Health Campus The Hague

Data Governance

Data-infrastructure Extramural LUMC Academic Network (ELAN)

24-03-2023

Table of contents

1. Objectives

The Extramural LUMC Academic Network (ELAN) data research project aims to contribute to healthcare innovation and innovation in the field of health and well-being in the Haaglanden and Hollands-Midden region. ELAN is developing a regional data infrastructure that securely links routine healthcare data from different domains in healthcare to other relevant data sources. This regional data infrastructure is a prerequisite for developing, indicating, implementing, and evaluating interventions. ELAN's perspective is broad, so healthcare is not only seen as healthcare provided in the (clinical) healthcare domain, but also includes associated social problems.

2. Embedding

ELAN was developed from the Health Campus with the objective of investigating health differences between populations in order to improve personalized healthcare. ELAN is filled with general practitioner data, hospital data, data from the municipality of The Hague (Law on social support, WMO in Dutch), Parnassia Group (mental healthcare), microdata from CBS, data from Hadoks (chronic and acute care), and Perined (perinatal data). ELAN uses routinely collected patient data and does not collect data itself. Therefore, ELAN has a mainly facilitating function. ELAN is carried out in a multidisciplinary manner under the direction of the Health Campus, in which the municipality of The Hague, The Hague hospitals (HMC and HAGA Hospital), Parnassia Group (mental healthcare), GGD, and LUMC are represented at the management level.

The ELAN data research project supports the regional program 'Healthy and Happy The Hague', among others. This regional program was started by a broad representation of healthcare and welfare, the municipality of The Hague, health insurers, and residents, in collaboration with the Health Campus. For the execution of research under the flag of the 'Healthy and Happy The Hague' program, data from ELAN can be used.
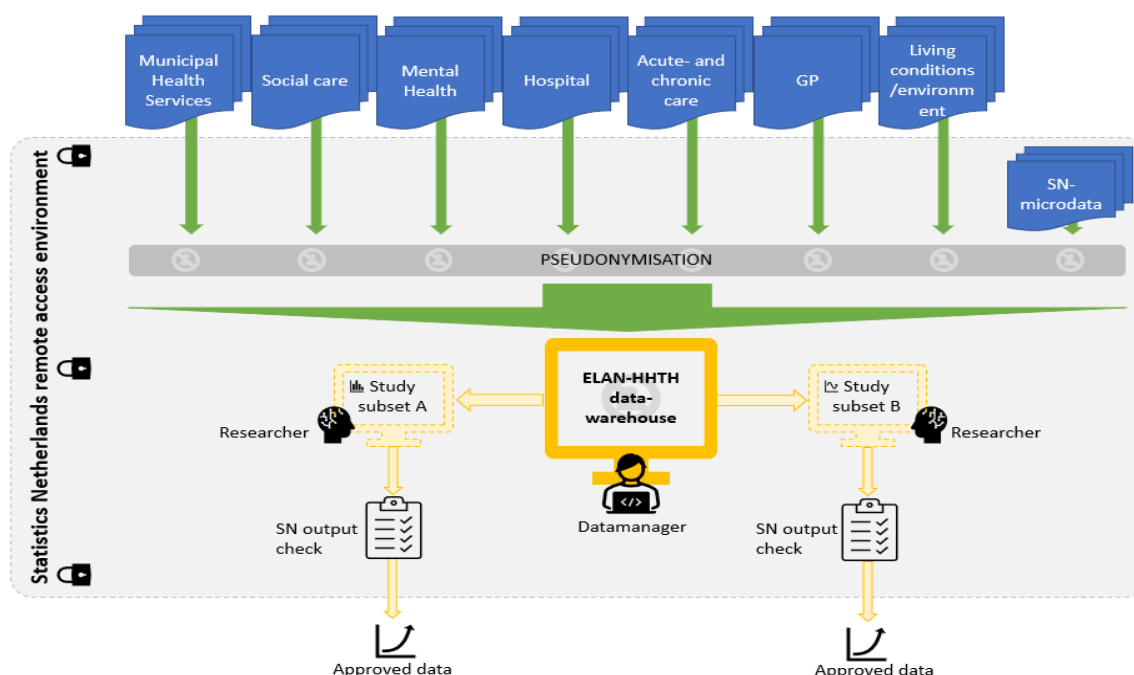
3. Infrastructure

3.1 Overview

Figure 1: Overview of the relationships between the different components in the ELAN infrastructure.

The infrastructure is designed in such a way that data can be transferred to the SN (Statistics Netherlands) Remote Access (RA) environment in a secure manner. The Electronic Health Record systems (EHRs) of the ELAN General Practitioners (GPs) participating in ELAN (https://www.lumc.nl/over-het-lumc/partners/partners-in-de-zorg/extramuraal-lumc-academisch-netwerk-elan/) are the primary care source systems for first-line data. The patient records are automatically extracted from the EHRs and transferred to a central STIZON database, from where the patient records are sent to the CBS.

Patient records from participating hospitals (HMC and Haga), Hadoks (Acute and Chronic Care), the municipality of The Hague (WMO), GGD (COVID-19), and Parnassia Groep (mental healthcare) are sent directly to the SN without the intervention of a third party. See the following link for a full description of the upload procedure: https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/microdata-zelf-onderzoek-doen.

Incoming data is pseudonymized ("de-identified") by SN and made available within a secure environment to ELAN data managers. If necessary, data managers clean up data and link it to SN-microdata using linkable data at the personal, company, and address level. These linked files result in a project database from which subsets can be made available to researchers for specific studies. The source files made available to SN for pseudonymization are pseudonymized within two months of delivery. CBS then deletes the dataset provided within four weeks after encryption.

After approval of the research proposal, the ELAN data manager creates a subset of data proportional to the research question and makes it available to authorized researchers after they have signed a data access agreement (see appendix 2). Researchers involved in the research

proposal analyze the data and prepare a report based on this, such as a scientific publication or an information product (e.g., management dashboard).

Once data is pseudonymized by SN, it can never leave the secure RA environment in its original or pseudonymized form. Researchers can ask SN to release their statistical analysis results so that researchers can take their (interim) results outside the secure RA environment for inclusion in a publication or information product. SN checks whether the results do not contain any disclosure risks and are no longer attributable to a person, company, or institution. Results are always based on aggregated data that is no longer attributable to individual persons, companies, or institutions. See the following link for a complete description of the data export procedure: https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/microdata-zelf-onderzoek-doen/export-van-gegevens.

3.2. Security

The ELAN data research project aims to ensure maximum safety of the data. To achieve this goal, all storage locations and transfer routes are adequately secured.

1. The SN is certified with ISO 9001:2015 and ISO 27001.

2. To access the data, a VPN connection must be established between the user and the SN-RA environment. Additionally, the SN utilizes two-factor authentication for logging in.

3. When datasets are released on the SN platform, only authorized individuals are granted access to the data in the SN "remote access environment."

3.3 Accessibility

The entire project database is only accessible to the ELAN data managers (n>=2). After requesting (and approval of) a subset of data for research purposes (see paragraph 4.2), it is made available to the researcher in the SN-RA environment. When datasets are released, only authorized individuals are granted access to the data they have requested. Authorization and confidentiality must be provided in two forms before researchers can access a subset:

Researchers and their supervisors sign a confidentiality agreement to access SN-microdata. The rules and measures used by the SN are described in the following link: https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/microdata-zelf-onderzoek-doen/regels-en-maatregelen. SN stores the confidentiality agreements of all researchers within their own infrastructure.

Researchers sign a Data Access Agreement (DAA) to access ELAN data from other suppliers than SN. See Appendix 1 for the text of this document. All DAA's from researchers are stored on a secure drive accessible only to the ELAN data manager and project leader.

3.4 Responsibility

The following responsibilities can be distinguished with regard to ELAN data:

The control over the data of the parties involved in the ELAN data research project remains with the supplying parties.

The supplying parties retain control over the use of the data they provide for research purposes.

The supplying parties are the data controllers for the dataset until the dataset is sent to SN. SN is the data controller from the moment the dataset is sent until the dataset is destroyed. From the destruction onwards, the supplying parties and SN are jointly responsible for processing the dataset. The administrative responsibility for building the data infrastructure lies with the board of directors of LUMC and the program board of the Health Campus The Hague.

3.5 Communication

The supplying parties are obliged to inform the individuals concerned about the purpose of ELAN and the use of their data for that purpose in accordance with applicable laws and regulations. They must provide the opportunity to object to the disclosure of their data.

The supplying parties can include a reference to ELAN on their website. Questions and complaints are collected and forwarded to the relevant persons, and the response time is monitored. A brief overview of the complaints and their handling is made annually for the program board of Health Campus The Hague.

4. Data

4.1 Pseudonymization

Data is encrypted through one-way pseudonymization. This means that all identifiable data is removed, and pseudonymized data cannot be de-pseudonymized. This is done through the following subprocesses:

1. Removal of personal identifying data such as Citizen Service Numbers (BSN in Dutch);
2. Removal of address data;
3. Removal of unique numbers, such as internal IDs and patient numbers, that may be traceable to individuals.

Data is pseudonymized after it is uploaded to SN, and before it is made available to researchers. Pseudonymization can be completed in two ways:

a) Pseudonymization of individuals is based on BSN or key pair (e.g. postcode, date of birth, gender, house number, and year of validity of the postcode).

b) Pseudonymization of addresses is based on the complete address, postcode, house number, house letter, house number addition, and year of validity of the address. The result of pseudonymization is a so-called Record Identification Number (RIN) that serves as a unique ID within the SN-environment. RINs can be used to link data within the SN-environment. This applies to both SN-microdata or third-party data.

4.2 Data requests and issuance

When researchers want to request access to data from ELAN, they must follow the following steps:

1. SN states that researchers can access the data infrastructure when the relevant knowledge institute where the researcher works has an agreement with SN. This does not mean that researchers from, for example, hospitals, where the hospital has no agreement with SN, cannot gain access. These

researchers can connect to LUMC by, for example, signing a zero-hours or internship contract. Therefore, researchers from LUMC can always access ELAN. For researchers from other institutes, it is always important to check if there is an agreement.

2. Ensure that your research protocol has been reviewed by a scientific committee and the Medical Ethical Committee (METC).

3. If the scientific committee and the METC approve, you can submit a ticket through https://www.elanresearch.nl/.

4. If access needs to be requested from external data managers of the data sources, you need to initiate that yourself (but only after ELAN have approved your research).

5. If ELAN approves, the data manager will start the application process by sending a form to the researcher requesting personal information. After this form is returned to the data manager, it takes about two weeks for the researcher to have access to a project within the SN data infrastructure. The researchers' completion forms are stored on a secure disk accessible only to the data manager and the ELAN project leader. Written consent for external data sources should be sent to the ELAN data manager The written consent and the METC application will be reviewed by both the data manager and the ELAN project leader for completeness (4-eyes principle).

6. After obtaining and verifying the written consent, the researcher will receive a Data Access Agreement (DAA) (see attachment 1) from ELAN and a SN completion form. The DAA sets clear conditions for the use of ELAN data. The SN completion form is necessary to access the SN RA-environment. Both documents must be completed and returned to the sender.

7. While the researcher is being registered with SN, the requested subset is compiled by an ELAN data manager. The actual issuance to the researcher occurs within the SN RA environment, where the researcher can perform their analyses.

4.3 Issuance

The issuance of the requested dataset takes the form of one or more encrypted files. From the moment the dataset is released to the researcher within the SN RA environment, the researcher is responsible for the new created datasets. See appendix 2 for a more detailed description of the subset issuance procedure for a funded research.

4.4 Publications and information products

Aggregated results of statistical analyses can be released for use in publications by the SN. The procedure guarantees that data that can be traced back to individuals, households, institutions or companies cannot be used outside the RA environment. For a complete description of the procedure for exporting the data, see the following link: https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/microdata-zelf-onderzoek-doen.

4.5 Archiving

As described in Appendix 2, the original dataset issued by the ELAN data manager is archived. It is mandatory for the respective researcher, under the framework of Good Research Practice (GRP), to save the analysis steps within the research. This means that all SPSS syntaxes, R scripts, or similar

files must be saved so that all steps within the analyses (including recodings and exclusions) can be reproducible by third parties. Data and syntaxes used for research must be kept for legal reasons for fifteen years. This archiving takes place within the secure SN-RA environment.

5. References

1. https://www.cbs.nl/-/media/7ecc662b3d604685a49d00d7d4d42cc7.ashx
2. https://www.cbs.nl/-/media/imported/onze%20diensten/methoden/gevalideerde%20methoden/throughput/documents/2010/18/2010-x37-13-pub.pdf
3. https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/microdata-zelf-onderzoek-doen/import-van-eigen-bestanden
4. https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/microdata-zelf-onderzoek-doen/export-van-gegevens
5. https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/microdata-zelf-onderzoek-doen/regels-en-maatregelen
6. https://www.stizon.nl/data-governance/

6. Appendices

Appendix 1: Data Access Agreement

Data Access Agreement ELAN-HHTH data-infrastructure

Prof. M.E. Numans
Associate Prof. J.N. Struijs
Department of Public Health and Primary Care
Leiden University Medical Center, Leiden, the Netherlands

Conditions

1. The ELAN-HHTH (Happy and Healthy The Hague, in Dutch: ELAN) data manager will provide the researcher with a subset of data within the Remote Access (RA) environment of Statistics Netherlands (Dutch abbreviation: CBS) that conforms to the query specified in the researchers data application.
2. The researcher will not provide or give access these data to another researcher, will not access data from another researcher, and will comply the CBS guidelines concerning data safety.
3. The researcher is familiar with the Dutch GDPR (General Data Regulation Regulation) and will act in the spirit of these guidelines.
4. The researcher will not disclose any information received in confidentiality related to the research project. The same applies to information put at the researchers disposal of which can be expected to understand the confidential character.
5. The Researcher is required to provide annual updates of the progress on the project via the general email of ELAN-HHTH (ELAN@lumc.nl).
6. Before submitting a paper that contains ELAN-HHTH data, the researcher has to send a concept of the paper, and after publication, a final version of the paper to ELAN@lumc.nl.
7. When analyses have been completed, the researcher is responsible for the reproducibility of the results as presented in the publication by delivering syntaxes and datasets to the ELAN-HHTH by informing ELAN-HHTH via ELAN@lumc.nl.
8. The researcher is responsible for financing the costs incurred within ELAN-HHTH.
9. The supervisor will become responsible for the above mentioned conditions in case the researcher is not able to finish the research project.

By signing this agreement form, the ELAN-HHTH representative and the LUMC give their permission to access the data-infrastructure of ELAN-HHTH according to the above mentioned conditions.

Researcher:

Name: Signature researcher:
Institute:
E-Mail:

Supervisor researcher

Name: Signature supervisor:
Institute:
E-Mail:

Appendix 2: Step-by-Step Plan to get access to the ELAN-data

1. Before the research data can be used for research purposes, the project applicant must have met the following conditions:

- There is a METC approval (WMO, or non-WMO declaration of no objection).

- The project applicant has signed and submitted the application form and the ELAN data access agreement (DAA).

- The project applicant has had an introductory conversation with the ELAN data manager and project leader.

- The required consent forms from the different data providers are available (when applicable).

- The project applicant must have authorization via SN to be able to log in to the SN Remote Access (RA) environment.

2. The ELAN data manager creates a dataset specific to the research and archives a copy of this original dataset. The format of these datasets will always be SPSS.

3. In the digital ELAN logbook (i.e. Access database for logistics), the following is noted:

   - Based on which SN microdata extraction the research dataset was created.

   - Research name

   - Location on the network where data is placed

   - Date of data issuance

4. The ELAN data manager informs the project applicant that the dataset is ready. The ELAN data manager stores the issued research dataset in the SN RA environment.

5. The project applicant is obligated under Good Research Practice (GRP) to save the analysis steps within the research. This means that all SPSS syntaxes, R scripts, or similar files must be saved so that all steps within the analyses (including recodings and exclusions) can be reproduced by third parties. As the project applicant, make sure to provide a clear description of the scripts (e.g. script order, start and end dataset name).

6. At the end of the research project, when all research results are reported/published, the project applicant submits all analysis scripts to the ELAN data manager. The project applicant checks if the analysis scripts generate any error messages. If this is the case, it is the responsibility of the project applicant to ensure that the scripts are modified. The ELAN data manager saves the analysis scripts with the original released research dataset, so that the research results can be reproduced at all times.

7. The project applicant deletes all data files from the project folder once their research is completed.

8. The ELAN data manager notes in the digital logbook that the research is completed, the analysis scripts have been received.

9. 15 years (the standard retention period for research data) after completing the study, the research dataset is also removed from the SN disk.