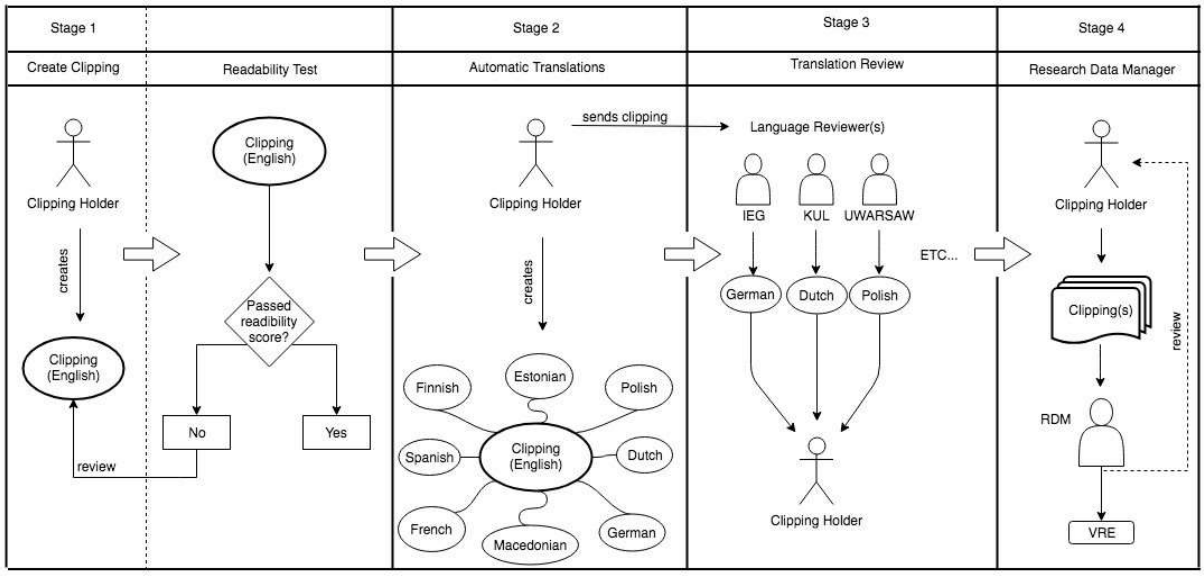
# 4. Data handling and workflow plan for Clippings

This section represents the workflow between the work packages with particular attention to creation, translation, review and final upload of clippings. This workflow applies to all written content as well as to metadata keywords.

The structure of the workflow was as follows:

- Creating the clipping
- Testing the clipping for readability
- Translating the clipping & metadata/keywords
- Sending the clipping to the partners for translation review
- Sending the clipping to the Research Data Manager

This structure can be visualized as follows:



Please consider the following terms:

| | |
|---|---|
| Clipping bundle | A bundle consists of the original clipping and the clippings translations in eight languages. |
| Clipping Holder | Individual in charge of and responsible for the quality and correctness of the clipping(s). |
| Language Reviewer | Individual nominated by the designated institution, in charge of and responsible for the quality and correctness of the clipping's translation in the appointed language. |

## Stage 1: Creating the clipping

Clippings may consist of different file types and appear in various formats. All the clippings will have a structured metadata vocabulary. Further information about the translation and the metadata structure will be in explained in "stage 2" and at this link respectively.

Clippings with written content are divided as follow: for longer and complex texts a readability test is necessary to match pupil's reading level; for every other written part (i.e. image captions, metadata keywords, etc…) follow the instructions in "stage 2".

Procedure for video documents will be defined later on.

## Stage 2: Testing the clipping for readability

This stage is especially relevant for clippings containing mainly written text. It is recommend to analyze its reading ease and grade level by testing its readability. The readability test and the score results are based on and apply to English written texts only. This way the Clipping Holder will be able to ensure the appropriate understanding level for its target. The project's target group's age will be between 13 and 18 years, which corresponds to children attending classes between the seventh to the twelve grade. The clippings can be tested via this easy-to-use online tool.

The relevant scores for each test in prioritized order are as follows:

| | Flesch-Kincaid reading ease score | ARI (automated readability index) | SMOG index | Flesch- Kincaid grade level | Gunning Fog index | Coleman- Liau index |
|---|---|---|---|---|---|---|
| Score | 90.0 - 50.0 | 7 - 12 | 6-12 | 6-12 | 6-12 | 6-12 |

If the results of at least four of six tests are in the above range, the clipping can be considered as an understandable text for the target group. If the results extremely differ from the above scores, the

Clipping Holder should consider to re-edit the clipping.

Here is a brief example on how the scores can vary by simply splitting a long sentence in two shorter ones:

The Peace of Augsburg was a treaty signed between Charles V, Holy Roman Emperor, and the forces of the Schmalkaldic League on September 25, 1555 at the city of Augsburg in Germany.

The Peace of Augsburg was a treaty signed between Charles V, Holy Roman Emperor, and the forces of the Schmalkaldic League. The treaty was signed on September 25, 1555 at the city of Augsburg in Germany.

| | Flesch-Kincaid reading ease score | ARI (automated readability index) | SMOG index | Flesch- Kincaid grade level | Gunning Fog index | Coleman- Liau index |
|---|---|---|---|---|---|---|
| Score | 90.0 - 50.0 | 7 - 12 | 6-12 | 6-12 | 6-12 | 6-12 |
| Example 1 | 44.8 | 15 | 11.6 | 15 | 12.8 | 9.8 |
| Example 2 | 61.7 | 8.1 | 8.3 | 9 | 7.2 | 9.9 |

## Stage 3: Translating the clipping & metadata/keywords

The clippings needed to be translated in eight different languages: German, French, Spanish, Polish, Dutch, Macedonian, Estonian and Finnish. To accelerate the clipping's translations from the English source to the above mentioned languages, the following tools were used:

- DeepL. for translations from English to German, French, Spanish, Polish, Dutch
- Google Translate for translations from English to Macedonian and Estonian

Likewise all metadata keywords are translated in the mentioned eight languages with Babelnet.

## Stage 4: Sending the clipping to the partners for translation review

Subsequently the automated translations, the clippings presumably require a minor grammatical and semantic revision. Each institution nominated a Language Reviewer, who will be responsible for the quality and correctness of the translated clipping. The Clipping Holder sends the automatically translated clipping to the Language Reviewer. The Language Reviewers will afterwards send the examined and reviewed version of the clipping back to the Clipping Holder.

| Language | Partner | Contact |
|---|---|---|
| German | IEG | juergens@ieg-mainz.de |
| French | Le Foyer | ann.trappers@foyer.be |
| Spanish | UGR | earigita@ugr.es |
| Polish | Uniwarsaw | michal.choptiany@al.uw.edu.pl |
| Dutch | KUL | bram.deridder@kuleuven.be / Christophe Schellekens |
| Macedonian | MLIMC/MCIC | georgieva@fzf.ukim.edu.mk |
| Estonian | Utartu | riho.altnurme@ut.ee |
| Finnish | HELUNI | arpad.welker@gmail.com |

### Stage 5: Sending the clipping to the Research Data Manager

The Clipping Holder returned the clipping bundle to the Research Data Manager (RDM - see contacts below), who re-examined the clipping bundle, before inserting and integrating the data in the VRE. If the clipping bundle presented imprecisions, the RDM returned the clipping bundle back to the Clipping Holder for reexamination.

| | | |
|---|---|---|
| Laudito, Ilenia | IEG | laudito@ieg-mainz.de |

## 5. Metadata and Translation of Clippings

The metadata templates were examined and inspected by the RDM, transforming the un-normalised metadata sheets into harmonised metadata. This included checking spelling and punctuation mistakes, to avoid formatting issues when ingesting the csv files, verifying links and webpages, and correcting double categorisations, tags and redundancies. Questions, mistakes and other issues were discussed with the researcher.

After receiving confirmation from the researcher, the harmonised sheets were re-sorted and prepared for the first upload in Omeka. During the first upload, only the following tags were included: Alternative Title, Title, Content, Description, Context focus, Questions, Temporal Coverage, Date, Spatial Coverage, Subject, Source, Bibliographic Citation, Transcript of, Translation of, Audience, Creator, Contributor.

The automatically generated metadata was obtained by retrieving the relavant data from BabelNet using the Java API.

The excel/numbers sheets were then exported into a csv file, conforming Omeka's prescriptions. The csv files were uploaded in Omeka and the <context> and <content> tags transformed into HTML code, so that Omeka classifies it as "media" (in contra to "item").

## 6. Ingestion of the Docutubes

A total of 26 docutubes were created by the participants, which are currently displayed in a dedicated section of RETOPEA's website along with an explanatory text giving context to RETOPEA's general plan and this particular activity.

A more detailed description of the ingestion has been described in deliverable 4.13.

## 7. Costs and Allocations of Resources

The costs of the data management practices and the allocation of resources were covered by the Grant Agreement.