Assessment Information

[CoreTrustSeal Requirements 2017–2019](#)

Repository:	International Institute of Social History (IISH)
Website:	https://iisg.amsterdam
Certification Date:	13 July 2020

This repository is owned by: **Royal Academy of Sciences (Koninklijke Academie van Wetenschappen, KNAW)**

International Institute of Social History (IISH)

Notes Before Completing the Application

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

True

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, Library/Museum/Archives, Other (Please describe below)

Brief Description of Repository

The International Institute of Social History (IISH) is a research and cultural heritage institute in the field of global history of labour and labour relations (see also R1). The institute acquires, manages and preserves archives, library material and audiovisual material in this field, as well research data. The IISH is a 'private archive' and not a state archive. It has no legal deposit function and materials are collected outside the public sphere of government. Archives can be acquired from organisations (like for instance trade unions or activist groups) or private persons. Likewise, the subject matter of archives can be about organisations or individuals (as for instance Karl Marx, Mikhail Bakunin or Rosa Luxemburg).

Brief Description of the Repository's Designated Community.

- Researchers (mostly students and scientists from the humanities and social sciences and research journalists)
- General public with (professional) interest in the IISH collections
- Archival donors

Level of Curation Performed. Select all relevant types from:

C. Enhanced curation – e.g. conversion to new formats; enhancement of documentation

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

Comments

Reviewer 1

Comments:

Reviewer 2

Comments:

Outsource Partners. If applicable, please list them.

The IISH uses the digital storage services of the KNAW (Royal Netherlands Academy of Arts and Sciences, see point 2 directly below) ICT services for primary storage of the digital collections. This is formalised in an SLA between the KNAW ICT services and the IISH. The KNAW outsources the storage to the datacenter 'Vancis' (https://vancis.nl/). For secondary storage the IISH uses the services of 'SURFsara' data storage (https://www.surf.nl/en/services-and-products/data-archive/index.html). Both parties are closely linked to the Dutch university sphere in which security, trustworthiness and authenticity of data are of eminent value and have to meet strict demands on information security. Both parties meet the relevant ISO 27001 standard family (https://en.wikipedia.org/wiki/ISO/IEC_27001). 'Vancis': https://vancis.nl/over-vancis/certificeringen/ . 'Surf': https://www.surf.nl/en/services-and-products/data-archive/data-security-and-privacy/index.html.

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

Other Relevant Information.

1. In a small selection of CTS criteria a distinction is made between the IISH collection data and IISH research data. By the first is meant all data concerning the archival, library and audio-visual collections (which can either be digitised or born digital materials). By the second the data which are the result of IISH initiated research.

2. Important to mention is that IISH is a Royal Netherlands Academy of Arts and Sciences (KNAW) institute. This means that part of the IISH services and policies are dependent on broader KNAW policies. Responsibility for collection management is however in the hands of an independent body - the IISH Foundation. Ownership of the research data that are created by IISH researchers lies with the KNAW Institute IISH. Since 2017 the IISH is also part of the Humanities Cluster (HUC) - an alliance between the KNAW institutes IISH, the Meertens

Institute (https://www.meertens.knaw.nl/cms/en/) and Huygens-ING institute (https://www.huygens.knaw.nl/?lang=en). This alliance has a goal to stimulate cooperation between these institutes and promote innovation on the terrain of technical infrastructure and digital humanities research. This wider KNAW and HUC context gives the IISH more backbone/elbow room concerning issues like office IT, information security, digital humanities tools and research and knowledge management. When relevant this wider context will be mentioned in the criteria below.

3. The IISH uses Archivematica - an open source digital archiving workflow software which follows OAIS principles (https://www.archivematica.org/en/) - to manage the transfer and ingest of digitized and born digital content and supports the preservation watch function of the digital repository. In 2020 Archivematica will also be implemented to archive the IISH research data. Archivematica is therefore a core application in the IISH digital archiving process. Since 2019 the IISH is also part of the Archivematica Support Group which supports and monitors the developments of new features and contributes to the Archivematica development roadmap.

4. Important information for the reviewers:

a.Almost all documentation mentioned in the requirements is found on the public part of IISH Confluence pages: https://confluence.socialhistoryservices.org/display/CTS/Certification+of+IISH+Digital+Repository+Home. Where this is not the case this is mentioned explicitly.

b. The plain text of this application form can also be read in a formatted and slightly easier to read version on the IISH confluence: https://confluence.socialhistoryservices.org/display/CTS/Core+Trust+Seal+%28CTS%29+certificering.

c. The DOI mentioned in the organizational profile only points to the research data archive of the IISH. See also point 1 directly above.

d. The plain text of this application form can also be read in a formatted and slightly easier to read version on the IISH confluence: https://confluence.socialhistoryservices.org/display/CTS/Core+Trust+Seal+%28CTS%29+certification.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

ORGANIZATIONAL INFRASTRUCTURE

I. 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

Response:

In the IISH mission statement is as following: "The IISH is a unique institute, serving science and society on a global scale. At an international level, we generate and offer reliable information and insights about the (long-term) origins, effects and consequences of social inequality.

To promote this, we form an international hub for social historians worldwide. We offer and produce historical sources and data, facilitate social-history research and collaborate internationally in ground breaking research projects.

Moreover, by preserving the heritage of often oppressed social movements, the Institute serves the quality of the world's memory. With our work we hope to contribute to a vibrant civil society." Source: https://iisg.amsterdam/en/about/mission.

In this statement especially the phrases "We offer and produce historical sources and data" and "Moreover, by preserving the heritage of often oppressed social movements, the Institute serves the quality of the world's memory" refer to long term stewardship of the institute.

In the IISH Strategic Plan 2019 - 2023 (https://confluence.socialhistoryservices.org/download/attachments/32703335/strategic_plan_2018-2023_iisg_engels.pdf) the ambition for long term stewardship and a trustworthy digital repository is addressed (in paragraphs 3.1, 3.2 and especially measure 5.2 (page 26): "We develop an internationally accredited repository for the sustainable storage of digital collections."). Also in the Collection Policy 2015-2020 (https://confluence.socialhistoryservices.org/download/attachments/42568284/20180110_collectieplan_rev_2018.pdf?api=v2) clearly points towards the ambition of long term digital

stewardship, while relating to the overall IISH mission. In the plan it is stated as such (page 5): "in 2020 all digital objects will be stored in a Trusted Digital Repository". Paragraph 3.8.3 (page 23) contains the most relevant information about long term access to digital collections of the IISH.

It is also important to mention that between 2017-2025 the IISH will receive ample funds from the KNAW to make the organisational and infrastructural transition from paper to long term digital archiving. This is a clear mandate from the umbrella organisation of the Institute to invest in the people, knowledge, software and hardware needed to help with this transition. The ideas behind this can be found in the (Dutch) application document Van archief/bibliotheek naar humanities research infrastructure, Aanvraag vernieuwingsinvestering KNAW (From archive/library to a humanities research infrastructure, application renewal investment KNAW): https://confluence.socialhistoryservices.org/download/attachments/42568284/20161203_investeren_in_collecties_intranetversie_0.pdf.

II. 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

Response:

Metadata records about collections and collection items contain clear information regarding access, restrictions and permissions, copyright holders and - when relevant - other licenses such as Creative Commons (CC). Based on this information internal procedures and automatic responses take place. The metadata themselves are published under a CC0 license. Research data are always published with a CC license.

Some examples:

- Example of an open archive ("not restricted"): http://hdl.handle.net/10622/ARCH00860
- Example of a (partly) restricted archive: http://hdl.handle.net/10622/ARCH00200
- Example of photos distributed under BY-NC-SA 4.0 license: http://hdl.handle.net/10622/COLL00293
- Example of an IISH dataset published under a CC0 license: http://hdl.handle.net/10622/BYFPBK
- Example of an IISH dataset published under a CC BY-NC license: http://hdl.handle.net/10622/D5WXZZ

All descriptive metadata are available under a CC0 license: https://iisg.amsterdam/en/collections/use/api-linked-data.

In the case of a restricted collection, data non-compliance is no issue as these collections are simply not available to the users. In the unlikely case our systems or processes fail we shall act as promptly as possible and repair breaches quickly. In the case of copyright violation we remind users of their legal obligations. See our 'Policy in case of noncompliance of licenses and copyright violations': https://confluence.socialhistoryservices.org/display/CTS/Policy+in+case+of+noncompliance+of+licences+and+copyright+violations.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

III. 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

Response:

The IISH has a long-standing tradition (since 1935) in preserving and ensuring access to its (world famous) collections. It is at the core of the institute's activities and a guiding principle for all policy and strategy documents. The plans for ongoing access to collections and research data can be found in multiple places but are most clearly stated in the Collection Policy Plan 2015 - 2020 (https://confluence.socialhistoryservices.org/download/attachments/42568284/20161203_investeren_in _collecties_intranetversie_0.pdf). A citation from the IISH collection plan:

" 1. The IISH aspires to be (in addition to being a first-class research institute) a world-leading centre for preserving and making available sources in the field of socio-economical history, in particular the world history of labour and industrial relations.

2. The collection is primarily formed, made accessible and made available for the purpose of scientific use, but it simultaneously forms a socially valuable heritage collection relevant to a broader public.

3. Highly decisive for realising the aforementioned ambitions are:

a. the volume and quality of the collection;

b. the dependability of safe storage, taking into account security, privacy and the trust of the archival creators.

c. the level of the archive's discoverability and findability;

d. the quality and openness of the services used in making the archive available (online);

e. the measures we take to maintain the collection;

f. the extent to which the IISH leads nationally and internationally in using cutting-edge methods."

IISH collections are given on loan or as a deposit to a separate and independent foundation - the IISH Foundation. The IISH foundation is responsible for the collection management, the KNAW institute IISH is responsible for carrying out this management. For research data, the KNAW institute IISH is the owner and is responsible for carrying out this management. Being a KNAW institute and being part of the HUC (see R0), guarantees the continuity of the IISH and its collections. The sharing of (for instance IT) services and knowledge (i.e. information security) between the KNAW and

HUC institutes also creates the environment for the IISH to perform at a higher level that it could do on its own.

Since the foundation of the IISH, more than 100 million euro has been spent on the collections and the infrastructural facilities to acquire, manage and preserve them. Yearly, around 60 percent of the budget for IT investments has been spent on the sustainability and development of the collection and data infrastructure that supports the work-processes, including digital preservation. It is unlikely that funds will be spend on something completely different or that this funding will significantly drop to a point where this is a serious risk.

In the formal contract between IISH and KNAW the importance of the collections and the need for its availability to researchers is recognized, ensuring and acknowledging the role of the IISH in the continuing process of collecting, processing and making available of the collections as one of its primary tasks. It is also stated that the IISH receives financial means from the KNAW – which itself receives the financial means from the Dutch state – to enable the institute to carry out its tasks.

In the unlikely event that funding should run out, the IISG will turn to the KNAW to come up with means to continue taking care of both collections and research data.

From the perspective of risk management, a yearly updated risk assessment document ('Risicomanagementplan 2020 Humanities Cluster', meaning 'Risk managementplan', latest version September 2019 - available on request) offers important evidence regarding the long-term conservation of the collection (see also R 16). In this document the risk of data loss and cyber-attacks is indicated and insight is given on how these risks are controlled.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

IV. 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

Response:

Looking at the ethics concerning the collection in general, two things are important to mention:

- The IISH adheres to the ICA Code of Ethics (http://www.ica.org/en/ica-code-ethics) from which article 7 ("Archivists should respect both access and privacy, and act within the boundaries of relevant legislation") is especially important for this requirement.
- Every IISH employee signs a confidentiality agreement in which the ethical conduct as required by the IISH is fully described.

Looking at the different phases of the archival process the following criteria are important when talking about collection ethics:

Creation of archives:

- The IISH collects private archives and therefore has no influence on the creation of archival records.
- The institute does not request confirmation from the archival donor that the data was collected or created in accordance with legal and ethical criteria. The IISH collects materials on labour and labour relations. In some of the countries of origin, the data could be considered to be illegal or unethical (i.e. reports on working conditions in factories, on child labour etc).
- Another form of creating digital collections is by digitising existing IISH collections. The possible disclosure or copyrights issues of the digitised collections are taken into account before starting the digitising process.

Curation/collection of archives:

- New acquisitions of archives and research data sets are accepted on the basis of individual agreements with the donors. Access regulations are an explicit part of the agreement, and are agreed upon by both donor and IISH staff. There is a standard contract available but for every archival donor contracts are tailor made.
- The IISH will invest in a clear and secure procedure for transfer of collections from the donor to the institute.
- During talks with the archival donor the most sensitive parts of the archive are identified. This might mean that these parts are not transferred to the IISH or that these parts will be closed for a certain period of time (see below). This has been common practice since the institute started archiving collections in 1930's and worked well for both paper and digital collections.

Pre-ingest (quick scan and appraisal and selection):

During appraisal and selection of the to be ingested archives sensitive parts might be identified and, in consultation with the archival donor, be de-selected or closed for a certain period of time (see below under 'access'). The first step in this proces is the so called 'quick scan' of the new archive by the Collection Development employee who has acquired the archive. This procedure is described in the document First inspection/quick scan of a new archive procedure: https://confluence.socialhistoryservices.org/pages/viewpage.action?pageId=60653936. On the basis of this quick scan it can be decided to do a more elaborate appraisal and selection session. During the quick scan the archive is also checked for possible privacy or other disclosure issues. This is done on a high, general level. In practice this means that not every individual file will be checked but instead a generated file and directory list is scanned to try and identify any of these issues.

Ingest:

- During ingest born digital archives are not scanned any further for privacy or other disclosure issues. This is a conscious choice because of:
a. The sensitive parts of the archive are - on a high/medium level - detected during talks with the archival donor (see directly above).
b. Privacy or disclosure issues can also be identified during appraisal and selection.
c. The (highly) restricted access to born digital archives (see below under 'access').
d. The nature of born digital archives: the discovery of all privacy related information within unordered born digital content is highly problematic. This would need the appliance of named entity recognition on document level which for may reasons would be hard to implement (see directly below) and certainly not perfect (in the sense that it would find all such information).
e. Technical reasons: full text indexing (which you would need to check for privacy or other disclosure issues on document level) of the born digital content is not yet part of the ingest procedure (or any later part of the workflow).
- During ingest checks for viruses and malware are routinely applied (see also R12).

Access:

- To disclose archival collections the appropriate access policy is included in the archival description. Some (digital) collections are only accessible after a formal request and can only be viewed in the reading room (example: http://hdl.handle.net/10622/ARCH00200, see under "tab access and use").
- Collections with highly sensitive materials are not accessible at all (for the period of time which has been agreed upon with the archival donor).
- For now born digital archival collections will only be available in the reading room and only after explicit permission from the archival donor (see above). If access is given researchers do have to conform to strict rules concerning the re-use of personal information. See the Agreement end user of IISH born digital collections:

https://confluence.socialhistoryservices.org/display/CTS/Agreement+end+user+of+IISH+born+digital+collections.
- All procedures regarding the disclosure of material, restricted access and permissions are documented and fully described for internal use on the central knowledge base (Confluence) collection platform (CODI: https://confluence.socialhistoryservices.org/display/coditest/Home - open on request). See especially the procedures for archives which need special permission before gaining access ("Toestemmingen" - permissions: https://confluence.socialhistoryservices.org/display/CTS/Toestemmingen) and for copyrights (https://confluence.socialhistoryservices.org/display/CTS/Copyright). This documentation is all in Dutch.
- A general copyright statement is published on the website (https://iisg.amsterdam/en/collections/using/reproductions - under header 'copyrights') and when the copyright holder is known a copyright statement is included at item level.

In the IISH 'risk management plan' document (see R3) there is ample attention given to the risks of disclosure and what is done to mitigate these. Also the legal department of the KNAW has checked all contracts and procedures. Most important to mention here is that for born digital content the standard procedure is that only an AIP is created (and the AIP storage not connected to the outside world) but not automatically a DIP (which can be available in the reading room or - in seldom cases - online) as well. Only after a conscious decision by the collection acquisition and public services staff a DIP is made 'manually'.

With the implementation of the European legislation concerning General Data Protection Regulation (GDPR) (May 2018) the KNAW has taken the lead in making sure all KNAW institutes do comply with the GPDR. Some of the necessary provisions (KNAW privacy statement - https://www.knaw.nl/en/about-us/academy-privacy-statement -, organisational procedures, data protection officer etc.) have been taken at KNAW level and apply to all the institutes. As for the collections, the Department of Collections at the IISH has made an inventory of workflows where personal data is involved and has taken the necessary measures to ensure compliance. Regarding personal data that is present in the collections of the institute, a working group within the KNAW has been established to determine necessary measures and monitor developments. Institution-specific documentation on GPDR, such as the exceptions in collecting materials with personal data valid for scientific and historical research, and how these apply to the IISH, is available on Confluence.

Specifically for research data the following is relevant:

For research data, disclosure issues between depositor and IISH are laid down in the Provisions Data Deposit Agreement document (https://confluence.socialhistoryservices.org/pages/viewpageattachments.action?pageId=32703335&preview=/32703335/48988571/20160900_provisions_data_deposit_agreement.docx). The Agreement contains the following: "The Depositor declares that the dataset contains no data or other elements that are, either in themselves or in the event of their publication, contrary to Dutch law." In other words: the same checks for disclosure issues apply for research data as for collection data.

Comments:
Accept

V. 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:

4 – The guideline has been fully implemented in the repository

Response:

- The repository is hosted by the IISH. The IISH is an organisation with a clear profile (see mission statement mentioned in R1) and is part of the KNAW and the KNAW Humanities Cluster (https://huc.knaw.nl/). The collections are owned by, or given in loan to, the IISH foundation - except for research data which is owned by the KNAW institute IISH. See also R0 and the description on the IISH website about the organisation (with more information about the KNAW, scientific board, the KNAW Humanities Cluster, the IISH foundation and the board of directors) and staff - respectively: https://iisg.amsterdam/en/about/organization and https://iisg.amsterdam/en/about/staff.
- As the IISH is part of the KNAW and its collection is renowned, there are no signs of an end to the funding horizon. Yearly the IISH is funded by the government through the KNAW. Funding does not depend on a temporary budget and is guaranteed for as long as the IISH and its collections exist. Since 1935 approximately 100 million euro's have been spent on the collections and its infrastructure. Digital preservation is a crucial element in keeping the collections available for research. This is present in the mission of the IISH as well as in strategic and policy plans. See for instance Strategic Plan 2018-2023 (goal 3.1, 3.2 and 5.5.2): https://confluence.socialhistoryservices.org/download/attachments/32703335/strategic_plan_2018-2023_iisg_engels.pdf.
- The description of tasks and expertise needed in connection to the repository are part of the description of the workflow processes (R12 Workflows) and functional descriptions of jobs. Roles and responsibilities are also described in the Digital Preservation Policy, chapter 13: https://confluence.socialhistoryservices.org/display/CTS/Digital+Preservation+Policy+201

9-2022#DigitalPreservationPolicy2019-2022-13.Organisation.

- As a middle sized institute, the "range and depth of expertise" is limited to the needs and mission of the organisation. This means that there is in-depth knowledge of the systems, software (like Archivematica) and workflows connected to the digital repository as all this is hosted by the institute itself. Also there is in-house knowledge of the OAIS model and specific preservation issues (i.e. file formats) related to the institute's digital collections. The ambition and expertise of the IISH do not extend to in-depth research on preservation issues or to ambitious software development of - for instance - digital preservation tools. This is clearly seen as the task of bigger institutes or collaborative efforts in national or international projects and of the open source communities.

- From a Dutch perspective the IISH takes an active role in the Dutch Digital Heritage Network (NDE http://www.netwerkdigitaalerfgoed.nl/). The NDE "is a partnership in the Netherlands that focuses on developing a system of national facilities and services for improving the visibility, usability, and sustainability of digital heritage". Together with the KNAW-HUC partners the IISH is a so-called 'hub' in the NDE network and by doing so is sharing and gaining knowledge.

VI. 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

Response:

- Aware of the challenges that come with digital preservation, the IISH appointed a digital archivist in 2016. The digital archivist (https://iisg.amsterdam/en/about/staff/robert-gillesse) is dedicated to all aspects of long term preservation within the institute. One of the tasks of the digital archivist of the IISH is to monitor technical developments concerning file formats, (pre-)ingest, preservation and access. He has a network of other national (see below) and international experts who he can consult. Other IISH staff monitor changes in organisational digital strategies and practices (of our archival donors), changes in metadata standards, storage and access systems, security and workflow management.
- The data officer (https://iisg.amsterdam/en/about/staff/richard-zijdeman) of the IISH monitors changes in research data and is deeply involved in Dutch and European projects concerning data research infrastructure (CLARIAH https://www.clariah.nl/over/wie-is-wie/wp4/richard-zijdeman and DARIAH).
- The IISH has in active role in the Dutch Netwerk Digitaal Erfgoed (NDE, network digital heritage). This network contains a nationwide pool of expertise on digital preservation (see also R5). The IISH digital archivist represents the KNAW-HUC in the domain group ('board') on digital preservation.
- The IISH also strengthens its expertise by being an active member- or partner of the following organisations: ICA (International Council of Archives http://www.ica.org/), BRAIN/KVAN (branch organisation and association of Dutch archives, https://www.kvanbrain.nl/), IALHI (The International Association of Labour History Institutions http://www.ialhi.org/) and OCLC research library partnership (http://www.oclc.org/research/partnership.html).
- As part of national and international research infrastructures the IISH is deeply embedded in the research community. Communication with the designated community is therefore an everyday occurrence. The IISH-HUC is also closely involved in many digital humanities initiatives and research-infrastructures like CLARIAH (i.a. as part of the board https://www.clariah.nl/en/about/organisation/board) an DARIAH (as national coordinator https://www.dariah.eu/network/partners-countries/netherlands/).

DIGITAL OBJECT MANAGEMENT

VII. 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

Response:

Data integrity checks by the IISH:

Fixity:

- When the archival donor delivers the digital collection in an Archival Bag (in which checksums are included) fixity is checked as soon as the bag reaches the institute. If not, the institute will create an archival bag with MD5 checksums after arrival which is then ingested. In that case any data corruption that took place before the bag reached the institute falls outside the responsibility of the institute.
- During the pre-ingest (resulting in a SIP) and ingest (resulting in an AIP and DIP) the checksums are validated by Archivematica.
- During the pre-ingest, Archivematica produces SHA-256 checksums (https://www.archivematica.org/en/docs/archivematica-1.7/user-manual/transfer/transfer/#process-transfer) for each of the files.
- After the creation of the AIP, Archivematica performs a bag check which includes a final fixity check.
- After the storage of the AIP the fixity is regularly checked by the IISH (https://confluence.socialhistoryservices.org/display/CTS/Fixity+checking).

The integrity of the bitstreams is guaranteed by the institute's data provider. An independent comparison by the system between metadata and hashed content is considered.

Version control:

- Original files and metadata persist in the AIP.
- Changes over time in file formats are recorded and stored in the AIP METS file using PREMIS events.
- Other than file format changes, AIPs are immutable; changes result in a new AIP and the relation is documented in the technical and structural metadata (METS).

Standards/conventions:

The IISH uses international standards for preservation of digital objects (OAIS, PREMIS, METS).

Data authenticity measures by the IISH:

- A strategy for the planned change of digital objects (through careful and logged migration steps, the original will always be kept) can be found in the IISH preservation policy 2019-2022 (https://confluence.socialhistoryservices.org/display/CTS/Digital+Preservation+Policy+2019-2022). The Preservation Policy states the following about the authenticity of the digital object: "the informational value within the digital object is first and foremost guaranteed by always keeping the original digital object. If this object threatens to go obsolete, the object is migrated to a preservation copy that - as much as possible - preserves the informational value of the original object. Also an important part of guaranteeing the authenticity of the object is the creation of technical metadata. This metadata gives in-depth technical information about the object and is the basis for possible future preservation actions. Technical metadata is acquired and stored (as PREMIS metadata in a METS wrapper) during the preservation workflow described below. The authenticity of a digital object can only be fully guaranteed if the file format can be identified."
- Provenance metadata are made in two senses:
a. The origin of the digital object: the context of the object is described in the archival description which is published on the IISH website.
b. Every step in the archiving workflow (audit trail) is logged by Archivematica in PREMIS metadata. Eventual migration to a new format - during the archiving process or at a later date - will be logged in the PREMIS metadata.
- The link between digital objects and metadata is maintained/guaranteed by the use of a PID (Persistent Identifier) through the Handle service (https://www.handle.net). See the IISH PID declaration & binding workflow for digitized and born digital collections (https://confluence.socialhistoryservices.org/pages/viewpage.action?pageId=48989155). See also CTS R13 for an in-depth description of the Handle service.
- Essential properties of categories of digital objects (for instance word processing files) are seen by the IISH as a somewhat problematic concept as these are very much dependent on the context. The institute rather talks about preservation intent which is described as followed in the Digital Preservation policy (https://confluence.socialhistoryservices.org/display/CTS/Digital+Preservation+Policy+2019-2022#DigitalPreservationPolicy2019-2022-5.3Preservationintent):"
a. All digital objects will be kept and (as much as possible) preserved in their original form.
b. But the IISH will, in the end, give priority to the informational value within these objects. This means that objects may be migrated to other file formats - as the original format is obscure and/or obsolete - as long as it can be guaranteed that the information within these files is still authentic.
c. These migrated files are called preservation copies.
d. As the proof of authenticity will, in some cases, be a challenge the original object will always be available as a fall back file.
e. The digital object is always shown within the right context and is findable through the correct contextual information."
- The identity of the IISH archival donors are known to the institute and agreements are formalised in a contract between donor and institute.

Specifically, for research data the following is relevant:

At the moment of writing research data are not yet being archived by the use of the Archivematica (workflow). This is planned for the beginning of 2020 in a separate Dataverse Archivematica integration project. When this work is finished all of the above will apply for research data as well. For now, research data are published online via the Dataverse platform and the preservation level is that of bit preservation. This means that research data are stored on a replicated storage system and the fixity of the files are regularly checked. See also the schematic view of the four main digital archival workflows:

https://confluence.socialhistoryservices.org/display/CTS/Scheme+of+four+main+IISH+digital+archival+workflows.

Reviewer Entry
Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

VIII. 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

Reviewer 2

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

Response:

- The IISH has a collection policy document (Collectiebeleid 2015-2020,

https://confluence.socialhistoryservices.org/download/attachments/42568284/20180110_collectieplan_rev_2018.pdf)

which guides the Institute in the acquisition of new archives. A well-defined collection profile ("labour related network organisations and persons involved in these organisations") assures that the Institute can make sharp choices when acquiring collections.

- As the IISH is a private archive the Institute is not in the position to enforce any criteria on the transfer of (digital) archives.
- The IISH strives to carry out the appraisal and selection in close dialogue with the archival donor. If, due to circumstances, this is not possible, the IISH will accept the collection as it is and will run a quick scan to see if and what appraisal and selection steps are necessary.
- During the reception and pre-ingest phase objects will be checked for - to the degree where the IISH can check these things - completeness and understandability.
- For reasons mentioned above the IISH cannot enforce the use of preferable file formats by the archival donor. All files will be ingested as such and for some file formats extra preservation and access copies will be made during the ingest process.
- To make sure our designated communities can use and understand the digital objects, the IISH uses the most common files formats for the dissemination of these objects, while offering metadata on how and why the choice for these formats was made. When permitted and technically feasible the user can also have access to the originally ingested file. See the IISH File format policy for born digital collections (https://confluence.socialhistoryservices.org/display/CTS/File+format+policy+for+born+digital+collections).
- When objects are delivered in an obscure file format which cannot be recognized by the (current) characterization tools (the IISH uses the PRONOM file format registry http://www.nationalarchives.gov.uk/PRONOM/Default.aspx within Archivematica) the files will be stored as such. For these files the authenticity cannot be guaranteed and only bit preservation is possible. As these files are marked as unrecognized, they might be re-ingested at a later date when characterization tools are updated.

Research data:

The appraisal of research data follow a slightly different route as they are all produced by IISH researchers themselves. That means that these data automatically fall within the collection profile (as IISH research is all done following the IISH research programme: https://iisg.amsterdam/en/research/programme). Checks for completeness and understandability (as the adding of descriptive metadata) are done before the data is published on the IISH Dataverse platform. IISH research data are published using open formats. As explained in R7 in 2020 research data will also be ingested via a full preservation Archivematica 'fuelled' workflow.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

IX. 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

Reviewer 2

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

Response:

The primary archival storage of the IISH is provided by KNAW ICT Services. The AIPs are stored in two different data centers - provided by Vancis - in the Netherlands. Because of the nature of the materials and the agreements with archival donors, storage under Dutch law is a strict requirement of the IISH when selecting archival storage. For additional safety, Vancis also provides a 30-day backup of the data. As an extra security measure, the AIPs are also stored on secondary storage provided by SURFSara, including a backup. The secondary storage system is only accessible as a WORM storage, which does not allow deletes and does not allow overwriting AIPs. In this sense, the AIPs are always stored in versions; a new version of the same AIP does not overwrite the old AIP, but writes a new version of the AIP. See document: Multiple copies and
backup: https://confluence.socialhistoryservices.org/display/CTS/Multiple+copies+and+backup.

As the data is replicated across two different storage providers and mirrored across various data centers, this allows for a high level of protection against data loss. See document: Security against data
loss: https://confluence.socialhistoryservices.org/display/CTS/Security+against+data+loss.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

X. Preservation plan

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

Compliance Level:

4 – The guideline has been fully implemented in the repository

Response:

The Digital Preservation Policy 2019-2022 is published on the IISH Confluence wiki (https://confluence.socialhistoryservices.org/display/CTS/Digital+Preservation+Policy+2019-2022).

There are two preservation levels which will be used in the IISH:

1. Full preservation level: next to fixity checks and multiple secure storage the objects that can be recognized by file format characterisation tools during ingest will be preserved on the highest level. A preservation policy exists for each category of files (text, spreadsheets, databases, email etc). See the File Format Policy: https://confluence.socialhistoryservices.org/display/CTS/File+format+policy+for+born+digital+collections
2. Bit preservation level: for files formats that cannot be recognized during ingest only fixity checking and secure/multiple storage is possible.

Collection data are preserved on 'full preservation level', research data - at the moment of writing - on 'bit preservation level'. In 2020, when research data will also be processed with the 'Archivematica workflow', the full preservation level will also apply for research data (see also R12).

For the contract between the archive and the archival donor/depositor a distinction has to be made between collection and research data:

Collection data:

- The standard and generic "act of deposit" between depositor and repository that is used by the IISH states that the Institute takes up the obligation to keep the collection in good condition ("The Stichting IISG undertakes to keep the archives in a good condition"). The contract offers the possibility to expand to a specific agreement between the repository and the archival donor.
- As to rights to transform the archive: In the standard contract one of the conditions states that the "appraisal of the archives can take place without the special permission of the depositor".
- As to the rights to give access and copy the archive the contract states the following: "The Stichting IISG will be permitted to make the Archives available to third parties through the channels commonly, including, but not limited to, reading rooms and website. The Foundation IISG will also be permitted to make copies of the Archives for conservation purposes and availability under the same conditions that apply for the original Archives".

Research data:

When the institute acquires research datasets from outside the institute a data deposit agreement is drawn up between the IISH and researcher/depositor: https://confluence.socialhistoryservices.org/download/attachments/42568284/2016090 0_-_data_deposit_agreement.docx . In this standard contract all provisions are laid down regarding the deposit of dataset. The provisions data deposit agreement (https://confluence.socialhistoryservices.org/download/attachments/42568284/20 160900_provisions_data_deposit_agreement%20%281%29.docx) contains the following relevant quote:"

a. The Repository shall ensure, to the best of its ability and resources, that the deposited dataset is archived in a sustainable manner and remains legible and accessible.
b. The Repository shall, as far as possible, preserve the dataset unchanged in its original software format, taking account of current technology and the costs of implementation. The Repository has the right to modify the format and/or functionality of the dataset if this is necessary in order to facilitate the digital sustainability, distribution or re-use of the dataset.
c. If the access category is Open Access for registered users "under any of its three conditions, as specified in the Data Deposit Agreement at the end of these Provisions, are selected, the Repository shall, to the best of its ability and resources, ensure that effective technical and other measures are in place to prevent unauthorised third parties from gaining access to and/or consulting the dataset or substantial parts thereof."

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
At the time of next renewal, please provide evidence that the Archivematica workflow has been fully implemented for research data as well.

XI. 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

Reviewer 2

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

Response:

Collection data:

- The IISH aims to give proper access to all collections by adding sufficient descriptive metadata so they can be easily found in the IISH internal and external search systems and engines. If the depositor has delivered metadata we will use that as a starting point for creating descriptive metadata. But, as noted earlier, as a private archive the Institute has no real influence on the use, the form and the amount of metadata the depositor delivers with the archive. In many cases the descriptive metadata will be created by the institute itself.
- The approach to the quality of collection describing metadata is pragmatical in the sense that the IISH follows standards that are widely used within the cultural heritage community. Findability and interoperability of the collections are key to the allocated metadata. Obvious standards used here are archival standards ISAD(G), EAD and the library standard MARC 21. Also the institute makes extensive use of well-established vocabularies like VIAF (http://viaf.org/) and AAT (http://www.getty.edu/research/tools/vocabularies/aat/), Library of Congress Subject Headings and GeoNames. These are at the basis of the search functionalities of the linked data powered IISH website.
- The institute uses standardized Open Source applications like Evergreen and Archivespace for descriptive metadata, which are built upon the internationally accepted metadata standards like MARC 21 and EAD.
- For technical and preservation metadata the PREMIS (embedded in METS metadata) is used. This metadata is generated by Archivematica during the different steps in the archiving workflow.
- Search results and records are always presented in context and with suggestions to related works. Example: https://hdl.handle.net/10622/3C26EA12-6914-4C3D-B3E6-D22EFF25B6E3.

Research data:

- The IISH uses the Dataverse platform for giving access to the research data: https://datasets.iisg.amsterdam/. It is one of the standard applications maintained by the newly formed 'Collections and data department". Metadata and files are part of the standard workflows, dedicated staff members will ensure the quality.
- Metadata and file formats are according to international standards. The DDI (Data Documentation Initiative) standard is used for describing the datasets.

XII. Workflows

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

Compliance Level:

3 – The repository is in the implementation phase

Response:

Collection data:

The IISH has two workflows for the ingest of collection data in the repository:

1. A workflow for born digital materials: in this - Archivematica supported - workflow the ingest of born digital materials into the digital repository is described. This involves - among other things - pre-ingest steps as the assignation of checksums, checks for viruses, file format characterisation and validation. If necessary, appraisal and selection are (as much as possible) done between the pre-ingest (Transfer) and the SIP creation steps. The most important steps during the ingest phase are the normalisation of files for access and preservation, the allocation of a Handle PID's to files and folders, the creation of the AIP and the DIP (when allowed, see R4) and the storage of these packages. All steps of the workflow are logged and documented in PREMIS metadata wrapped in METS. Documentation: Functional description of the workflow for pre-ingest and ingest of Born Digital Collections https://confluence.socialhistoryservices.org/display/CTS/Functional+d escription+of+the+workflow+for+pre-ingest+and+ingest+of+Born+Digital+Collections. Schematic representation of the born digital workflow: https://confluence.socialhistoryservices.org/pages/viewpage.action?pageId=48989289.
2. A workflow for digitized materials: in this workflow the ingest into the archive and dissemination of digitized collections is described. This involves - among other things - the checking of structural metadata, the creation of METS, creation of derivates, the allocation of Handle PID's and the allowed access status (resolution of the image the Institute has allowed for viewing). Documentation: https://confluence.socialhistoryservices.org/display/CTS/Flow+14 . This workflow will have migrated in the second half of 2019 to an Archivematica driven workflow, so the individual steps in the workflow can be better monitored, managed and changed.

Security and impact on workflows:

- In the dissemination part of the workflow for digital archival materials, privacy and copyrights status determine whether the archive is open or not and, if open, under which access regime the images (or otherwise) may be presented. On the highest level of the archival description under "access" the user receives a statement on the openness of the archive (restricted / not restricted). Under the Access and use tab the exact status can be found.
- Access regimes can be based on copyright and privacy laws and/or on the individual agreements with the archival donor.
- A general IISH copyright statement can be found on the website:
https://iisg.amsterdam/en/collections/using/reproductions.

Approaches towards data outside the mission/collection profile:

The method of, and depth of, appraisal and selection of born digital collections is an ongoing discussion and adaptive process because of the changing nature of the materials and the organisations and persons that provide the materials. Having said this, the following materials will be deselected as a rule:

- Computer programs
- ISO images of hard disks or otherwise
- Virus infected data
- Password protected files

In the near future also extra functionality will be developed within Archivematica for identification and removal of duplicates files.

Change management of the workflows:
Important changes of the two workflows will be logged and documented on the IISH Confluence pages that describe these workflows (see above).

Research data:

The archival workflow for research data is, for an important part, determined by the Dataverse platform (https://datasets.socialhistory.org/) that the IISH uses for dissemination of the datasets: http://dataverse-guides.readthedocs.io/en/latest/user/dataset-management.html. Files are stored on a file server (of which back-ups are made). At the moment of writing this means that, for the IISH datasets, only the integrity of the files can be guaranteed (bit preservation level - see R7). In 2020 the datasets will be ingested through an Archivematica workflow which means that the datasets will be archived on full preservation level (see R7). Research data sets uploaded in Dataverse will be transferred to the Archivematica workflow and follow the standard ingest. The data will be published if access is open. See the also scheme of the four main IISH digital archival workflows: https://confluence.socialhistoryservices.org/display/CTS/Scheme+of+four+main+IISH+digital+archival+workflows.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

XIII. 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

Reviewer 2

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

Response:

Collection data:

- The IISH collection can be discovered through the IISH website (https://search.iisg.amsterdam/) and portals in which the institute takes part - i.a.: Europeana (http://www.europeana.eu/portal/nl/search?q=europeana_collectionName%3A%282022086_Ag_EU_HOPE_IISHInternational%29), the Archives Portal Europe (https://www.archivesportaleurope.net/), the Social History Portal (http://www.socialhistoryportal.org/), Worldcat (https://www.worldcat.org/libraries/91211) and Archivegrid (https://researchworks.oclc.org/archivegrid/?q=contributor:301). An overview of all portals can be found here: https://confluence.socialhistoryservices.org/display/CTS/Externe+portals (in Dutch).
- As of 2018 IISH collection and research data can also be discovered as linked data: https://druid.datalegend.net/IISG/iisg-kg.
- Descriptive metadata are stored as MARC 21 (library and individual objects) and EAD (archive collections and multiple objects): https://confluence.socialhistoryservices.org/display/coditest/Beschrijving+collecties+-+metadata (Dutch documentation).
- The IISH offers a webservice (API) which enables querying of the collection metadata records through the SRU/SRW protocol and a harvesting of the records through the OAI-PMH protocol: https://confluence.socialhistoryservices.org/display/coditest/Metadata+harvesting. Technical REST API documentation: https://confluence.socialhistoryservices.org/display/DPEM/REST+API+documentation. Online access to IISH API: https://iisg.amsterdam/en/collections/using/machine-access.
- The catalogue offers and an automated citing service as an export of the record to Endnote and Refworks. For instance: https://hdl.handle.net/10622/3C26EA12-6914-4C3D-B3E6-D22EFF25B6E3.
- All collection items (high level archival descriptions and individual digital items to a very granular level) are resolved by a PID Handle service. This is documented on IISH PID declaration & binding workflow for digitized and born digital collections IISH (PID declaration & binding workflow for digitized and born digital collections), PID Handle Service https://pid.socialhistoryservices.org/, PIDS at the IISH (https://confluence.socialhistoryservices.org/display/CTS/PIDS+at+IISH) and Rebind PID procedure (https://confluence.socialhistoryservices.org/display/CTS/Rebind+PID+procedure).
a. Example of PID on high level archival description: http://hdl.handle.net/10622/ARCH01993.
b. Example of PID on inventory number level: http://hdl.handle.net/10622/ARCH01993.1.
c. Example of PID on image individual image level: http://hdl.handle.net/10622/08D61C12-F514-4A5D-A4D3-C2BB7CCAD8C4?locatt=view:level2.

Research data:

- All IISH research data can be discovered through the IISH Dataverse: https://datasets.socialhistory.org/ and from 2019 onwards also through the IISH general website (via Linked data).
- Every dataset is resolved through its own PID (example: http://hdl.handle.net/10622/3SG3KY) as is every file within the dataset (example: http://hdl.handle.net/10622/9FIDWF).
- Metadata assigned conform to the DDI metadata standards.
- All the datasets are available as linked data: https://druid.datalegend.net/IISG/iisg-kg (see under datasets).

XIV. 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

Response:

Metadata of archival donors:

See answer to R11: "The IISH aims to give proper access to all collections by adding sufficient descriptive metadata... created by the institute itself."

Data provided for the Designated Community:

- The digitised collections are presented - when still images - in the common JPEG format and - in case of objects which consist of more than one image like for instance books - PDF format. Video is presented as streaming MP4 and audio as MP3 files. When users need high resolution and/or uncompressed files they can order these files on the website.
- The born digital collections will be publicly available, but we will use authentication and probably on-site access only to give access to the users. The Institute strives to - when permitted and technically feasible - to give access to all objects in the original format (for instance MsWord .doc file) and, when necessary, in an at least one normalized access copy (for instance PDF).
- Both digitized and born collections will be made available through a IIIF (https://iiif.io/) compliant viewer. This will heighten the interoperability of the collections considerably.
- The research data are published on the IISH dataverse platform in common formats used within social history research. Metadata about the datasets can be exported as DC, DDI, Datacite, JSON, OAI_ORE, OpenAIRE and JSON-LD (see for instance https://hdl.handle.net/10622/FHJJYK under button 'export metadata'). The metadata are exported as linked data and will in 2020 be available through the integral data search on the IISH website.
- The IISH API can deliver descriptive metadata in EAD, DC and MARC (see for example: https://search.socialhistory.org/Record/COLL00293/Export).

Measures taken into account for the possible evolution of formats:

With the help of Archivematica the archive is monitored regularly for the possible obsolescence of the preservation formats. When a format threatens to become obsolete a migration workflow is started to create new preservation copies of the originals. In this case a new AIP is created. For reason of authenticity and provenance the original file will always be preserved (as are the earlier preservation formats). The new preservation format will be selected on the basis of the regularly updated IISH File format policy for born digital collections (https://confluence.socialhistoryservices.org/display/CTS/File+format+policy+for+born+digital+collections). The migration process (preservation action) will be documented in technical metadata (PREMIS).

Understandability of the data:

- By updating or migrating file formats that threaten to go obsolete (see above) the IISH ensures that file formats can be read by current software. Also, by offering comprehensive context in the form of descriptive metadata and metadata that proves the authenticity of the digital object the understandability is guaranteed. Preservation planning and action are described in the preservation policy.
- By using open and common (with social history research) standards for research data the understandability of the data is guaranteed on the short and medium term. On the long-term research data can follow the procedure described directly above whereby file formats will be migrated to then common formats.

Understandability of the metadata:

All collection data are described following international standards MARC21 and EAD (see R13). For MARC21 descriptions the IISH follows RDA (Resource Description and Access) description guidelines which, among other things, guarantee the authenticity and interoperability of the metadata. EAD guarantees the standardized description of archival collections. Each archival description comes with elaborate context information (for instance http://hdl.handle.net/10622/ARCH02639 under header 'content and context'). Additional IISH description guidelines further standardize the use of the metadata within the institute. Research data are described using the DDI standard via the IISH Dataverse platform (https://datasets.socialhistory.org/). This ensures that the data are described in a standardized and interoperable fashion.

TECHNOLOGY

XV. 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

Response:

The IISH uses the OAIS functional model
(https://en.wikipedia.org/wiki/Open_Archival_Information_System#The_functional_model) as guiding principle for the elements which have to be covered by the repository. For the recording and management of new acquisitions the IISH Acquisition database (Github: https://github.com/IISH/acquisition_database) is used. Digital packages are produced using the BagIt format using the tool Exactly (Github: https://github.com/IISH/uk-exactly). The archival bags are ingested into Archivematica (https://www.archivematica.org/en/) using the accession number provided by the Acquisition database.

Archivematica serves as the OAIS ingest workflow and preservation planning software tool. The ingest procedure in Archivematica is provided as a workflow through a number of microservices, each providing the necessary tools for the required checks, information extractions, transformations and normalization steps during the ingest procedure. At the end of the ingest procedure, Archivematica produces an AIP, which also conforms to the BagIt format. This AIP includes a METS document as a wrapper for all structural and technical metadata. All preservation metadata is included in the METS using the PREMIS data model. All customizations of rules and tooling to use for various file formats during the ingest procedure is configured using the preservation planning tools provided by Archivematica.

During the ingest procedure persistent identifiers (see R13) are assigned using the PID handle
service (Githbub https://github.com/IISH/PID-webservice). For persistent identifiers and resolving the IISH uses the Handle System. All digital objects and descriptive metadata are made accessible through persistent identifiers. See also R13.

All descriptive metadata are described using either MARC or EAD. MARC is used for describing the library holdings and the descriptions of individuals objects. Evergreen (http://evergreen-ils.org/) is used for the management of the MARC inventory. EAD is used for describing the archival descriptions. An XML editor XMetal (https://xmetal.com/) is used for the management of the EAD descriptions. All descriptive metadata can be accessed through OAI-PMH and SRU/SRW: https://iisg.amsterdam/en/collections/using/machine-access.

The IISH tries as much as possible to use open source (community supported) software. Especially software concerning collection management and dissemination is mostly all open source. All software, documentation and test material produced by the IISH is open source and available through our GitHub repository (https://github.com/iish). A list of software used and relevant technical documentation is managed and monitored by the Digital Infrastructure Department of the KNAW Humanities Cluster (HUC).

For real-time to near real-time data streams the provision of around-the-clock connectivity to public and private networks is at a bandwidth that is sufficient to meet the global and/or regional responsibilities of the repository. The IISH is connected to the 100Gb/s port of the Amsterdam Internet Exchange, which is considered to be one of the largest and fasted data hub transports in the world. Agreements of the KNAW with SURF Sara ensure this connection to a high quality and broad bandwidth connection.

Future developments are (among others):

- A further refinement of the pre-ingest procedure for digitized and born digital materials. This will be finished in the beginning of 2020.
- In 2020 research data will be ingested via Archivematica in this way be preserved in such a way that authenticity can be (as much as possible) guaranteed (see also R7).
- Also in 2020 the Acquisitions database and XMetal (both mentioned above) will be replaced by ArchivesSpace - an open source archives information management application. As ArchivesSpace can be integrated with Archivematica the workflows will be more efficient, especially with respect to born digital collections.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Accept

XVI. Security

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

Compliance Level:

3 – The repository is in the implementation phase

Reviewer Entry

Reviewer 1

Comments:
3 – The repository is in the implementation phase

Reviewer 2

Comments:
3 – The repository is in the implementation phase

Response:

The repository system is actively monitored and subject to standard security policies and procedures maintained at the Royal Dutch Academy of Arts and Sciences (KNAW).

Organizational systems are regularly subjected to vulnerability scans and to a yearly audit process (SURFaudit: https://www.surf.nl/en/surfaudit). In terms of SURFaudit recommendations all KNAW organizations are expected to operate at level 3 ('embedded in the organization'). Progress or deviations from this expected level of outcome are monitored on a yearly basis and results and improvement points are communicated in a yearly Security Workplan.

All security incidents are registered, coordinated and handled by the Computer Security Incident Response Team of the KNAW in accordance with the process for handling information security incidents ('Proces voor afhandeling informatiebeveiligingsincident CSIRT (nov 2015)') Technical administrators in collaboration with functional administrators resolve incidents. Each institute within the KNAW maintains an Information Security Officer acting as an intermediary between the central CSIRT group and institute. If an incident is reported the following standard procedures are followed:

1. Identification
2. Damage assessment and control
3. Repair activities
4. Evaluation and reporting

Each step is further subdivided into a specific set of actions related to the incident level.

As of January 1st 2016 each organization is obliged to report data leaks with the Autoriteit Persoonsgegevens (Authority personal data https://autoriteitpersoonsgegevens.nl/en) if a serious data leak has been discovered. See also the implementation of the GPDR and the measures taken by the KNAW and the Institute (as described in R4).

All of our servers are provided by the ICT Services of the KNAW and covered via Service Level Agreements. These also cover recuperation procedures in case of an organization wide outage. Risk management procedures are described at the organizational level.

To safeguard data deposited into the repository system multiple copies are maintained at off site locations. All data is furthermore replicated (see R9) in AIP packages that include metadata, data and authorization information. In case of a system outage all data can thus be retrieved from several locations.

From the perspective of risk management a yearly updated risk assessment document (Risicoscan IISG , "risk scan IISH", latest version April 2018) offers important evidence regarding the long term conservation of the collections. In this document the risk of data loss and cyber attacks is indicated and insight given on how these risks are controlled.

Although the R16 Security Level can be considered to be compliance level 4, the IISH would like to see this confirmed by the Surf Audit in 2020 (which at the moment of writing still has to take place) that will take into account the infrastructural

setup and measures taken over the last year when implementing Archivematica and new storage facilities. With the renewed confirmation of Surf Audit the level 4 is also confirmed.

APPLICANT FEEDBACK

Comments/feedback

These requirements are not seen as final, and we value your input to improve the core certification procedure. To this end, please leave any comments you wish to make on both the quality of the Catalogue and its relevance to your organization, as well as any other related thoughts.

Response: