



## Assessment Information

[CoreTrustSeal Requirements 2020–2022](#)

Repository:

APIS - Arquivo Português de Informação Social

Website:

<http://www.apis.ics.ulisboa.pt/en/>

Certification Date:

1 September 2021

This repository is owned by:

**Instituto de Ciências Sociais da Universidade de Lisboa  
(ICS-UL)**

**CoreTrustSeal Board**

W [www.coretrustseal.org](http://www.coretrustseal.org)

E [info@coretrustseal.org](mailto:info@coretrustseal.org)



# APIS - Arquivo Português de Informação Social

## Notes Before Completing the Application

*We have read and understood the notes concerning our application submission.*

True

### *Reviewer Entry*

#### **Reviewer 1**

Comments:  
Accept

#### **Reviewer 2**

Comments:  
Accept

## CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

### Background & General Guidance

### Glossary of Terms

## BACKGROUND INFORMATION

### Context

*R0. Please provide context for your repository.*

## ***Repository Type. Select all relevant types from:***

Domain or subject-based repository, Archive

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

Accept

#### **Reviewer 2**

Comments:

Accept

## ***Brief Description of Repository***

The Portuguese Archive of Social Information (APIS) is committed to preserving and giving access to digital research data (survey data in particular) for the purposes of secondary analysis, public consultation and pedagogical use.

The archive comprises a range of datasets provided by research projects of the national scientific community, mainly covering the following subjects: politics; health; social stratification and groupings; society and culture.

APIS is a social science data archive based at Instituto de Ciências Sociais, University of Lisbon (ICS-ULisboa).

APIS, together with the national node of the European Social Survey - ERIC, form the Production and Archive of Social Science Data (PASSDA) which is part of the Portuguese Roadmap of Research-Infrastructures, launched by the national funding agency FCT in 2015, being the main goal the production, analyses and archiving of data provided by national and international social sciences studies.

PASSDA was created as a partnership between research centres and units of the University of Lisbon, ISCTE-IUL, the University of Coimbra and the University of Porto.

ICS-ULisboa is the leader and host institution of PASSDA consortium, being the principal provider of PASSDA and, consequently, of APIS. Thus, APIS benefits of its support, in terms of human and technologic resources.

We should also quote from ICS-Ulisboa website:

“ICS has always been rated as an Institution of Excellence (the highest rating) by international FCT panels (1996, 1999, 2005, 2008, 2014 and 2019). These evaluations have provided a stimulus for innovative and interdisciplinary research, for high-quality teaching and for enhancing the social relevance of knowledge produced by ICS. Over fifty years, the driving force behind ICS has been the combination of researchers and the ideas they investigate with a scientific culture characterised by methodological and theoretical diversity.”.

Data within APIS is acquired through network activities, in close relationship with the research community and the depositors, mostly because there is no self-archiving system yet.

The archive staff enhances all datasets by editing the data itself, producing metadata according to standards such as DDI

and documenting it with other files such as method reports.

The data collection of APIS operates under an open access license (CC-BY, more details in R2. Licenses) and with no access costs for the end user. All data holdings are available through the online catalogue and can be analysed using Nesstar. Whenever the user wants to download data, s/he will be redirected to the Scientific Data Hosting Service owned by the Portuguese Foundation for National Scientific Computing (FCCN) as explained in the “outsource partners” section. In RCAAP there are different options available when depositing: open access, restricted access, embargo or closed. All end users are required to comply with the conditions of use of the data.

APIS website

<http://www.apis.ics.ulisboa.pt/en/>

PASSDA websites

<http://passda.pt/?lang=en>

<https://www.ics.ulisboa.pt/en/projeto/production-and-archive-social-science-data>

ICS-ULisboa website

<https://www.ics.ulisboa.pt/en>

Nesstar WebView page with APIS studies

<http://nesstar.ics.ul.pt/webview/>

Scientific Data Hosting Service

<http://projeto.rcaap.pt/index.php/lang-en/sobre-o-rcaap/servicos/sardc>

Portuguese Roadmap of Research Infrastructures

[https://www.fct.pt/apoios/equipamento/roteiro/2013/docs/Portuguese\\_Roadmap\\_of\\_Research\\_Infrastructures.pdf](https://www.fct.pt/apoios/equipamento/roteiro/2013/docs/Portuguese_Roadmap_of_Research_Infrastructures.pdf)

Foundation for Science and Technology

<http://www.fct.pt/>

APIS has outsource partners - as it will be better explained in (5) - and a whole research network. Please check the link to the website: <http://www.apis.ics.ulisboa.pt/en/organizational-infrastructure/>.

### *Reviewer Entry*

#### **Reviewer 1**

Comments:

Accept

#### **Reviewer 2**

Comments:

Accept

## ***Brief Description of the Repository's Designated Community.***

We may say that APIS designated community is mainly quantitative social science researchers from the fields of Sociology, Political Sciences and Social Psychology. We say “quantitative” as most of the data that we archive and publish is quantitative, rather than qualitative.

Hence, we can distinguish two components in the Designated Community of APIS: a) the depositors and b) the data users.

a) Researchers associated with the research centres of PASSDA and working in Social Sciences, as well as all academics developing projects in Social Sciences, producing mainly quantitative data (from representative surveys or micro data from experimental studies in Social Psychology field);

b) Data users, including not only researchers, but also PhD students, teachers of universities or even high school teachers and students interested in Social Sciences.

Our designated community is mainly from Portugal, as you may observe in the statistics as follows:

<https://dados.rcaap.pt/handle/10400.20/1/statistics?locale=en>

We can document APIS recognition in the designated community with the link to the statistics of our collection in RCAAP:

<http://dados.rcaap.pt/handle/10400.20/1/statistics>.

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

Accept

#### **Reviewer 2**

Comments:

Accept

## ***Level of Curation Performed. Select all relevant types from:***

B. Basic curation – e.g. brief checking; addition of basic metadata or documentation, C. Enhanced curation – e.g. conversion to new formats; enhancement of documentation, D. Data-level curation – as in C above; but with additional editing of deposited data for accuracy

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

Accept

#### **Reviewer 2**

Comments:

Accept

## **Comments**

Different levels of curation are applied to different datasets, depending on the data quality of the material that APIS does receive. To ensure data integrity, we store all incoming content in a Submission Information Package (SIP). The SIP contains data files, metadata information and additional documentation, e.g. the deposit form and the study report. The datasets, metadata and documentation are reviewed by the data manager, which ensures the anonymization of data. Enhanced curation (level of curation C) includes the conversion to new formats and quality control checks (please see R8. Appraisal). Documentation is compared to the data to ensure that everything is accurate. A codebook (in pdf.) is generated for the depositor and also for dissemination. Formats like .csv and .xml files are kept in AIP. In some cases, we have level of curation D - data-level curation, by altering the data (in accordance with the signed licence) to publish it in our catalogue and RCAAP.

### **Reviewer Entry**

#### **Reviewer 1**

Comments:  
Accept

#### **Reviewer 2**

Comments:  
Accept

## ***Insource/Outsource Partners. If applicable, please list them.***

APIS, through ICS-ULisboa, has a contractual relationship with the Foundation for National Scientific Computing (FCCN), which is the national authority for scientific computing. FCCN is a unit within the Foundation for Science and Technology (FCT) - the national funding agency and the body that coordinates the national scientific system - and it is responsible for managing the National Research and Education Network.

The contract with FCCN was signed in the framework of RCAAP (Scientific Open Access Repositories of Portugal, in Portuguese, “Repositórios Científicos de Acesso Aberto de Portugal”) which is the key infrastructure of the national scientific and technological system.

The mission of RCAAP is “to promote, support and ease the adoption of open access to scientific knowledge in Portugal and to store, make available and preserve the scientific production” (RCAAP, 2016). The project widened its activity to research data in 2010 by creating the Scientific Data Hosting Service (SARDC).

While RCAAP is a national repository for general data (any type of data can be deposited there), SARDC is a service that makes part of RCAAP but specifically for research data, in which is included APIS research data:

<https://dados.rcaap.pt/handle/10400.20/2>.

For instance, in OAIS terms (and better explained in R12), submission and ingest are prepared by the archive; then we deposit the data in RCAAP/SARDC (in Portuguese “Serviço de Alojamento de Repositórios de Dados Científicos”; in

english “Scientific Data Hosting Service”, as it is designated in our website workflow - “SDHS”), a repository including other repositories aggregation facilities. This service intends to make available an open access platform to the scientific data produced at national level.

Archival storage is granted by APIS server (in ULisboa) and the platform DSpace in RCAAP/SARDC (better explained in R15) which gives support to all data maintenance components, such as backups, updates, security, etc.

Dissemination is assured in APIS catalogue (also discovery), RCAAP website (discovery and safeguarded data download), Nesstar webview (online analysis), all of them platforms where users can access data in different ways.

The contract with FCCN was signed in 2012 and is renewed on an annual basis, a fairly common way to maintain a contract in Portugal, as in general the contracts are renewed annually and automatically. The only thing that could stop the renewal could be, in a hypothetical situation, the closure of RCAAP and consequently, SARDC.

For a better analysis of the roles of each partner we provide the following information:

#### FCCN

- One of the most important obligations in terms of APIS’ sustainability is to “store and make available online the scientific data repository of the entity member through an operational infrastructure which guarantees the availability of the service”. For this purpose, FCCN has to install and keep updated the necessary operating systems and other core infrastructural hardware and software technologies appropriate for the storage of the data, ensuring access to data in the long-term.
- FCCN offers technological security - through the installation, update and renewal of the security certificates - and keeps pace with technological developments, providing the protection of the facility and its data, products, services, and users.

#### RCAAP (which includes SARDC)

- The service SARDC aims to make accessible the deposited research data in the repository via the RCAAP portal.  
<http://projeto.rcaap.pt/index.php/lang-pt/sobre-o-rcaap/servicos/sardc>
- Specifically, Scientific Data Hosting Service (SARDC) allows:
  - the download of research data;
  - the long-term storage of research data;
  - the attribution of persistent identifiers for each dataset;
  - backup copies;
  - download records;
  - indexing, discovery and retrieval of the research data on the RCAAP portal.

#### APIS

- Depositing data in the repository (RCAAP/SARDC on a regular basis, i.e. around 20/30 datasets (plus the additional materials, such as questionnaires) per year.
- Managing the pages in the repository, in what concerns the online catalogue on APIS website, but also information about

APIS and all the datasets in RCAAP.

- APIS has the obligation to respect the instructions of FCCN in the RCAAP project, for instance APIS needs to ensure the good practices in what refers to (meta)data which enable interoperability.
- It is APIS' responsibility to guarantee that all the datasets, which are deposited in the repository, respect the legislation concerning intellectual property of data and GDPR (data protection).
- Finally, APIS should promote the open access of the data and its dissemination within the designated community and other relevant public.

■ Important websites regarding APIS partners:

FCCN website

<https://www.fccn.pt/en/>

RCAAP website

<https://www.rcaap.pt/>

Relation between FCNN and RCAAP

<https://fccn.pt/en/knowledge/rcaap/>

Scientific Data Hosting Service (SARDC)

<http://projeto.rcaap.pt/index.php/lang-en/sobre-o-rcaap/servicos/sardc>

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Accept

##### **Reviewer 2**

Comments:  
Accept

### ***Summary of Significant Changes Since Last Application (if applicable).***

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Not applicable

##### **Reviewer 2**

Comments:  
N.A.



## ***Other Relevant Information.***

We have already mentioned the integration of APIS in PASSDA, but we can again refer to it as a national network of research centres and units from different Portuguese universities or national research infrastructures.

In the final FCT evaluation report of Research Infrastructures in Portugal 2019, PASSDA got the highest score.

Due to future procedures and regulations, we have made an effort to register APIS in re3data.org. It was successfully achieved and the entry is now online: <https://www.re3data.org/repository/r3d100013420>.

APIS submits in RCAAP approximately 20/30 datasets every year. In the next years, our goal is getting this number of uploads higher. Concerning the usage and impact of the repository data holdings, each "handle" or landing page has a link to a statistical page (e.g.: <https://dados.rcaap.pt/handle/10400.20/2080/statistics>).

To confirm the reliability and visibility of APIS in the designated community we share the link of RCAAP (Repositórios Científicos de Acesso Aberto de Portugal); these statistics represent the number of users or visitants, by month: <https://dados.rcaap.pt/handle/10400.20/1/statistics?locale=en>

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

Accept

#### **Reviewer 2**

Comments:

Accept

## **ORGANIZATIONAL INFRASTRUCTURE**

### **1. Mission/Scope**

***R1. The repository has an explicit mission to provide access to and preserve data in its domain.***

#### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

## **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

## ***Response:***

“APIS mission is to increase the use of data resulting from social research carried out in Portugal in particular from representative surveys. Through rigorous data curation and processing, APIS ensures long-term preservation and online dissemination of the data, making sure all data is reliable and easily accessed for the purposes of public consultation, secondary analysis and pedagogical use.”

Depositing data in APIS does not imply the transfer of any rights. APIS is only responsible for a copy of the original data (please see also R2).

APIS works with a network of partners and potential partners in order to accomplish its mission (please see R0 - (5) Insource/Outsource Partners). We are incorporated in the national roadmap as part of the Research Infrastructure PASSDA, one of the highest ranked RI in the Social Sciences & Humanities area in Portugal.

As we already explained in R0, ICS-ULisboa is the leader and host institution of PASSDA consortium. Thus, APIS, being one of the central components of PASSDA, uses some human resources and technology to its advantage.

ICS-ULisboa mission is to produce social data about Portuguese society, as well as to promote the Social Sciences research in the country.

Accordingly, ICS is committed to the preservation and dissemination of social research data, thus assuring the future of Social Science Research in Portugal. ICS's duty to promote and ensure the access to data is accomplished through APIS.

The Portuguese Archive of Social Information (APIS) has been approved as a Service Provider of CESSDA in the CESSDA ERIC 2nd General Assembly Meeting (2017). Our role in CESSDA is to represent the Portuguese collection of research data and contribute to the discovery, as well as the reuse of social sciences data all over Europe, so that researchers and the scientific community can search and (re)use data across borders.

APIS mission

<http://www.apis.ics.ulisboa.pt/en/mission/>

ICS-ULisboa mission

<https://www.ics.ulisboa.pt/en/institute/mission>

RCAAP mission

<http://projeto.rcaap.pt/index.php/lang-en/sobre-o-rcaap/missao-e-objectivos>

CESSDA

<http://cessda.net/National-Data-Services/Other-Data-Services/Portugal>

Portuguese Roadmap of Research Infrastructures

[https://www.fct.pt/apoios/equipamento/roteiro/2013/docs/Portuguese\\_Roadmap\\_of\\_Research\\_Infrastructures.pdf](https://www.fct.pt/apoios/equipamento/roteiro/2013/docs/Portuguese_Roadmap_of_Research_Infrastructures.pdf)

*Reviewer Entry*

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

## 2. Licenses

*R2. The repository maintains all applicable licenses covering data access and use and monitors compliance.*

***Compliance Level:***

4 – The guideline has been fully implemented in the repository

*Reviewer Entry*

**Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

**Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

***Response:***

Depositing data in APIS does not imply the transfer of copyright. However, APIS requires all depositors to sign a declaration of deposit, which gives APIS a non-exclusive licence to distribute data, allowing the archive to store, copy, digitize, modify, migrate and use the data in other collections.

The declaration of deposit [1] also informs the depositor about the main responsibilities of APIS:

1. To store, document and preserve the data according to the best practices in data management;
2. To disseminate the data in APIS catalogue and also in portals such as RCAAP and CDC (CESSDA Data Catalogue);
3. To preserve data in a long-term basis;
4. To acknowledge the authors of the data by means of bibliographic citation;
5. To promote the proper use of the data;
6. To notify the Principal Investigator, whenever possible, about the use of data in new publications;
7. To notify the Principal Investigator about eventual misuse of the data;
8. To notify the Principal Investigator in case of any changes in APIS governance.

From: APIS - access conditions

<http://www.apis.ics.ulisboa.pt/en/access-conditions/>

The depositor ensures in the declaration of deposit that:

1. Data was collected in an ethical way;
2. Data is fully anonymised.

The end-user of the data is informed about the following conditions of use:

- a. To quote the authors of the data in any publication stemming from the secondary use of the data according to the prevailing citation rules in social sciences;
- b. To notify APIS of any publication in any form stemming from the secondary use of the data in order to report the Principal Investigator of such;
- c. To use the data in an ethical way, according to the professional codes of ethic of each discipline;
- d. The end-user undertakes to notify APIS of any inaccuracy or error found in the datasets.

APIS and the authors of the data should not become responsible for the analysis of the data made by the end-user or for his or her opinions expressed in publications.

From the Ethics response (R4), the non-compliance with the conditions of use will be communicated to the ICS-ULisboa Ethics Committee and will immediately lead to the notification of the occurrence next to the Principal Investigator of the study by means of email. At the same time, the user will be advised to comply with the conditions of use of the data. If the user does not cooperate, the following sanctions will be applied:

1. Notification of the occurrence next to the journal, editor or any other entity responsible for the publication;
2. Notification of the occurrence next to the host institution of the user;
3. Notification of the occurrence next to the professional association of the discipline of the user.

In RCAAP (Repositórios Científicos de Acesso Aberto de Portugal) there are Common Licences applied, since when

depositing data, we associate a creative commons license, usually CC-BY according to researcher guidance. For instance, in the deposit agreement, depositors choose if data will be open access or restricted access, available to download upon request, to registered users in RCAAP, or under an embargo period. APIS recommends the use of CC-BY, nonetheless allows other options. Depositors have always chosen CC-BY, as it is justified through the norms of FCT. Therefore, in these terms, the declaration of deposit determines the type of access users will have in the repository RCAAP.

We have translated the most important guidance regarding the adoption of CC-BY license, from the portuguese link in FCT (not available in English).

The emerging consensus on the adoption of the CC-BY license by a number of science funders reflects the fact that restrictions resulting from the application of other licenses unnecessarily limit the possible reuse of published research. By choosing to require a CC-BY license, FCT is also implementing one of the principles recommended by Science Europe in defining minimum Open Access publishing services. According to these principles, European public science funders must demand the application of this license when they sponsor content published in Open Access.

From FCT website:

<https://www.fct.pt/faq/pr.phtml.pt?idFaq=CDEE01DD-2A17-4568-B549-D6CCE226FC8C>

APIS - access conditions

<http://www.apis.ics.ulisboa.pt/en/access-conditions/>

[1] APIS - Declaration of deposit (in Portuguese)

<http://www.apis.ics.ulisboa.pt/wp-content/uploads/2020/02/Declaraçãode-depósito-2020.docx>

APIS - conditions of use

<http://www.apis.ics.ulisboa.pt/en/conditions-of-use/>

FCT Open Access Policy

<https://www.fct.pt/acessoaberto/index.phtml.en>

### *Reviewer Entry*

#### **Reviewer 1**

Comments:

Accept

#### **Reviewer 2**

Comments:

Accept

## **3. Continuity of access**

### ***R3. The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.***

#### ***Compliance Level:***

3 – The repository is in the implementation phase

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

3 – The repository is in the implementation phase

Accept

##### **Reviewer 2**

Comments:

3 – The repository is in the implementation phase

Accept

#### ***Response:***

APIS is responsible for the promotion of continuous access and digital preservation of the data, as stated in the declaration of deposit. APIS mission is to increase the use of data resulting from social research carried out in Portugal in particular from representative surveys.

These tasks are done through rigorous processing, long-term preservation and dissemination of the data, ensuring that they are reliable and can be openly and friendly used for the purposes of public consultation, secondary analysis and pedagogical use.

We disseminate data in RCAAP/SARDC on a long-term basis and also the data is kept in our archive in the University of Lisbon. When depositing any materials in RCAAP, the repository provides a PID (handle) related to the dataset (and associated metadata and related materials to the study).

<http://projecto.rcaap.pt/index.php/lang-pt/consultar-recursos-de-apoio/faqs/3-repositorios-cientificos/24-que-tipo-de-identificadores-persistentes-sao-usados-no-repositorio>

ICS-ULisboa is the host institution of APIS, providing facilities, services and punctual financial support.

Since APIS collection does not contain sensitive data and that it is open access (or to registered users in RCAAP), in case of unexpected or advance ceasing of funding, the continuity of access to the data and its preservation will be ensured by the RCAAP infrastructure of the outsource partner (FCCN). The data is also included in the University of Lisbon (ULisboa) since its backup is always guaranteed due to the fact that our host institution (ICS-UL) makes part of University of Lisbon.

The mission, continuity, transition of activity and succession in case of unexpected circumstances will be decided by ICS-ULisboa (the host institution) and also by the research infrastructure PASSDA (Production and Archive of Social Science Data) where APIS is integrated. Ultimately, ICS-ULisboa as the host institution will be responsible for the

governance of APIS, and for the continuity of its activities as a research infrastructure.

In terms of certification, it should be noted that RCAAP (a SARI repository) was submitted to an audit with reference to ISO 16363:2012 - Audit and certification of trustworthy digital repositories. It was under a large-scale audit initiative of the Institutional Scientific Repositories of Portugal that covered around 28 repositories that revealed reliability in the management and preservation of digital objects.

You can see the entire article (2015) here: <https://openaccess.sdum.uminho.pt/?p=4133>.

The main conclusion can be summarized like this: “The internal audit revealed a high level of maturity regarding the fulfilment of the normative requirements, (...) and being well prepared in terms of infrastructure, reliability and security in the preservation of digital objects.” (translated from the original version in Portuguese).

In order to reach a higher compliance level in the future, APIS is committed to (1.) the improvement of metadata and all the components which we may say are included in a technical level; also (2.) to managing APIS enhancement at institutional level.

#### 1. Technical/Metadata quality:

- a) Development of metadata, complying with DDI vocabularies, what can potentially make our data fairer, assuring continuity of access;
- b) Preparation of Dataverse installation, which can provide better versions of DDI, more compliant with CESSDA standards, also in what refers to discovery (with “doi” instead of “handle”).

#### 2. Institutional level:

- a) The main objective of having FCT mandate to consider APIS as a reference data archive in Social Sciences in Portugal, also regarding Data Management;
- b) Attainment of financial sustainability (by diverse means, e.g., public long-term funding, or/and the increased support from PASSDA partners);
- c) The participation of new research centres in PASSDA consortium;
- d) The integration in advisory bodies and participation in consultative meetings regarding the definition and development of research data policies;
- e) The strengthening of international networking, for example, in terms of participation in RDM networks, like Research Data Alliance (RDA) or European Open Science Cloud (EOSC).

### *Reviewer Entry*

#### **Reviewer 1**

Comments:  
Accept

#### **Reviewer 2**

Comments:  
Accept

## 4. Confidentiality/Ethics

***R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.***

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository  
Accept

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository  
Accept

### ***Response:***

Although social sciences research in Portugal does not have a long tradition on ethics committees, the situation is changing and nowadays research projects have to be presented to an ethics committee. Therefore, APIS does require compliance with that norm to all depositors.

Moreover, ICS-ULisboa has an Ethics Board that conforms to the rules of the European Commission in terms of data collection, protection and storage.

The Ethics Board also verifies if the requirements of the Portuguese Data Protection Law (Law no. 67/68 October 26 are respected: [http://www.cnpd.pt/bin/legis/nacional/lei\\_6798.htm](http://www.cnpd.pt/bin/legis/nacional/lei_6798.htm)).

The following procedures must be contemplated in all projects design:

- 1) total confidentiality in the analysis of the data collected;
- 2) personal data/identifiers used to conduct the surveys/interviews will be eliminated before the archiving of the data at the ICS-ULisboa;
- 3) voluntary and informed consent will be sought in implementing the surveys/interviews, in particular the interviewee will be asked to give his/her express consent before filling in the online questionnaire;
- 4) at all stages participants will be informed about the purposes of the research project;
- 5) only anonymised data will be released onto the public website open access, subject only to straightforward registration.



At APIS, confidentiality and the ethical processing of the data is guaranteed and under ongoing monitoring, in order to avoid constraints to its reuse in open access.

The main stages can be grouped and described in the following way:

## Deposit

The processing of personal data is a responsibility of the researcher who confirms in the declaration of deposit that data has been ethically collected and that it is fully anonymised. In addition, we do not deal with historical data, so it minimizes the risk of having personal data in datasets. APIS makes available to researchers (through the website) the main steps for the anonymization of the data before depositing it:

- Removal of any variable that may lead to the direct identification of a respondent (ex.: names, addresses, postcode, telephone number, email).
- Recode of variables that may lead to the indirect identification of a respondent (ex.: school, occupation, income, age).
- Recode of open questions that may present a risk of identification of a respondent. In this case, the depositor has to describe the way that the recodification was made (ex.: grouping the most mentioned answers).

We do also have in our webpage information about good practices in research; there is info about the DMP and researchers can download it. We also provide some information about data protection and informed consent.

<http://www.apis.ics.ulisboa.pt/consentimento-informado/>.

## Ingest

The staff ensures that there are no direct or indirect identifiers in the deposited data sets. If the deposited dataset has personal data, the archive staff will send it back to the depositor requesting the anonymization of the data as instructed in the website. Therefore, concerning sensitive data and anonymization, there is a good communication channel between the researcher/depositor and the data manager, so that ethics in pre-deposit and deposit phases are guaranteed.

## Access (post-deposit Ethics)

All datasets are anonymised and open access. The archive informs the end-user about the conditions of use of the data and has sanctions in place in case of non-compliance.

Non-compliance with the conditions of use will be communicated to the ICS-ULisboa Ethics Committee and will immediately lead to the notification of the occurrence next to the Principal Investigator (PI) of the study by means of email. At the same time, the user will be advised to comply with the conditions of use of the data. If the user does not cooperate, the following sanctions will be applied:

1. Notification of the occurrence next to the journal, editor or any other entity responsible for the publication;
2. Notification of the occurrence next to the host institution of the user;
3. Notification of the occurrence next to the professional association of the discipline of the user.

ICS-ULisboa Ethics Board

[https://www.ics.ulisboa.pt/en/organizacao/orgaos\\_consultivos](https://www.ics.ulisboa.pt/en/organizacao/orgaos_consultivos)

Data anonymization for depositors

<http://www.apis.ics.ulisboa.pt/en/preparing-data-and-documentation/>

Conditions of use

<http://www.apis.ics.ulisboa.pt/en/conditions-of-use/>

GDPR/Data Protection + Data Protection Officer (DPO)

<https://www.ics.ulisboa.pt/en/research/ethics/data-protection>

About the ICS-ULisboa Ethics Committee (legal documents on the right side)

<https://www.ics.ulisboa.pt/en/research/ethics/ethics-comission>

*Reviewer Entry*

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

## 5. Organizational infrastructure

*R5. The repository has adequate funding and sufficient numbers of qualified staff managed through a clear system of governance to effectively carry out the mission.*

## ***Compliance Level:***

3 – The repository is in the implementation phase

### ***Reviewer Entry***

#### **Reviewer 1**

Comments:

3 – The repository is in the implementation phase  
Accept

#### **Reviewer 2**

Comments:

3 – The repository is in the implementation phase  
Accept

## ***Response:***

APIS is hosted by ICS-ULisboa. ICS-ULisboa was evaluated as an institution of excellence in all research assessments carried out by the Foundation for Science and Technology (FCT) since 1996.

The funding of APIS is sufficient for the fulfilment of its mission and core activities. Such activities have low maintenance costs and have been supported by research projects and the host institution which provides APIS with administrative and IT resources. Currently, there is an adequate budget for attending national and international meetings.

APIS does rely on the success of the funding of the research infrastructure PASSDA (structural funds). The funding is subject to an assessment that usually is made every three years. In last evaluation (RESEARCH INFRASTRUCTURES – MATURITY EVALUATION 2019 / FINAL REPORT) PASSDA has got the highest score “very high”, being the only RI in social sciences that reached the maximum: “The RI has a very high degree of maturity.”. Even though, we should reiterate that the ultimate institution that we depend on (in case of PASSDA not getting the funding) is ICS-ULisboa.

Our staff consists of:

- on a permanent basis:

1 Manager

1 Data Manager

- not regular basis:

1 IT technician

1 administrative officer

The manager of the archive is a researcher who allocates most of his time work to APIS; the data manager works full-time; the administrative officer as well as the IT technician give punctual support whenever is needed.

Our organizational organogram is in APIS website:

<http://www.apis.ics.ulisboa.pt/en/organizational-infrastructure/>.

It shows the relations between APIS and other organizations/research infrastructures. In this page on the website you can check APIS staff: <http://www.apis.ics.ulisboa.pt/en/apis-team/>.

We may say that the staff has been sufficient due to the amount of datasets received per year.

Nevertheless, the process of enhancing the human resources in the archive has already began, as we have already started accessing the recruitment of more people during 2021 through the following sources:

- financial budget of PASSDA due to the very positive evaluation by FCT will enable hiring one more person working full-time;
- negotiation with ICS' Executive Board in terms of allocation of more working hours from the technical staff (IT specialists) to the archive;
- the members of PASSDA, which are research centres, will also be able to collaborate with APIS, through their library staff.

The staff has appropriate qualifications to perform the tasks they are assigned to, mostly because they are qualified in the fields of Sociology. APIS ensures the training of the manager and data manager as evidenced by their participation in training promoted by CESSDA and CESSDA partners. The training covered subjects such as research data management, the Open Archival Information System, Data Documentation Initiative, etc.

Training priorities are the following:

- Research Data Management
- Data Curation
- Metadata (Data Documentation Initiative)
- Online Content Management

In the framework of CESSDA, APIS is involved in the following working groups:

- CESSDA Working Group – Communications Group
- CESSDA Service Providers' Forum

### *Reviewer Entry*

#### **Reviewer 1**

Comments:  
Accept

#### **Reviewer 2**

Comments:  
Accept

## **6. Expert guidance**

***R6. The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either inhouse or external, including scientific guidance, if relevant).***

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

### ***Response:***

APIS has informal consultants and formal channels in order to receive technical expertise, namely FCCN. We also receive feedback from the researchers and students from ICS-ULisboa who are closer to the archive and work with the datasets.

The archive communicates with the Designated Community for feedback by email or directly in meetings or events. We can document APIS recognition in the designated community with the link to the statistics of our collection in RCAAP: <http://dados.rcaap.pt/handle/10400.20/1/statistics>.

Our staff attends CESSDA expert seminars and training workshops and webinars, so that data management expertise is acquired.

APIS staff also attends workshops and seminars in the host institution so that we are aware of trends and developments in the research community.

In November 2019 we organized in ICS-ULisboa, where APIS is based, a training workshop on Data Management Expert Guide (DMEG). It focused on disseminating the tools provided by the CESSDA Training Working Group. This kind of initiative allows us to communicate with the researchers and scientific community the main aspects included in what is generally considered “best practices”; namely the use of Data Management Plan (DMP) and also the compliance with the FAIR Principles. APIS has thus an

important role among the scientific community in social sciences. The participants found the workshop very useful and with relevant topics, since “The CESSDA Data Management Expert Guide aims to put social scientists at the heart of making their research data findable, understandable, sustainably accessible and reusable”.

APIS has also the Ethics Committee of ICS-ULisboa at its disposal, and the Scientific Advisory Board.

Thus, we can summarize two components in the process of getting guidance expertise in APIS:

#### 1) Feedback

- a) informal meetings with researchers and students;
- b) formal and periodic meetings with PASSDA, in which all centres of the consortium are present to discuss their needs, confronting with APIS actions and procedures;
- c) in the future we may create a mechanism to measure if we are responding to the needs of our designated community with a brief, random sample and periodic survey for the research centres, as well as for data users and data depositors in APIS.

#### 2) Technical expertise

- a) Training and activities regarding the development of data management expertise through CESSDA channels;
- b) Meetings with FCT/FCCN representatives, usually from RCAAP, which allows APIS to be updated in what refers to technology and preservation of data;
- c) Participation in national networks that function at international level, such as RDA-PT;
- d) Available support from technical Infrastructures included in Portuguese Roadmap of Research Infrastructures (link below), namely:
  - INCD - the Portuguese National Distributed Computing Infrastructure, a virtual RI - available technical support (including specialized IT support) to every infrastructure in the roadmap;
  - RCTS - the Science, Technology and Society Network, a single-sited RI (also referred in R9 and R16) - it is the Portuguese Research and Education Network (NREN), a dedicated high-performance network to serve researchers, teachers and students with greater demands, acting as a test platform for advanced communications services and applications. RCTS provides a privileged collaboration channel for Portuguese researchers to access foreign research infrastructures, data sets and services.

Links to events that APIS have participated:

<http://sodanet.gr/en/training/workshops/cessda-training-workshop-athens-2019>

<https://www.cessda.eu/News-Events/Events/CESSDA-Trust-Workshop>

<https://www.opensciencefair.eu/>

<https://forumgdi.rcaap.pt/5forum/>

<https://www.cessda.eu/News-Events/Events/CESSDA-Trust-Workshop2>

Ethics Commission of ICS-ULisboa

<https://www.ics.ulisboa.pt/en/research/ethics/ethics-comission>

Consultive Bodies (ICS-ULisboa):

<https://www.ics.ulisboa.pt/en/institute/general>

EXTERNAL ADVISORY BOARD

PASSDA - Board, Scientific Council and Advisory Council

<http://passda.pt/quem-somos/?lang=en>

Portuguese Roadmap of Research Infrastructures

[https://www.fct.pt/apoios/equipamento/roteiro/2013/docs/Portuguese\\_Roadmap\\_of\\_Research\\_Infrastructures.pdf](https://www.fct.pt/apoios/equipamento/roteiro/2013/docs/Portuguese_Roadmap_of_Research_Infrastructures.pdf)

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

Accept

##### **Reviewer 2**

Comments:

Accept

## **DIGITAL OBJECT MANAGEMENT**

### **7. Data integrity and authenticity**

*R7. The repository guarantees the integrity and authenticity of the data.*

#### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

## ***Response:***

### Data and Metadata - completeness and changes

To ensure the integrity of the data, the data manager follows the documented procedures described in APIS workflow. During ingest, the metadata table and submitted files are checked for completeness. The data manager processes SIPs into AIPs and DIPs in accordance with detailed instructions in APIS internal document (please see attachment), corresponding to the APIS workflow for “internal guidance”; it is like a lively document, updated when needed.

Checksums are run monthly and they confirm the integrity of deposited and disseminated data files in RCAAP/DSpace. If the checksums do not match, files are investigated for inconsistencies, mitigating the risk of inadvertent or unauthorised changes to the data. Thus, in order to guarantee the integrity of the data, i.e. ensure data was not erroneously altered or changed without permission in any moment (ingest, archival, storage, data access), APIS is now using MD5 Command Line Message Digest Utility. RCAAP uses DSpace and this has automatic md5 checksums. The process is better explained in this link: <https://comum.rcaap.pt/help/index.html>.

Metadata are always linked to data with the ID number provided for the study by APIS. As the Nesstar software provides DDI version 1.2, APIS data is in compliance with the Data Documentation Initiative (DDI), which is “(...) an international standard for describing the data produced by surveys and other observational methods in the social, behavioral, economic, and health sciences.” (<https://ddialliance.org/>)

In the near future, we will install a new software that will allow us to adopt a higher version of DDI.

Currently, all changes made to data with the purpose of quality control are registered in a report that lists all changes. The procedure is saving the reports in the Data Management folder identified with the study ID. In this way, 'authenticity' (in terms of the repository-side of provenance) starts here.

Version responsibility is assigned to one member of the staff, the data manager, the only person who is authorized to access the data and change it, always communicating with the depositor and/or the author of the study. Every change that is made to the data file or associated metadata is reported in the study report, also elaborated by the data manager. Our policy of versioning considers minor and major changes in data and/or metadata, which will have different implications in terms of numeric classification.

### Provenance Data

The authenticity of the data is primarily guaranteed by the depositor who deposits the dataset and associated materials. We require the depositor to guarantee that data was ethically produced and afterwards APIS also does the curation work (please see R8).

The declaration of deposit is written in Portuguese, nevertheless you can check the content of the document in this link: <http://www.apis.ics.ulisboa.pt/wp-content/uploads/2020/02/Declaração-de-depósito-2020.docx>.



APIS ensures file provenance through the following steps: the first ingest step is the storage of provenance data, assigning an archival number (ID) to the respective dataset (e.g. APIS0031), which is stored with a report containing the author, the title, abstract, the name of the data depositor, etc. In the SIP, all authentic formats we received from the depositor are archived. In the AIP, the declaration of deposit and all processed files are stored and documented. Lastly, access is guaranteed through RCAAP.

#### Identity Checks of Depositors

The identity of the depositor is checked by consulting the website of his/her host institution, the ORCID when possible, contacting him or her directly or checking his/her email domain, which has to be institutional. This might be sufficient in our archive due to the fact that the social science research community in Portugal is relatively small and most of it is integrated in PASSDA research centres.

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Accept

##### **Reviewer 2**

Comments:  
Accept

## 8. Appraisal

*R8. The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for data users.*

### *Compliance Level:*

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
4 – The guideline has been fully implemented in the repository  
Accept

##### **Reviewer 2**

Comments:  
4 – The guideline has been fully implemented in the repository  
Accept

## *Response:*

### Collection development policy

Research Data collected by social scientists, preferably:

- Research projects with competitive funding (e.g. FCT open calls);
- Scientific/relevant value (based on the publications/articles related to the deposited dataset; and also in some feedback from our regular meetings with research centres of PASSDA);
- National coverage (representative surveys);
- High expected use considering the designated community (based on the subjects considered more interesting by social researchers in the past few years, for searching, downloading and (re)using);
- Should be related with the topics covered by the archive (fields of Sociology, Political Sciences and Social Psychology).

### Quality control checks

- To verify the completeness of the file received, both case numbers and variable numbers;
- To verify the correct correspondence between the survey and the dataset variables, namely what concerns the proper classification of the type of variable (numeric or string) and its categories, or even the way missing data were managed;
- To verify the existence of missing values and its clear identification, particularly in regard to “no answers”;
- To verify the consistency of the answers that depend on individual features such as age, gender, civil status or level of education;
- To verify the appropriateness of recoding and to ensure that it is properly documented;
- To verify compliance with legal principles concerning data protection, by checking if the dataset is anonymised.

### Metadata requirements

The archive requires the depositor to submit some metadata related to the study, according to CESSDA mandatory fields, based on ERIC CMM. The staff documents the data with more metadata based on the documentation (such as method reports) provided by the depositor, so that the study metadata in Nesstar can be the most exhaustive as possible.

Recently, associated with publishing some studies in CDC (CESSDA Data Catalogue), there was the need to have at least the metadata which is there provided.

If the metadata is not enough, the depositor is contacted in order to supply the archive with more information and those contacts are registered in the study report.

### List of preferred formats

- SPSS
- Excel
- Stata
- PDF
- Word

By preferred formats we mean formats whose process of curation, dissemination and preservation are guaranteed by the archive. Every other format is accepted and migrated to preferred formats whenever possible. Otherwise, data is returned to the depositor.

This information is available on the APIS website (following links).

APIS - Data and Metadata criteria

<http://www.apis.ics.ulisboa.pt/en/data-and-metadata-criteria/>

APIS mission and workflow

<http://www.apis.ics.ulisboa.pt/en/mission/>

How to deposit

<http://www.apis.ics.ulisboa.pt/en/deposit/>

Preparing Data and Documentation

<http://www.apis.ics.ulisboa.pt/en/preparing-data-and-documentation/>

Preferred file formats

<http://www.apis.ics.ulisboa.pt/en/preparing-data-and-documentation/>

Metadata form

<http://www.apis.ics.ulisboa.pt/metadados/>

*Reviewer Entry*

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

## 9. Documented storage procedures

*R9. The repository applies documented processes and procedures in managing archival storage of the data.*

**Compliance Level:**

4 – The guideline has been fully implemented in the repository

### *Reviewer Entry*

#### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

#### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

### *Response:*

We refer to the documentation of APIS practices from deposit to the point that it passes storage responsibility onto a partner and the documented storage information about that partners' approach. For that, we have an internal handbook/manual in order to document our processes (document attached).

About our partners' security storage procedures, please see R16.

APIS mission

<http://www.apis.ics.ulisboa.pt/en/mission/>

All data is securely handled and organised according to the OAIS model. The storage procedures are well defined in APIS workflow (R12).

APIS holdings are stored in two different locations with different functions: a) the host institution server (ICS-ULisboa) associated with SIP and AIP; and b) RCAAP/SARDC, aimed at storing the DIP.

The Archival Information Package (AIP) is stored in APIS server (in the host institution) for preservation purpose, providing the base for the Dissemination Information Package (DIP) in Scientific Data Hosting Service (SARDC/RCAAP).

The dataset is uploaded to the Scientific Data Hosting Service in RCAAP (<http://dados.rcaap.pt>) which gives a persistent identifier to each dataset (handles), so that the APIS online collection is updated and the data manager is notified of such by email.

In terms of backup strategies, we do also have different locations: the host institution server, the central services of University of Lisbon and APIS google drive; RCAAP/SARDC is responsible for backup copies of deposited research data (as referred in R0).

Data is backed up on a weekly basis in the server of the host institution (ICS-ULisboa) and in the central services of University of Lisbon; also on a daily basis in the APIS archive's drive (which is the google drive - cloud - only shared between the manager and the data manager).

Scientific Data Hosting Service (SARDC) of RCAAP, as a research data repository, is also responsible for long-term storage. Therefore, the outsource partner (FCCN) is part of our preservation policy, namely in terms of data storage, backup/multiple copies, data recovery provisions and risk management (see also R15 and R16). RCAAP/SARDC has its own backup strategy.

FCCN practices deal with security of data (R16) and involve a Trusted Certificate Service (TCS) which ensures the supply of several types of digital certificates to the national entities connected to RCTS - Science, Technology and Society Network.

The links below document these processes:

<https://www.fccn.pt/en/security/rcts-cert/>

<https://www.fccn.pt/en/security/rcts-certificates/>

<https://docplayer.net/18228041-Trusted-certificate-service-tcs.html>

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

Accept

##### **Reviewer 2**

Comments:

Accept

## **10. Preservation plan**

***R10. The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.***

### ***Compliance Level:***

3 – The repository is in the implementation phase

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

3 – The repository is in the implementation phase

Accept

## Reviewer 2

Comments:

3 – The repository is in the implementation phase

Accept

## *Response:*

APIS is developing and implementing the preservation policy, since we are making efforts to develop new ways of accomplish long-term preservation, i.e. considering responding to designated community feedback (e.g. demand for new formats) and maintenance of metadata.

The declaration of deposit which is signed by the researcher/depositor authorizes APIS to proceed to all modifications in order to ensure continuous access and preservation of the information in the long term.

<http://www.apis.ics.ulisboa.pt/wp-content/uploads/2020/02/Declaração-de-depósito-2020.docx>

Depositing data in APIS does not imply the transfer of copyright. However, APIS requires all depositors to sign a declaration of deposit, which gives APIS a non-exclusive licence to distribute data, allowing the archive to store, copy, digitize, modify, migrate and use the data in other collections. Without this declaration, digital preservation of the data would be compromised.

From APIS website:

<http://www.apis.ics.ulisboa.pt/en/access-conditions/>

Datasets are the main focus of the preservation policy. To that end, one of the preferred formats of datasets to the archive is SPSS, a software which is widely used by the designated community. Migration to different versions or formats occurs every time is necessary. For example, to import data to Nesstar, we convert the dataset file that originally was in SPSS (.sav) format to Stata (.dta).

In what refers to preservation and particularly to AIP, our preferred format is .csv (comma-separated values file). Also, APIS produces to each study (option in Nesstar “export DDI”) a .xml file for each dataset (plus metadata); the .xml file and the codebook are kept in the AIP folder.

The .csv file is important and indispensable in terms of preservation policy, due to the fact that in case of something unexpected can happen, this format will help us to create something new.

## Legal Requirements for Preservation

The “contract” between depositor and APIS, in our case the transfer and licence agreements, provide for all actions necessary to meet the responsibilities for data preservation on a long-term basis, including the rights to copy, transform, and store the items as well as provide access.

APIS holdings are stored in two different locations: in the host institution server (ICS-ULisboa) and on RCAAP/SARDC. In

terms of backup strategies, we do also have different locations: the host institution server, the central services of University of Lisbon and APIS google drive.

Besides, RCAAP has its own way of digital preservation with DSpace software backup: <http://projecto.rcaap.pt/index.php/ang-pt/consultar-recursos-de-apoio/faqs/3-repositorios-cientificos/18-como-e-que-o-dspace-preserva-o-material-digital>.

#### Plan for long-term preservation

In the near future, we intend to make improvements at information level required in data description, as well as to develop a service specifically intended for the long-term preservation of data, which requires harmonising file formats and complementing the metadata with more preservation descriptive information and provenance data.

To ensure the usability of the data in the long term, APIS is also intending to include in the curation process the production of detailed metadata in Portuguese and English, which can also be enabled with the installation of a new software. As indicated in R3, the possible installation of a more advance software could provide richer versions of DDI, enabling the discovery, reuse and a better preservation of data, as more technical information is included. Thus, the compliance with FAIR Principles would be also dependent from this new software, which could provide (meta)data in machine readable formats.

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:  
Accept

##### **Reviewer 2**

Comments:  
Accept

## 11. Data quality

***R11. The repository has appropriate expertise to address technical data and metadata quality and ensures that sufficient information is available for end users to make quality-related evaluations.***

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

## **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

## ***Response:***

Data quality is addressed during Ingest and Data Management functions in curation process.

### Ingest

In this phase, the quality control is guaranteed by the data manager, who:

- Verifies the declaration of deposit;
- Verifies if the metadata form is complete (if it has the mandatory fields to CDC - CESSDA Data Catalogue - and if it complying with DDI controlled vocabularies, accordingly to CESSDA ERIC Metadata Model - CMM 1.0);
- Verifies if all files can be open;
- Verifies the format in which the data was deposited and if it complies with the list of preferable formats (and whenever is needed and possible, formats are change);
- Confirms there are no direct or indirect identifiers of the respondents and in case there are, the dataset is returned to the researcher who is required to anonymise the data.

### Data Management

APIS team provides guidance and instructions not only about how to prepare data and related documentation, but also emphasizes the importance of compliance to the FAIR principles. Thus, in order to ensure the metadata completeness and understandably, APIS stimulates the cooperation with the depositor to organize and produce the metadata for the study. As we indicated in R10, we produce a xml file (with DDI metadata) for each deposited dataset.

Besides that, another way to promote and control the quality of our data is through workshops and seminars about RDM. For instance, APIS delivered in November a workshop focused in the best practices of social science research, regarding the organization and deposit of data, demonstrating how DMEG (Data Management Expert Guide) and DMP (Data Management Plan) are used in order to help the scientific community to make their research data Findable, Accessible, Interoperable and Reusable (FAIR).

### Meeting the expectations of the Designated Community

APIS encourages the feedback from researchers by the means referred in R6 and R8. Taking advantage from the fact that we are located in ICS, a research centre, there is always an attempt to get informal opinions about data quality from some of our main depositors. Also our regular meetings with members from research centres, in scope of PASSDA consortium,



contribute to get critical opinions concerning data quality, and the ways in which data is provided.

The archive also stimulates the cooperation with the community showing the benefits of depositing data, namely providing more visibility of research projects, especially in international terms through CESSDA Data Catalogue.

The sociological background of APIS staff, namely its coordinator and data manager, helps to be closer to the researchers, sharing the same scientific language and knowledge, as most of them are from Sociology field.

In order to improve the quality of information about the research data, RCAAP provides for each study, besides data, some documentation that is central to the research project, for example, the questionnaire, the codebook, the methodological report, etc. Moreover, we guarantee that metadata are always public and openly available (with CC-BY license).

In addition, RCAAP also provides some statistics of visits to the link to access data and about the download of each file deposited in the repository.

(e.g. <https://dados.rcaap.pt/handle/10400.20/2080/statistics>)

In the near future, APIS intends to apply a random survey to users/researchers, to check if we are responding to the needs of our designated community in terms of quality of data and metadata.

It is also important to develop better ways of discovery and search data, by improving the functionalities and options that are available on APIS website; note that this may include an enrichment at metadata level and the installation of a new software that may allow a more diverse and deeper search.

FAIR in “Preparing Data and Documentation”

<http://www.apis.ics.ulisboa.pt/en/preparing-data-and-documentation/>

Best Practices

<http://www.apis.ics.ulisboa.pt/en/best-practices/>

Workshop on DMEG (and materials)

<http://www.apis.ics.ulisboa.pt/en/workshop-on-dmeg-06-11-2019/>

*Reviewer Entry*

**Reviewer 1**

Comments:  
Accept

**Reviewer 2**

Comments:  
Accept

## 12. Workflows

## ***R12. Archiving takes place according to defined workflows from ingest to dissemination.***

### ***Compliance Level:***

3 – The repository is in the implementation phase

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

3 – The repository is in the implementation phase

Accept

##### **Reviewer 2**

Comments:

3 – The repository is in the implementation phase

Accept

### ***Response:***

The Open Archival Information System (OAIS) is the reference model for the workflow of APIS and it is displayed in our website, in "Mission and Workflow": <http://www.apis.ics.ulisboa.pt/en/mission/>.

The information related to APIS workflow which is of the most interest for the designated community is available across our website pages. Furthermore, we have been working in an internal document (please see the attachment) describing the technical phases of the workflow and providing general guidance of the archive organization.

In the beginning of 2020 we have published in the APIS website the information below, as well as the workflow diagram.

#### **1. INGEST FUNCTION**

When a Submission Information Package (SIP) is sent to the archive, the staff performs data quality control checks:

- To check the consistency of the dataset and the existence of errors in order to have a first overview of the data quality.
- To check the amount of non-answers, filters and missing values.
- To check how many variables are missing.
- To check the quality of the metadata by comparing the metadata provided with the metadata requested.

All types of errors are corrected and subsequent changes of the data are registered in a report.

Whenever needed, the depositor is contacted using the channels of communication provided in the declaration of deposit.

This can lead to the deposit of a new Submission Information Package (SIP). Accordingly, the email is used where

required to comment data or metadata, but if possible the data manager contacts directly the depositor of the study.

The Submission Information Package (SIP) is stored in two different locations: in APIS server and also in the data centre of the University of Lisbon.

## 2. DATA MANAGEMENT FUNCTION

Version 1 of the Archival Information Package is created. The folder and the study are named according to the archive standards. Then, the data manager enhances all datasets by producing metadata according to standards such as DDI and documenting it with other files such as method reports.

All datasets are fully documented using Nesstar Publisher which is compliant with the DDI standard.

The archive uses:

- CESSDA mandatory (for CDC) and some of recommended metadata fields;
- CESSDA controlled vocabularies.

In case of not having all the metadata required, we review with the depositor the missing metadata, or when not possible, we just publish it in RCAAP.

We should also enlighten that APIS contributed in the working group of CESSDA MDO (Metadata Office) to the translation (to Portuguese - PT) of controlled vocabularies.

This is the main link: <https://vocabularies.cessda.eu/#!discover>.

## 3. ARCHIVAL STORAGE FUNCTION

The dataset is uploaded to the Scientific Data Hosting Service in RCAAP (<http://dados.rcaap.pt>) which gives a persistent identifier to each dataset (handles). The online catalogue is updated and the data manager is notified of such by email.

The Archival Information Package (AIP) is stored in two different locations: in APIS server (in the host institution) and in Scientific Data Hosting Service (SARDC/RCAAP).

## 4. ACCESS FUNCTION

APIS website is the main hub to access data. All data holdings are available through the online catalogue and can be analysed using the Nesstar tool. The catalogue gives access to some metadata describing the contents of the study.

Whenever the user wants to download data, s/he will be redirected to the Scientific Data Hosting Service (RCAAP). Users are required to comply with “conditions of use” both in the APIS website and at Scientific Data Hosting Service. The entire collection of APIS is open access and with no costs for the end user.

APIS Mission and Workflow

<http://www.apis.ics.ulisboa.pt/en/mission/>

CESSDA/DDI Controlled Vocabularies

<https://vocabularies.cessda.eu/#!discover>

APIS Online Catalogue

<http://www.apis.ics.ulisboa.pt/en/catalogue/>

APIS Nesstar WebView

<http://nesstar.ics.ul.pt/webview/>

Scientific Data Hosting Service

<https://dados.rcaap.pt/>

APIS in SARDC/RCAAP

<https://dados.rcaap.pt/handle/10400.20/1>

*Reviewer Entry*

**Reviewer 1**

Comments:

Accept

**Reviewer 2**

Comments:

Accept

## 13. Data discovery and identification

*R13. The repository enables users to discover the data and refer to them in a persistent way through proper citation.*

***Compliance Level:***

4 – The guideline has been fully implemented in the repository

*Reviewer Entry*

**Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository  
Accept

## **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository  
Accept.

## ***Response:***

All datasets deposited in APIS are available both in APIS online catalogue and in RCAAP website.

Accordingly, APIS offers two main search facilities:

- a) The online catalogue (users can search data by means of a simple search by words);
- b) Users can browse through the RCAAP online catalogue and find the APIS collection there (<https://dados.rcaap.pt/handle/10400.20/2?locale=en>).

A subset of APIS data can be discovered via the CESSDA Data Catalogue. Moreover, metadata are available to all RCAAP users and in nesstar webview (online analysis) without registration.

APIS's data holdings are also available in general search engines, as Google. In fact, CESSDA Data Catalogue has the main purpose of data discovery and supports the data reuse, making accessible the landing page of each APIS study, where the proper citation can be found.

In 2019 we had the first studies in CDC (CESSDA Data Catalogue); by now thirty-one studies are available. The search can be carried out by using some of the controlled vocabularies as filters.

The interface for harvesting our metadata into the CDC is Kuha, which is compliant with the OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting).

APIS, as a CESSDA Service Provider, uses internationally agreed upon metadata standards, so that consistent descriptive metadata is enabled by using the DDI's Controlled Vocabularies.

Thus, based on CESSDA ERIC Metadata Model (CMM 1.0), APIS promotes the use of the renewed version of CESSDA Topic Classification and DDI standardised vocabulary.

APIS has recently started using the ELSST (European Language Social Science Thesaurus) standard -

<https://thesauri.cessda.eu/elsst/en/> - to the keywords in both search facilities. Topic Classification is also used in APIS online catalogue, even though we can say that these two search facilities are limited in terms of metadata standards.

Both the CESSDA Data Catalogue and nesstar webview provide more advanced possibilities for data discovery. It is also of interest to mention nesstar webview as a resource, which should be potentiated due to its possibilities through "Advanced Search".

In the platform nesstar webview it is indeed possible to search studies by some of DDI standard metadata. In each study in our online catalogue, there is already a link to the study in nesstar webview and it is possible to use the "advanced search" through DDI metadata standards.

While processing metadata to the studies, APIS does use several controlled vocabularies: Analysis Unit, Mode of Collection, Sampling Procedure, Time Method, Type of Instrument and CESSDA Topic Classification (<https://vocabularies.CESSDA.eu/#!discover>). Most of these metadata are also available in CDC.

In the future, it will be important to develop better ways of discovering data in APIS website, by improving the functionalities and options available for online catalogue filtered search, by DDI metadata standards. Also when adopting another software for publishing (as Dataverse), we will be able to provide a more refined mechanism for searching, using DDI standards for metadata search.

In order to make data discoverable and identifiable, APIS has persistent identifiers (PID) - a handle provided by RCAAP - and provides a recommended citation format for each dataset.

This citation with the persistent identifier (handle) is included in the study description of archived and disseminated data, to ensure the correct citation of the data in all publications and presentations resulting from the data.

Concerning bibliographic citation, the archive uses the following elements: Authoring entity; Year of publication; Study title; Year of production; Distributor; Place; Serial number; Identifier.

Example:

Baptista, Luís Vicente; Resende, José Manuel; Rodrigues, Luís; Macedo, Paulo Machado (2012): Atitudes, expectativas e aspirações dos jovens face às Forças Armadas, 2011. Arquivo Português de Informação Social, Lisboa. APIS0031, <<http://hdl.handle.net/10400.20/53>>

The archive intends to provide Digital Object Identifiers (DOI) when the outsource partner does. In 2019, we have applied to have a “doi” instead of a “handle” PID, generated by RCAAP, but RCAAP decided to maintain the “handle”.

Besides the registration in RCAAP, and not excluding future participations in other listings, APIS is registered in re3data (Registry of Research Data Repositories):

<https://www.re3data.org/repository/r3d100013420>.

CESSDA vocabularies

<https://vocabularies.CESSDA.eu/#!discover>

ELSST

<https://thesauri.CESSDA.eu/elsst/en/>

APIS catalogue

<http://www.apis.ics.ulisboa.pt/en/catalogue/>

RCAAP portal

<https://www.rcaap.pt/>

Scientific Data Hosting Service

<https://dados.rcaap.pt/>

APIS in RCAAP

<https://dados.rcaap.pt/handle/10400.20/1>

APIS in Nesstar WebView

<http://nesstar.ics.ul.pt/webview/>

APIS in CDC (CESSDA Data Catalogue)

[https://datacatalogue.cessda.eu/?publisher.publisher\[0\]=Portuguese%20Archive%20of%20Social%20Information%20%28APIS%29](https://datacatalogue.cessda.eu/?publisher.publisher[0]=Portuguese%20Archive%20of%20Social%20Information%20%28APIS%29)

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

Accept

##### **Reviewer 2**

Comments:

Accept

## **14. Data reuse**

*R14. The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data.*

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

## ***Response:***

All depositors are required to fill a web deposit form which includes the CESSDA Mandatory Fields included in CDC (CESSDA Data Catalogue). If they have questions, the data manager will provide assistance with the metadata. The most important fields are the ones of CDC and they are compiled in a data table, which is available in the APIS website (link below).

<http://www.apis.ics.ulisboa.pt/wp-content/uploads/2020/02/TabelaMetadados.docx>

APIS receives and revises the study metadata, the data manager completes the metadata fields and Nesstar software collects the metadata related to the dataset, so it has some fields already harvested, such as the labels.

Metadata is available through a general search facility in our main page website and also in Nesstar WebView (also see R13).

For a better understanding of the process, when submitting the dataset to APIS, the depositor should fill the metadata form. This deposit requirement provides more information about the research team, the project and related publications, which can allow a better understanding and a wider contextualization of the research project when reusing data.

APIS, as a CESSDA Service Provider, uses internationally agreed upon metadata standards, so that consistent descriptive metadata is enabled by using the DDI's Controlled Vocabularies.

Thus, based on CESSDA ERIC Metadata Model (CMM 1.0), APIS promotes the use of the renewed version of CESSDA Topic Classification and DDI standardised vocabulary.

The DDI Alliance controlled vocabularies available in Portuguese are as follows: Analysis Unit, Commonality Type, Data Source Type, General Data Format, Mode Of Collection, Sampling Procedure, Summary Statistic Type, Time Method, Type of Instrument; and also CESSDA Controlled Vocabulary for CESSDA Topic Classification.

- <https://vocabularies.cessda.eu/#!discover> -

The data manager is always available to help the depositor complying with these vocabularies, especially regarding the mandatory ones, which are: Analysis Unit, Mode Of Collection, Sampling Procedure, Time Method, Type of Instrument and CESSDA Topic Classification.

CESSDA Vocabulary Service enables users to discover, browse, and download controlled vocabularies in a variety of languages.

APIS provides data in SPSS format, excel or even .csv format. Stata is a possible option although we usually do not hold datasets in this format; commonly it is used for Nesstar purposes. So far, SPSS has filled the needs of our designated community.

SPSS files are updated at the same pace as the software is in order to keep the data understandable and reusable (instead of only ensuring preservation through XML).



So, in Archival Information Package (AIP) we have all datasets in .csv format, also with the additional metadata: a) the xml file from Nesstar (DDI 1.2.2) and b) the codebook (in pdf format). The codebooks are always disseminated through RCAAP, which makes data and metadata reusable.

All the files disseminated in RCAAP (namely SPSS files available for download) are maintained in AIP.

These data formats also ensure possible migrations in the future.

From APIS website:

<http://www.apis.ics.ulisboa.pt/en/data-and-metadata-criteria/>

Metadata form

<http://www.apis.ics.ulisboa.pt/metadados/>

Nesstar WebView page with APIS studies

<http://nesstar.ics.ul.pt/webview/>

#### *Reviewer Entry*

##### **Reviewer 1**

Comments:

Accept

##### **Reviewer 2**

Comments:

Accept

## **TECHNOLOGY**

### **15. Technical infrastructure**

*R15. The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services it provides to its Designated Community.*

***Compliance Level:***

4 – The guideline has been fully implemented in the repository

### *Reviewer Entry*

#### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

#### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository

Accept

### *Response:*

The technical infrastructure of APIS is in two different locations:

- a) ICS-ULisboa;
- b) RCAAP.

a) The archive back office, including its website, is allocated in ICS-UL infrastructure, being the main characteristics as follows:

A VMware vSphere solution was implemented with a cluster composed of two ESXI 6.7 hosts (managed by vCenter Server Appliance 6.7) and an HPE MSA2050 shared storage that supports all infrastructure and production virtual machines.

The network for the entire infrastructure has its own VLAN, for management and for the production part.

A backup solution was set up with a Storeonce3100 unit, working with Veeam Backup, and reinforced with a NAS QNAP unit in production at other ULisboa facilities, which works as storage for remote file backups.

In addition, an IRS server (Insight Remote Support) was installed, which aims to collect the alarms of the HPE equipment and, in an automatic way, inform the manufacturer about the incidents that occurred. With this application it is possible to check the type of contract and its validity for each serial number.

The entire infrastructure is directly supported by two APC Smart UPS X3000 permanent power units, configured with the “PowerChute” management application.

b) RCAAP is based on OpenSource DSpace which is compliant with the Dublin Core metadata scheme and the OAI-PMH protocol; the system has OpenSource operating system Linux, CentOS distribution and redundant infrastructure with high-availability which is provided by FCT/FCCN.

In (5) Insource/Outsource Partners we explain the relation between partners (also in what refers to shared technology), including the obligations and roles that are in the contract (2012).

Evidences:

✓ Public information about the software and the standards:

<https://wiki.duraspace.org/display/DSPACE/Home> and <http://www.dspace.org/why-use>

✓ Information about the implemented standards (DC, MODS, METS):

<https://wiki.duraspace.org/display/DSPACE/Home> and <http://www.dspace.org/why-use>

<https://wiki.duraspace.org/display/DSDOC3x/Importing+and+Exporting+Content+via+Packages>

All data is securely handled by the archive in APIS google drive, where data folders are organized according to the OAIS model. We may also refer to the ISO 16363 compliance of RCAAP, being the main conclusion: "The internal audit revealed a high level of maturity regarding the fulfilment of the normative requirements (...) and being well prepared in terms of infrastructure, reliability and security in the preservation of digital objects."

More information regarding ISO 16363 compliance of RCAAP can be found in R3. (Continuity of access) and R16. (Security).

✓ The repository performs annual assessments on the capacity and forecasting (6-month weighted). The assessment concerns disk space, CPU usage, memory usage and bandwidths consumption.

✓ The information about the installed software is kept in a private wiki connected to the project. All software documentation used can be found here:

<https://wiki.duraspace.org/display/DSDOC3x/Dspace+3.x+Documentation>.

✓ The software is based on the DSpace community.

More information here: <http://www.dspace.org/why-use>.

✓ Information about the installed hardware:

The virtualization system that supports RCAAP virtual machines is based on the CITRIX XenServer solution, which implements a virtualization solution of enterprise-class. The solution offers all critical resources needed for any server implementation and virtualization.

The FCT|FCCN virtualization service is focused on business continuity and it implements a geographically redundant architecture (technical rooms; SE03 and GRID) composed by two Clusters of autonomous virtualization, without a single point of inter-cluster failure. In case of unavailability of one of the data centres, the interoperability between the two clusters establishes a Disaster Recovery Plan (DRP), which guarantees the continuity of service. The Recovery Time Objective (RTO) maximum level is 6 hours and the Recovery Point Objective is 24 hours.

### *Reviewer Entry*

#### **Reviewer 1**

Comments:

Accept

#### **Reviewer 2**

Comments:

Accept

## 16. Security

***R16. The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users.***

### ***Compliance Level:***

4 – The guideline has been fully implemented in the repository

#### ***Reviewer Entry***

##### **Reviewer 1**

Comments:

4 – The guideline has been fully implemented in the repository  
Accept

##### **Reviewer 2**

Comments:

4 – The guideline has been fully implemented in the repository  
Accept

### ***Response:***

In 2015, the repository (RCAAP) conducted a compliance assessment with ISO 16363. It was under a large-scale audit initiative audit of the Institutional Scientific Repositories of Portugal (covering around 28 repositories), which revealed reliability in the management and preservation of digital objects.

In the framework of RCAAP, the critical processes identified were those with direct impact in the:

- Compliance with the obligations established with the adherent entities;
- Legal compliance.

The security provision is ensured by RCAAP and also by ICS-ULisboa virtualization infrastructure previously explained in R15.

The roles and relations between APIS and RCAAP/SARDC are more detailed in R0. Context - (5) Insource/Outsource Partners.

We hold an internal document in portuguese (provided by RCAAP infrastructure management) with some principal aspects. We should note that this requirement involves technical details and more information in tables which are displayed in that internal document and now available on our website in english:

<http://www.apis.ics.ulisboa.pt/en/technology-technical-infrastructure-and-security/>.

There you can find and explore the different processes:

Risk analysis: Note that the risk analysis has been developed by RCAAP for all the repositories and not specifically for APIS. Additionally, APIS does not store sensitive data up to now.

The risk analysis is the result of a partnership with the TIMBUS project (Timeless Business Processes and Services - <http://timbusproject.net/>).

Context model: The context model shows the dependencies of the system – both in terms of infrastructure, software, operating system, data model and obligations – and then, after defining the requirements, a risk model was made. This was intended to know the risks that may affect the system in order to promote the preservation of RCAAP.

The methodology entails the following phases:

- Risk identification
- Risk analysis
- Risk assessment

Based on the work carried out in partnership with the TIMBUS Project, the model defines:

- The elements used to assess the risk
- The metadata used to describe the elements of the model

Contingency Plan: The Contingency Plan intends to anticipate and manage the impact of an eventual disaster with RCAAP services. According to ISO 16363 - 5.2.4, the repository must have a contingency plan that includes at least one copy of all digital information placed in a geographically distant location (including a copy of the contingency plan).

In this sense, recovery plans were made for 4 scenarios: 1) Unavailability of the HOST; 2) Permanent loss of the HOST; 3) Unavailability of the Backup; 4) Violation of the integrity of an Archival Information Package (AIP).

For each scenario, there are specific instructions and plans of action to replace the services.

There is also a set of procedures related to the following typologies: Service maintenance and content preservation; Content deposit; Support to the community; Risk management.

We should still highlight the following references concerning APIS outsource partners' security certifications:

Trust certification of FCCN: <https://docplayer.net/18228041-Trusted-certificate-service-tcs.html>

RCTS Certificates is the service that ensures the supply of several types of digital certificates to the entities connected to

RCTS. This service promotes the security, authenticity, confidentiality and integrity of information sent and received via the Internet. FCCN guarantees this service through the TCS (Trusted Certificate Service), provided by GÉANT Association (involved in EOSC and FAIRsFAIR project): <https://www.fccn.pt/en/security/rcts-certificates/>.

*Reviewer Entry*

**Reviewer 1**

Comments:  
Accept

**Reviewer 2**

Comments:  
Accept

## APPLICANT FEEDBACK

### Comments/feedback

*These Requirements are not seen as final, and we value your input to improve the CoreTrustSeal certification procedure. Any comments on the quality of the Requirements, their relevance to your organization, or any other contribution, will be considered as part of future iterations.*

*Response:*

*Reviewer Entry*

**Reviewer 1**

Comments:

**Reviewer 2**

Comments: