Implementation of the Data Seal of Approval

The Data Seal of Approval board hereby confirms that the Trusted Digital repository CINES : Long-term Preservation Platform (PAC) complies with the guidelines version 2014-2017 set by the Data Seal of Approval Board.
The afore-mentioned repository has therefore acquired the Data Seal of Approval of 2013 on March 6, 2015.

The Trusted Digital repository is allowed to place an image of the Data Seal of Approval logo corresponding to the guidelines version date on their website. This image must link to this file which is hosted on the Data Seal of Approval website.

Yours sincerely,

The Data Seal of Approval Board

Assessment Information

Guidelines Version: 2014-2017 | July 19, 2013
Guidelines Information Booklet: DSA-booklet_2014-2017.pdf
All Guidelines Documentation: Documentation

Repository: CINES : Long-term Preservation Platform (PAC)
Seal Acquiry Date: Mar. 06, 2015

For the latest version of the awarded DSA   http://assessment.datasealofapproval.org/seals/
for this repository please visit our website:

Previously Acquired Seals: Seal date: March 15, 2011
                                    Guidelines version: 2010 | June 1, 2010

This repository is owned by: **CINES**
                                 950 Rue de Saint Priest
                                 34097 Montpellier
                                 Hérault
                                 France

                                 T +33467141414
                                 F +33467523763
                                 E massol@cines.fr
                                 W https://www.cines.fr/

Assessment

0. Repository Context

Applicant Entry

Self-assessment statement:

PAC (Plateforme d'Archivage au CINES) is a platform dedicated to digital long-term preservation, built in 2006, and initially assigned to the conservation of digital thesis. Since then, PAC extended its missions with much more archiving projects related to the community of higher education and research (like universities' libraries), preserving digitized publications and scientific datasets.

PAC is developed by the Département Archivage et Diffusion (DAD, or Archiving and Diffusion Department), unit of the Centre Informatique National de l'Enseignement Supérieur (CINES), a public establishment under the administrative supervision of the French Ministry of Higher education and Research.

The platform is in compliance with the OAIS Model (three entities : transfer, storage and communication) and the metadata required for the description of the data are based on the Dublin Core Schema (see guideline 3).

For more informations, see also :

https://www.cines.fr/archivage/ (03.02.2015)

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

1. The data producer deposits the data in a data repository with sufficient information for others to assess the quality of the data, and compliance with disciplinary and ethical norms.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

The repository doesn't deal with data producer directly, but contracts with entities (government institutions, university libraries, etc.) responsible for gathering and/or digitizing data for which quality has already been validated (e.g. PhD theses, publications, etc).

Before the submission, the repository assists the data producer in selecting the appropriate information which will be used as metadata; the repository also validates the quality of the data to be preserved as only scientific and technical data are eligible for preservation.

At the time of submission, metadata and data are validated against defined criteria: compliance to a generic structure of submission information package (SIP), validation of file formats, etc.

Functional and technical documentation is available at the repository

(https://alfresco.cines.fr/alfresco/d/a/workspace/SpacesStore/6f5c38da-69e5-4d67-8f9c-ecd72d094534/PAC%20-%20Sp%c3%a9cificatio
(03.02.2015)

and

https://alfresco.cines.fr/alfresco/d/a/workspace/SpacesStore/3005eac0-dfbd-4d71-b936-39b28daf0e8e/PAC%20-%20Sp%c3%a9cification (03.02.2015))

to provide assistance to the data producer during the different steps of the submission (building of the information package, transfer, control, etc);

Here is a translation in english of each table of contents (with summaries) of the documents above:

Functional specifications:
https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/33920989-dd10-4ca6-b2c4-1c56604a1ee9/FunctionalSpecifications_resume.p (19.02.2015)

Technical specifications:
https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/726b74e2-9590-4dbc-9a40-a3a1c9aeddd6/TechnicalSpecifications_resume.p (19.02.2015)

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

2. The data producer provides the data in formats recommended by the data repository.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

A limited set of file formats has been selected by the repository to ensure full characterization of the files provided by the data producer and allow future format migrations. An online tool is also available for the repository to check the correctness of the files to archive (see FACILE).

Only data in preferred formats is accepted on the data repository, but new data formats can be added to the list of preferred file formats pending validation by an internal committee of experts.

More information :

Presentation of the CINES expertise : https://www.cines.fr/en/long-term-preservation/expertises/ (03.02.2015)

List of archivable formats :
https://www.cines.fr/en/long-term-preservation/expertises/formats-expertise/archiving-format-list/ (03.02.2015)

Online tool FACILE : http://facile.cines.fr/ (03.02.2015)

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

3. The data producer provides the data together with the metadata requested by the data repository.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

A generic structure of submission information package (SIP) has been put together and combines data along with metadata to be preserved. Any transfer must comply with this structure, otherwise it is rejected. The XML schema for the SIP is available here: http://www.cines.fr/pac/sip.xsd (03.02.2015)

Data format required by the repository is based on a list of preferred formats (see guideline 2).

Metadata format required by the repository is based on DCMI standard.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

4. The data repository has an explicit mission in the area of digital archiving and promulgates it.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

The repository has two main activities – as defined by the mandate given by the French ministry of Higher education & Research: high performance computing and long-term preservation of electronic documents.

A law published on August 7th, 2006 also reinforces the mission of the repository, stating that it is the official preservation centre for electronic PhD theses.

The repository promotes its activities in the domain of the preservation of digital documents through various working groups and seminars.

In December 2010, PAC has been given consent by the French National Archives to preserve digital material. In January 2014, this agreement has been renewed.

See also:

Mandate given by the French ministry of Higher education and Research:
https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/af694b04-f7e9-4404-890d-74842a1adbc7/lettre-DGRI-DGES-missions-du-C
 (03.02.2015)

Law of August 7th, 2006:
https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/325f4c83-b4f6-410a-930d-12684799fbd0/Arrete_du_7_aout_2006_version_c
 (03.02.2015)

Presentation of the platform: https://www.cines.fr/en/long-term-preservation/production-platform-2/ (03.02.2015)

French National Archives agreement:
http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000023278648&dateTexte=&categorieLien=id
 (03.02.2015)

Agreement renewal:
http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000028588141&fastPos=1&fastReqId=1351866607&categorieLien=
 (03.02.2015)

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

It would be helpful to note whether the archive has discussed succession planning for its digital assets. That is, if support from the ministry is withdrawn, what is the plan for the archival resources?

5. The data repository uses due diligence to ensure compliance with legal regulations and contracts including, when applicable, regulations governing the protection of human subjects.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

The repository is under the trusteeship of the French ministry of Higher education and Research. It is legally the official preservation centre for electronic PhD theses.

The repository proposes two types of contract, one for the long-term preservation of data and the other one for the temporarily preservation of public records (documents which have to be transferred to a state archives service at the end of the preservation period defined in the contract). Any data producer who deposits on the repository is under one of these contracts, in which roles and responsibilities of the parties involved (data producer, repository, data users) are defined. See guideline 9 to have a presentation of a contract (translation of the clauses' titles).

The repository conditions of use are published on the repository website (https://www.cines.fr/wp-content/uploads/2014/06/CHARTE0911201.pdf 02.03.2015), the interface specifications are also available online. Here's a translation of the main points of the conditions of use: https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/b195cb74-e153-4dca-8ee0-291819290f9d/condition_of_use.pdf (02.03.2015)

Recurrent reviews are scheduled with data producers to ensure the repository conditions are complied with.

A full time archivist also ensures that the repository processes and systems comply with national and international laws.

See also: https://www.cines.fr/archivage/ (03.02.2015)

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

6. The data repository applies documented processes and procedures for managing data storage.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

A document – "Politique d'archivage du CINES" – summarizes the repository strategy and objectives in terms of preservation. This is the basis for the risks analysis. A risk management plan has been put in place to ensure that events which could impact the repository preservation strategy are identified and managed.

There are also documented processes and procedures in place to ensure quality management for the storage of data: authentication of users through LDAP catalog, multiple on-line copies, tape backups, file integrity checks using hashing algorithms, disaster recovery plan. Online monitoring tools (Nagios, Cacti) have been put in place to check application and storage availability.

Since October 2012, data are replicated in the Centre de Calcul de l'IN2P3 (Institut national de physique nucléaire et de physique des particules), in Villeurbanne ; this institution, to reinforce preservation security, provides storage equipment for the archived data.

Presentation of the risks management plan:

https://www.cines.fr/en/long-term-preservation/production-platform-2/quality-solution/risks-management/

Risks management plan (example):

https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/048f543a-8d9d-4b08-ae68-0dccc99a65ad/PAC_Plan_gestion_risques_2014.x
 (03.02.2015)

Logical architecture:

https://www.cines.fr/en/long-term-preservation/production-platform-2/technical-responses/logical-architecture/

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

7. The data repository has a plan for long-term preservation of its digital assets.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

The "Politique d'archivage du CINES" document, which summarizes the preservation strategy of the repository can be accessed here :

https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/2422e254-ff26-4ffa-a7b7-c37498987343/CINES-DAD-PA.pdf
(03.02.2015)

Translation of the table of contents of the Archiving Policy (document above):
https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/3e45fa2f-97bb-4489-8f3e-68bf1dc6eee5/archiving_policy.pdf
 (24.02.2015)

A risk management plan has been put in place to identify and mitigate risks which could impact the ability of the repository to guarantee durable archiving.

A preservation plan is also being put in place to document technology watch, evaluate obsolete and emerging file formats, define physical & logical migrations procedures, etc.

The business processes for file format technical watching (obsolescence, etc) and file format migration are available here :

https://alfresco.cines.fr/alfresco/d/a/workspace/SpacesStore/21035109-9e6b-4845-a16b-7f3a0dbdc660/Expertise%20formats.pdf (03.02.2015)

https://alfresco.cines.fr/alfresco/d/a/workspace/SpacesStore/5feab7d9-fe4e-4b34-bb54-bc11df55e1f2/Migration%20logique.pdf (03.02.2015)

To comprehend the business processes, here's a link to a presentation of the rules defined by the CINES for the conception of its own metamodel. It presents the links between processes and sub-processes, the signification of the graphical elements...

https://www.cines.fr/en/long-term-preservation/expertises/business-process-managment/metamodel/ (24.02.2015)

Translation of the processes headings (only for the macro parts): https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/4adfc9fc-dc71-400f-a8b3-e393b2c06623/processus_explanation.pdf (24.02.2015)

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

8. Archiving takes place according to explicit work flows across the data life cycle.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Any data transferred has to comply with the following criterions: the SIP is well-formed and sticks to the structure defined by the repository; all files provided are compliant with the specifications of their format.

All workflows, selection process and document lifecycle are documented in the functional specifications of the repository. See
https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/6f5c38da-69e5-4d67-8f9c-ecd72d094534/PAC%20-%20Sp%C3%A9cificatio
 p 21-23 (03.02.2015)

Decision making processes have been documented for the file format expertise and watching, as well as for file format conversion, data access, etc. See
https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/8353d8d2-0757-4a26-99cc-1eb2a2b4f500/PAC_Access_Process_Flow.pdf
 for an example for the access process (03.02.2015).

See also
https://www.cines.fr/en/long-term-preservation/production-platform-2/technical-responses/logical-architecture/
(03.02.2015) for a description of the transfer process between the data producer and the data repository.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

The Logical Architecture page is very helpful and descriptive about CINES processes and procedures.

9. The data repository assumes responsibility from the data producers for access and availability of the digital objects.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Any data producer who deposits on the repository has a contractual agreement signed with the repository, in which roles and responsibilities of the parties involved (data producer, repository) as well as the data user community are defined. This agreement includes the nomination of an executive board in charge of periodic reviews & reports.

A contract is divided in 21 clauses :

Clause 1 : Definitions

Clause 2 : Object of the contract

Clause 3 : Execution terms and conditions

Clause 4 : Data property

Clause 5 : Data integrity and authenticity

Clause 6 : Responsibility of the Repository

Clause 7 : Limits of the Repository responsibility

Clause 8 : Ressources commitment

Clause 9 : Audit

Clause 10 : Retention period

Clause 11 : Data access

Clause 12 : Replacement of the Transferring agency

Clause 13 : Data restitution and destruction

Clause 14 : Financial terms

Clause 15 : Length of service

Clause 16 : Case of force majeure

Clause 17 : Insurance

Clause 18 : Invalidity

Clause 19 : Contractual documents

Clause 20 : Litigation

Clause 21 : Appendices

appendix 1 : Terms of service

-

appendix 2 : Archiving policy (see guideline 7)

-

appendix 3 : List of the file formats accepted

-

appendix 4 : Functional and technical specifications (see guideline 1)

-

appendix 5 : Liaison committee composition

-

appendix 6 : List of the essential informations which have to be preserved and composition of the experts committee, in case of a file format migration

-

appendix 7 : Targeted community and knowledge base identification

-

appendix 8 : Nature of the archives and maximal volumetry

-

appendix 9 : Communicability policy

-

To date, the only data users accessing the data repository are the data producers themselves – making the repository a dark archive.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

10. The data repository enables the users to discover and use the data and refer to them in a persistent way.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

The data repository provides access tools (search engine, etc) to browse the catalog of archived data and request copies of data sets.

The data provided is either the original version or the latest version of the archived documents (in the event migrations have occurred).

A search engine allows data users to find data by querying against the 50 metadata qualifying a document (see screenshots at

https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/98e31240-8fd6-479f-b5bd-6efb46ceedee/PAC_webtool_screenshot.pdf (03.02.2015)

Each data set has a unique identifier based on ARK so that it can be referenced in other publications or web pages.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

11. The data repository ensures the integrity of the digital objects and the metadata.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

The repository stores the electronic signature (SHA-256) of any data stored, and periodically checks that this signature remains the same. This applies to the files containing the metadata as well.

The ingest process also allows the data producer to provide an electronic signature (MD5, SHA-256, SHA-1) for each documents transferred on the repository. The repository will compute a new checksum with the same algorithm and compare it to the initial one to detect potential corruption during the transfer.

Availability is monitored using supervision tools (Nagios, Cacti).

The set of metadata defined to qualify documents includes version numbers to allow multiple versions of the data.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

You might link to the Logical Architecture page here.

12. The data repository ensures the authenticity of the digital objects and the metadata.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Authenticity and integrity policies are described in the contract between the data producer and the repository (clause 5). It says that the Transferring agency commits to deposit in the Repository the original and authentic copy of the data (the file format must be in compliance with the list given in the third appendix of the contract). The data have to be deposited with all the metadata required for its exploitation, like identification and traceability metadata. If not, the Repository won't assure the preservation of the data's authenticity and integrity.

When migrating data to an emerging format, links to the previous versions are maintained. Any operation on the data (metadata updates, migration, access, etc.) is logged. All repository logs are self-archived in the repository.

A brief explanation of the data migration, available on the CINES' website: https://www.cines.fr/en/long-term-preservation/production-platform-2/916-2/format-conversions/ (25.02.2015)

Versioning of data is permitted using metadata and unique identifiers (ark identifiers) provided by the repository, as well as relations between datasets. Relations are controlled after the submission to ensure the referred dataset does exist.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

13. The technical infrastructure explicitly supports the tasks and functions described in internationally accepted archival standards like OAIS.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

The technical infrastructure complies with the OAIS model – it is made of three logical servers (ingest, storage, access) on which the different functional entities are deployed.

See also:

Presentation of PAC logical architecture:
https://www.cines.fr/en/long-term-preservation/production-platform-2/technical-responses/logical-architecture/
(03.02.2015)

A brief presentation of the OAIS model:
https://www.cines.fr/en/long-term-preservation/a-concept-problems-2/reference-model-oais/ (03.02.2015)

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

14. The data consumer complies with access regulations set by the data repository.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

The repository access module can be accessed from the following link :

http://pac.cines.fr:9000/webarcsys/001-frmIdent.jspa (03.02.2015)

Though, there are security restrictions based on the IP address; in order to get access to the link above, auditors should contact the repository administrator.

Screenshots of the access module can be accessed here :

https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/98e31240-8fd6-479f-b5bd-6efb46ceedee/PAC_webtool_screenshot.pdf (03.02.2015)

At present, the data producers have their own platform to allow access to the information to users. Thus, the access module on the repository is only used by the data producers to get the archived copy of their data document, should they loose the one they had on their own platform – the repository is considered a dark archive.

In any case, the contract between the data producers and the data repository defines the community and the rules to deposit and access data. The access module will conform to this agreement.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

15. The data consumer conforms to and agrees with any codes of conduct that are generally accepted in the relevant sector for the exchange and proper use of knowledge and information.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

0. N/A: Not Applicable.

Self-assessment statement:

All data consumers are authenticated when accessing the repository catalog and have accepted the conditions of use as a prerequisite to get an account to browse it.

See:
https://alfresco.cines.fr/alfresco/d/d/workspace/SpacesStore/b195cb74-e153-4dca-8ee0-291819290f9d/condition_of_use.pdf

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

16. The data consumer respects the applicable licences of the data repository regarding the use of the data.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

0. N/A: Not Applicable.

Self-assessment statement:

At present, the data producers have their own platform to allow access to the information to users. Thus, the access module on the CINES repository is only used by the data producers to get the archived copy of their data document should they loose the one they had.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments: