# Plan Overview

*A Data Management Plan created using DMPonline*

**Title:** Copy of ECePS ERA Chair in e-Governance and Digital Public Services

**Creator:**Hector Charles Pagan

**Affiliation:** University of Tartu

**Funder:** European Commission

**Template:** Horizon 2020 Template

**Project abstract:**

The aim of the ECePS project – the ERA Chair in E-Governance and Digital Public Services is to strengthen the Center of IT Impact Studies (CITIS), a research unit within the Skytte Institute of Political Studies at the University of Tartu (UTARTU) so that it can act as a world leader in the field of research on e-governance, public e-services and data driven public innovation. We will do so by recruiting a leading expert in the field to serve as an ERA Chair for E-Governance and Digital Public Services who will in turn create a top-level research team capable of conducting cutting edge research that examines the fundamental questions of scientific and practical importance. The ERA Chair will trigger structural changes within UTARTU to support this effort by: • Initiating changes to CITIS research unit, including creation of a CITIS Supervisory Board, formation of advisory groups with members from industry, government and scientific community, and a Professorship position for the ERA Chair. • Integrating researchers from other departments relevant for e-governance research into the CITIS structure, including the SoBigData Research Infrastructure, the Institute of Social Studies, the School of Economics and the Faculty of Law as well as UTARTU's High Performance Computing Center. • Building partnerships with governments and leading technology companies to create new models for attracting public and private research funding. The ERA Chair will also act as a role model to produce spill-over benefits for UTARTU to modernize rules and practices regarding the recruitment and performance measurement of researchers and professors, systematically implement processes to address RRI priorities and improve UTARTU's gender policies and practices.

**ID:** 113130

**Start date:** 01-09-2019

**End date:** 31-08-2024

**Last modified:** 08-12-2022

**Grant number / URL:** 857622

the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

# Copy of ECePS ERA Chair in e-Governance and Digital Public Services - Initial DMP

## 1. Data summary

**Provide a summary of the data addressing the following issues:**

- **State the purpose of the data collection/generation**
- **Explain the relation to the objectives of the project**
- **Specify the types and formats of data generated/collected**
- **Specify if existing data is being re-used (if any)**
- **Specify the origin of the data**
- **State the expected size of the data (if known)**
- **Outline the data utility: to whom will it be useful**

**The purpose of the data collection/generation**

The purpose of data collection and generation is to develop new service prototypes as well as to analyse the impact and efficiency of new digital public services and to publish the results in high impact journals.

**The relation to the objectives of the project**

The current data management plan sets the data handling regulations for developing the research team on e-governance, public e-services and data-driven public innovation within the ECePS project. It will assist developing research on three fundamental questions of scientific and practical importance within the project:

- How can researchers and practitioners best utilize the vast amounts of data that is already being generated throughout the world (via existing public e-services)?
- How to harness the benefits of digital transformation of government while minimizing the associated risks and ensuring security, privacy and equal access?
- How can obstacles that prevent governments from implementing e-governance systems and solutions be overcome?

**The types and formats of data generated/collected**

The main data that will be collected and used within ECePS project, is:

- Anonymized log data on e-service usage in Estonia annually updated with the cooperation of data donors.
- More detailed usage data will be requested from specific e-service owners/providers in Estonia on a case by case basis when required to prototype 3rd generation e-services ideas.
- The ERA Chair staff will collect anonymized individual level survey data on service usage to better understand motivations behind uptake and usage depth.
- ERA Char staff will access registry data as part of applied research projects, this data will be accessed and analyzed through Statistics Estonia secure researcher data access workstations and will not be separately collected nor stored by the project participants.

Specify if existing data is being re-used (if any)

The ECePS project research team will use:

- anonymized log data from the X-road platform
- publicly available aggregated anonymized X-road log data accessible from State Information System Authority open data portal.
- anonymized survey data on internet voting usage in Estonia

**The origin of the data**

The service log data used in ECePS research will be collected via the X-road platform, which includes digital service call and reply queries generated by the usage of services. This data is held by the State Information System authority which makes it accessible for academic study based on Estonian Data Protection Inspectorate Permit. Registry data used by the project team through Statistics Estonia secure access workstations is generated by capturing data as part of state institutions providing given services to citizens. Anonymized survey data on service usage and attitudes is separately collected by survey organizations through interview processes in line with ethical guidelines on survey data collection, storage and usage.

**The expected size of the data**

As the specific research focus of the project can only be set after the ERA Chair has been recruited, the expected size of the data is currently not yet known.

**Outline the data utility: to whom will it be useful**

The aggregated and analysed summaries of the data will be useful for researchers, experts, state institutions and private companies working with e-governance and digital public services.

The data and resulting scientific findings would also be beneficial to public bodies implementing e-governance and digital public services.

## 2. FAIR data

### 2.1 Making data findable, including provisions for metadata:

- **Outline the discoverability of data (metadata provision)**
- **Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?**
- **Outline naming conventions used**
- **Outline the approach towards search keyword**
- **Outline the approach for clear versioning**
- **Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how**

**Discoverability of data (metadata provision)**

The metadata is stored in a searchable text format outlining dataset names, dates, formats, version and variable description, usage conditions and downloading options.

There will be a dedicated sub-page on the ECePS homepage where the overview of public metadata is provided and upon request the data will be securely distributed to other researchers.

**Identifiability of data and standard identification mechanism, such as persistent and unique identifiers (e.g. Digital Object Identifiers)**

UTARTU DSpace creates DOIs via DataCite for newly deposited research outputs.

**Naming conventions used**

All files will be named uniformly when storing them for public use, based upon the following criteria:

- No special characters such as "/ \ : * ? " < > [ ] & $ will be used in names.
- Underscores (_) will be used to separate terms not spaces
- Names will be 30 characters or less in length
- Names can specify the month and year of creation (MM-YYYY) at the end of the name.
- Names will be descriptive of what information they contain - so that they are understandable to someone who is unfamiliar with the research.

Some information could be described in metadata including the following:

- Project acronym
- Researcher name/initials
- Date or date range of the analysis
- Type of data
- Conditions
- Version number of the file
- Three-letter file extension for application-specific files

**Use of search keywords.**

Software library will provide basic discovery metadata online (title, author, subjects, keywords, department, etc.).

**Approach for clear versioning**

- Versions of files will only be stored and made available when relevant
- Whenever possible, obsolete versions will be discarded or deleted (while retaining the original 'raw' copy)
- Files with multiple versions will include the letter V and the number of the version before the date - for example V1 V2 V3, etc.

**Standards for metadata creation**

Survey data metadata will follow the Data Documentation Initiative (DDI) standard. Service log data metadata coming from Estonian e-governance systems which can be made available will also be created according to DDI standards as it contains log data on behavior in the aggregate.

**2.2 Making data openly accessible:**

- **Specify which data will be made openly available? If some data is kept closed provide rationale for doing so**
- **Specify how the data will be made available**
- **Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?**
- **Specify where the data and associated metadata, documentation and code are deposited**
- **Specify how access will be provided in case there are any restrictions**

**Data to be made openly available and to be kept closed.**

When possible, data will be available on Estonia’s Open Data Portal (opendata.riik.ee). All metadata will be made openly available via UTARTU DSpace and ECePS website sub-page.

Data that includes sensitive and personal information will not be displayed on the Open Data Portal, it will only be used for research purposes and will be stored in an encrypted form on UTARTU servers. Any data that will be kept closed will be done so either for intellectual property protection reasons or to protect confidentiality issues.

The underlying data will be made available for other researchers when permitted by the providers of the data, i.e. Estonian government. Analytical methods will be transparent and open, enabling other researchers to test and validate our findings and help increase acceptance within the wider scientific community.

**Form of data availability**

All the aggregated data will be available via DSpace and ECePS website sub-page.

**Methods/software tools needed to access the data. Required documentation and inclusion of the relevant software (e.g. in open source code)**

Metadata will be available in conventional DDI format; individual level data will be accessible in .CSV and .JSON formats. Documentation will not be required.

**Location of deposited data, associated metadata, documentation and code**

The data and associated metadata, documentation and code are accessible through DSpace and the ECePS website. Big data will be linked to hosting by the University of Tartu’s High Performance Computing Center.

**Access (restrictions) to data**

Restricted data is accessible via password protection and a data request through ECePS website sub-page. The access to the restricted data will be granted based upon data owners’ reasoned decision.

**2.3 Making data interoperable:**

- **Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.**
- **Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?**

**Interoperability of data, including data and metadata vocabularies, standards or methodologies to facilitate interoperability.**

The metadata is stored at first on the University of Tartu personal server and after data publication in DSpace.

Whenever possible interoperable file formats will be used, such as .CSV.

**Use of standard vocabulary for data to allow inter-disciplinary interoperability (including mapping to more commonly used ontologies)**

Wherever possible, standard vocabulary will be used for data sets. No mapping to more commonly used ontologies will be offered.

**2.4 Increase data re-use (through clarifying licenses):**

- **Specify how the data will be licenced to permit the widest reuse possible**
- **Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed**
- **Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why**
- **Describe data quality assurance processes**
- **Specify the length of time for which the data will remain re-usable**

Question not answered.

## 3. Allocation of resources

**Explain the allocation of resources, addressing the following issues:**

- **Estimate the costs for making your data FAIR. Describe how you intend to cover these costs**
- **Clearly identify responsibilities for data management in your project**
- **Describe costs and potential value of long term preservation**

**The costs for making our data FAIR.**

Free of cost.

**Responsibilities for data management in the ECePS project**

Data Management is handled by ECePS researchers during data collection and analysis, and by staff of UTARTU during preservation. The same staff will be responsible for producing standard conform metadata and storing and archiving the relevant datasets. Costs incurred for this have been budgeted and will be covered by project funding. Hosting metadata and data storage is ensured through DSpace which operates as a University of Tartu data and research object archive.

**Costs and potential value of long-term preservation**

No costs for long-term preservation are foreseen. The potential value of long-term preservation lies mostly in log and survey data which will allow the analysis of long-term changes in behavior and attitudes. Costs for long term storage are covered by archiving data in DSpace which is separately funded by University of Tartu research infrastructure funding.

It is anticipated that the size of the databases will continue to rise as information is collected either within Estonia by UTARTU or from other populations. Due to its potential uses in the public sector, UTARTU has additional incentives to maintain data after the official end-date of the ECePS project as it is anticipated that the ERA Chair and their team will continue their work.

## 4. Data security

### Address data recovery as well as secure storage and transfer of sensitive data

**Data recovery, secure storage and transfer of sensitive data**

The data as well as the research carried out by ECePS researchers will follow Estonian data protection regulations and relevant cybersecurity rules, such as all computers must have appropriate and up-to-date anti-virus and anti-spyware software. The ECePS team will work with IT specialists at UTARTU to ensure that the best possible secure data handling methods are used.

Sensitive data or data, which has been embargoed by the data donor, will be stored and archived in secure data storage facilities run by UTARTU High Performance Computing Centre.

## 5. Ethical aspects

### To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former

No research involving human participants is done.

## 6. Other

### Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

ECePS ERA Chair has dedicated Data protection protocols in place that add to the Data Management Plan stated here.

# Copy of ECePS ERA Chair in e-Governance and Digital Public Services - Detailed DMP

## 1. Data summary

### State the purpose of the data collection/generation

### The purpose of the data collection/generation

The purpose of data collection and generation is to develop new service prototypes as well as to analyse the impact and efficiency of new digital public services and to publish the results in high impact journals.

### Explain the relation to the objectives of the project

### The relation to the objectives of the project

The current data management plan sets the data handling regulations for developing the research team on e-governance, public e-services and data-driven public innovation within the ECePS project. It will assist developing research on three fundamental questions of scientific and practical importance within the project:

- How can researchers and practitioners best utilize the vast amounts of data that is already being generated throughout the world (via existing public e-services)?
- How to harness the benefits of digital transformation of government while minimizing the associated risks and ensuring security, privacy and equal access?
- How can obstacles that prevent governments from implementing e-governance systems and solutions be overcome?

### Specify the types and formats of data generated/collected

### The types and formats of data generated/collected

The main data that will be collected and used within ECePS project, is:

- Anonymized log data on e-service usage in Estonia annually updated with the cooperation of data donors.
- More detailed usage data will be requested from specific e-service owners/providers in Estonia on a case by case basis when required to prototype 3rd generation e-services ideas.
- The ERA Chair staff will collect anonymized individual level survey data on service usage to better understand motivations behind uptake and usage depth.
- ERA Chair staff will access registry data as part of applied research projects, this data will be accessed and analyzed through Statistics Estonia secure researcher data access workstations and will not be separately collected nor stored by the project participants.

### Specify if existing data is being re-used (if any)

The ECePS project research team will re-use following data:

- anonymized log data from the X-road platform
- publicly available aggregated anonymized X-road log data accessible from State Information System Authority open data portal.
- anonymized survey data on internet voting usage in Estonia

**Specify the origin of the data**

**The origin of the data**

The service log data used in ECePS research will be collected via the X-road platform, which includes digital service call and reply queries generated by the usage of services. This data is held by the State Information System authority which makes it accessible for academic study based on Estonian Data Protection Inspectorate Permit. Registry data used by the project team through Statistics Estonia secure access workstations is generated by capturing data as part of state institutions providing given services to citizens. Anonymized survey data on service usage and attitudes is separately collected by survey organizations through interview processes in line with ethical guidelines on survey data collection, storage and usage.

**State the expected size of the data (if known)**

**The expected size of the data**

As the specific research focus of the project can only be set after the ERA Chair has been recruited, the expected size of the data is currently not yet known.

**Outline the data utility: to whom will it be useful**

**Outline the data utility: to whom will it be useful**

The aggregated and analysed summaries of the data will be useful for researchers, experts, state institutions and private companies working with e-governance and digital public services.

The data and resulting scientific findings would also be beneficial to public bodies implementing e-governance and digital public services.

## 2.1 Making data findable, including provisions for metadata [FAIR data]

**Outline the discoverability of data (metadata provision)**

**Discoverability of data (metadata provision)**

The metadata is stored in a searchable text format outlining dataset names, dates, formats, version and variable description, usage conditions and downloading options.

There will be a dedicated sub-page on the ECePS homepage where the overview of public metadata is provided and upon request the data will be securely distributed to other researchers.

**Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?**

**Identifiability of data and standard identification mechanism, such as persistent and unique identifiers (e.g. Digital Object Identifiers)**

UTARTU DSpace creates DOIs via DataCite for newly deposited research outputs.

**Outline naming conventions used**

**Naming conventions used**

All files will be named uniformly when storing them for public use, based upon the following criteria:

- No special characters such as "/ \ : * ? " < > [ ] & $ will be used in names.
- Underscores (_) will be used to separate terms not spaces
- Names will be 30 characters or less in length
- Names can specify the month and year of creation (MM-YYYY) at the end of the name.
- Names will be descriptive of what information they contain - so that they are understandable to someone who is unfamiliar with the research.

Some information could be described in metadata including the following:

- Project acronym
- Researcher name/initials
- Date or date range of the analysis
- Type of data
- Conditions
- Version number of the file
- Three-letter file extension for application-specific files

**Outline the approach towards search keyword**

**Use of search keywords.**

Library will provide basic discovery of the metadata online (title, author, subjects, keywords, department, etc.).

**Outline the approach for clear versioning**

**Approach for clear versioning**

- Versions of files will only be stored and made available when relevant
- Whenever possible, obsolete versions will be discarded or deleted (while retaining the original 'raw' copy)

Files with multiple versions will include the letter V and the number of the version before the date - for example V1 V2 V3, etc

**Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how**

**Standards for metadata creation**

Survey data metadata will follow the Data Documentation Initiative (DDI) standard. Service log data metadata coming from Estonian e-governance systems which can be made available to interested researchers will also be created according to DDI standards as it contains log data on behavior in the aggregate.

## 2.2 Making data openly accessible [FAIR data]

**Specify which data will be made openly available? If some data is kept closed provide rationale for doing so**

**Data to be made openly available and to be kept closed.**

When possible, data will be available on Estonia’s Open Data Portal (opendata.riik.ee). All metadata will be made openly available via UTARTU DSpace and ECePS website sub-page.

Data that includes sensitive and personal information will not be displayed on the Open Data Portal, it will only be used for research purposes and will be stored in an encrypted form on UTARTU servers. Any data that will be kept closed will be done so either for intellectual property protection reasons or to protect confidentiality issues.

The underlying data will be made available for other researchers when permitted by the providers of the data, i.e. Estonian government. Analytical methods will be transparent and open, enabling other researchers to test and validate our findings and help increase acceptance within the wider scientific community.

**Specify how the data will be made available**

**Form of data availability**

All the aggregated data will be available via DSpace and ECePS website sub-page.

**Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?**

**Methods/software tools needed to access the data. Required documentation and inclusion of the relevant software (e.g. in open source code)**

Metadata will be available in conventional DDI format; individual level data will be accessible in .CSV and .JSON formats. Documentation and descriptions on how to access data will be provided with the

data.

**Specify where the data and associated metadata, documentation and code are deposited**

**Location of deposited data, associated metadata, documentation and code**

The data and associated metadata, documentation and code are accessible through DSpace and the ECePS website. Big data will be linked to hosting by the University of Tartu's High Performance Computing Center.

**Specify how access will be provided in case there are any restrictions**

**Access (restrictions) to data**

Restricted data is accessible via password protection and a data request through ECePS website sub-page. The access to the restricted data will be granted based upon data owners' reasoned decision.

## 2.3 Making data interoperable [FAIR data]

**Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.**

**Interoperability of data, including data and metadata vocabularies, standards or methodologies to facilitate interoperability.**

The metadata is stored at first on the University of Tartu personal server and after data publication in DSpace.

Whenever possible interoperable file formats will be used, such as .CSV.

**Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?**

**Use of standard vocabulary for data to allow inter-disciplinary interoperability (including mapping to more commonly used ontologies)**

Wherever possible, standard vocabulary will be used for data sets. No mapping to more commonly used ontologies will be offered.

## 2.4 Increase data re-use (through clarifying licenses) [FAIR data]

**Specify how the data will be licenced to permit the widest reuse possible**

**Licensing the data to permit the widest reuse possible**

Data collected as part of the project will be made available through a Creative Commons license, specifically Attribution + Noncommercial + ShareAlike (BY-NC-SA) license version. This will allow free usage for research purposes given that original attribution of the data is given. For the sake of tracking data usage and help to monitor project dissemination data requests will have to provide a short description of the intended usage and acknowledge the BY-NC-SA license.

**Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed**

**Making data available for re-use. If applicable, specify why and for what period a data embargo is needed**

There will be a dedicated sub-page on the ECePS homepage where the overview of public metadata is provided and upon request the data will be securely distributed to other researchers. Data will be embargoed prior to publication of related research by the project team in the international peer-reviewed journals and upon data donor request.  In case of non-infringement with the publication plan of the project team otherwise embargoed data can be made accessible on a case-by-case basis to third party researchers.

**Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why**

**When is the data produced and/or used in the project useable by third parties, in particular after the end of the project. Justifying restriction on re-use of some data.**

As ECePS project incorporates Sustainability plan, the continuation of research is foreseen after the project ends. Thus, the data will be available through aforementioned systems and stored on UT servers. All data subject to temporary embargoes during the duration of the project will be made freely available after the end of the project, except when data donors or data access agreements explicitly prohibit sharing with third party researchers. The latter might arise when the project uses some system log data that is not open access or data that has been granted explicitly for a particularly research project by the data owner, relevant data protection authority or research ethics committee.

**Describe data quality assurance processes**

**Data quality assurance processes**

ECePS research team will use data that has already gone through multiple quality testing stages within the X-Road platform.

Data quality of the X-Road environment is handled by Nordic Institute for Interoperability Solutions (NIIS). NIIS is responsible for testing the X-Road core that includes Central Server, Configuration Proxy, Security Server and external APIs and interfaces provided by the core components. Organizations operating their X-Road environments are responsible for testing any specific hardware and their own extensions built on top of the X-Road, e.g.: API catalog, environmental monitoring tools, operational monitoring tools etc.

Most of the tests are automated and there are three alternative execution schedules:

- when new code is submitted

- daily
- before releasing a new version.

New releases must pass all the categories and quality tests before they are published and distributed to NIIS members and X-Road community.

Survey data collection for the purpose of the project will be commissioned from opinion research companies adhering to MSR and ESOMAR professional codes and holding ISO certificates on international market research quality standards, quality management and personal data security.

Registry data requested for the purpose of the project will already have passed the quality insurance and data integrity checks by the relevant registry keeper. All research data will still be examined by experienced data analysts and researchers employed by the project to ensure the needed data quality for the project.

**Specify the length of time for which the data will remain re-usable**

**The length of time for which the data will remain re-usable**

The length of time the data will remain re-usable will depend on the exact content of the data, as some of it is prone to age faster as new data becomes available. For ongoing research projects aimed at piloting data driven digital services regular data updates will be requested as part of the project to ensure services would be deployable by the digital service process owners. The project team has also experience in analyzing technology diffusion patterns in log and survey data, for this type of research data will stay re-usable even after aging for up to date behavioral modeling. The length of time the data will stay usable between these two extremes will depend on the precise research question at hand. The team will make an effort in any case to keep all data sources updated as regularly as possible.

## 3. Allocation of resources

**Estimate the costs for making your data FAIR. Describe how you intend to cover these costs**

**The costs for making our data FAIR.**

There will be no foreseeable additional costs for making our data FAIR, as all the activities are already incorporated in the everyday work of the ECePS research team.

**Clearly identify responsibilities for data management in your project**

**Responsibilities for data management in the ECePS project**

Data Management is handled by ECePS researchers during data collection and analysis, and by staff of UTARTU during preservation. The same staff will be responsible for producing standard conform metadata and storing and archiving the relevant datasets. Costs incurred for this have been budgeted and will be covered by project funding. Hosting metadata and data storage is ensured through DSpace, which operates as a University of Tartu data and research object archive.

**Describe costs and potential value of long term preservation**

**Costs and potential value of long-term preservation**

No costs for long-term preservation are foreseen. The potential value of long-term preservation lies mostly in log and survey data which will allow the analysis of long-term changes in behavior and attitudes. Costs for long term storage are covered by archiving data in DSpace which is separately funded by University of Tartu research infrastructure funding.

It is anticipated that the size of the databases will continue to rise as information is collected either within Estonia by UTARTU or from other populations. Due to its potential uses in the public sector, UTARTU has additional incentives to maintain data after the official end-date of the ECePS project as it is anticipated that the ERA Chair and their team will continue their work.

## 4. Data security

**Address data recovery as well as secure storage and transfer of sensitive data**

**Data recovery, secure storage and transfer of sensitive data**

The data as well as the research carried out by ECePS researchers will follow Estonian data protection regulations and relevant cybersecurity rules, such as all computers must have appropriate and up-to-date anti-virus and anti-spyware software. The ECePS team will work with IT specialists at UTARTU to ensure that the best possible secure data handling methods are used.

Sensitive data or data, which has been embargoed by the data donor, will be stored and archived in secure data storage facilities run by UTARTU High Performance Computing Centre.

## 5. Ethical aspects

**To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former**

No research involving human participants will be done.

The ECePS ERA chair team will use anonymized X-road data, which does not violate the privacy of e-governance users.

Should there be any need to include non-anonymized data in the research, the ECePS Data Protection Protocols will be followed, which give overview of relevant legislative regulations in the EU, Estonia and University of Tartu concerning ethical aspects of different stages of research, among other things.

In addition, as also stated in the Data Protection Protocols, every time during research there is a need to process any type of personal data or there is possibility of identification of the anonymised personal information, then either the Research Ethics committee of the University of Tartu or the Ethics Committee of the Estonian Social Work Association will be consulted (depending on the research topic). Appropriate permission will be obtained before conducting any research if necessary to ensure the protection of fundamental rights and freedoms of natural persons.

All research in UT must also follow the Estonian Code of Conduct for Research Integrity, which states the ethical responsibilities of the researcher. The Code of Conduct can be accessed here: https://www.eetika.ee/en/ethics-estonia/estonian-code-conduct-research-integrity.

## 6. Other

### Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

### Other national/funder/sectorial/departmental procedures for data management that you are using

ECePS ERA Chair has dedicated Data Protection Protocols in place that complement current Data Management Plan.

# Copy of ECePS ERA Chair in e-Governance and Digital Public Services - Final review DMP

## 1. Data summary

### State the purpose of the data collection/generation

Question not answered.

### Explain the relation to the objectives of the project

Question not answered.

### Specify the types and formats of data generated/collected

Question not answered.

### Specify if existing data is being re-used (if any)

Question not answered.

### Specify the origin of the data

Question not answered.

### State the expected size of the data (if known)

Question not answered.

### Outline the data utility: to whom will it be useful

Question not answered.

## 2.1 Making data findable, including provisions for metadata [FAIR data]

**Outline the discoverability of data (metadata provision)**

Question not answered.

**Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?**

Question not answered.

**Outline naming conventions used**

Question not answered.

**Outline the approach towards search keyword**

Question not answered.

**Outline the approach for clear versioning**

Question not answered.

**Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how**

Question not answered.

## 2.2 Making data openly accessible [FAIR data]

**Specify which data will be made openly available? If some data is kept closed provide rationale for doing so**

Question not answered.

**Specify how the data will be made available**

Question not answered.

**Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?**

Question not answered.

**Specify where the data and associated metadata, documentation and code are deposited**

Question not answered.

**Specify how access will be provided in case there are any restrictions**

Question not answered.

## 2.3 Making data interoperable [FAIR data]

**Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.**

Question not answered.

**Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?**

Question not answered.

## 2.4 Increase data re-use (through clarifying licenses) [FAIR data]

**Specify how the data will be licenced to permit the widest reuse possible**

Question not answered.

**Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed**

Question not answered.

**Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why**

Question not answered.

**Describe data quality assurance processes**

Question not answered.

**Specify the length of time for which the data will remain re-usable**

Question not answered.

## 3. Allocation of resources

**Estimate the costs for making your data FAIR. Describe how you intend to cover these costs**

Question not answered.

**Clearly identify responsibilities for data management in your project**

Question not answered.

**Describe costs and potential value of long term preservation**

Question not answered.

## 4. Data security

**Address data recovery as well as secure storage and transfer of sensitive data**

Question not answered.

## 5. Ethical aspects

**To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former**

Question not answered.

## 6. Other

**Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)**

Question not answered.