Roadmap for Preservation and Curation in the SSH*

Background and description

In the Social Sciences and Humanities research is increasingly driven by the availability of a variety of digital resources, which exhibit an escalating internal complexity as well as diverse external relationships. The data production, management and dissemination processes are organized in a distributed manner, both within and between data producing institutions and data repositories. This fragmentation should be taken into account when designing or developing research infrastructures and data repositories for the corresponding scientific disciplines. The data produced within different data communities should be made available to their respective designated communities via environments that implement discipline specific workflows in a *trustworthy* manner.

A well-defined set of guidelines for digital curation is crucial to the continued viability and trust of digital materials. This report assesses existing data repository models, or frameworks that provide checklists or guidelines to data preservation entities. Based on the assessment of the various models it can be considered as a "guideline of guidelines" that can be used for a broad spectrum of digital long-term repositories that aim to retain their validity over a longer period. By identifying and discussing selected benchmark guidelines and standards for trusted digital repositories, the report can work as an implementable resource for existing and emerging repositories that seek to provide trustworthy long-term data repository services.

Findings

The guideline and standards are organized in a five-step 'trust maturity' model. This model is based on the five organizational stages of digital preservation (Kenney & McGovern, 2003), the CMMI five organizational maturity levels (SEI/Carnegie Mellon, 2010), and the Trusted Digital Repository framework that consist of three levels of trustworthiness.

The assumption behind the model is that all data repositories that seek to become trusted digital repositories should assess their organization and aim for a certain level of trust maturity.

Level 3, peer-reviewed self-assessment through the Data Seal of Approval (DSA), is considered the benchmark level as it contains trust requirements that are both necessary and sufficient for most data repositories within the SSH area. CLARIN and CESSDA require DSA level and have some additional community specific requirements.

External review and formal certification at level 5 is fairly expensive, time consuming and requires a high level of expertise. ISO 16363 certification should therefore only be considered by repositories that have special data holdings and special needs.

Table 1: Summary of maturity levels and key guidelines

Trust Maturity Level	Key Guideline	Guideline Source
1. OAIS Core Conformance	Support OAIS Information Model	OAIS Information Model: Section 2.2 of CCSDS 650.0-M-2 / ISO 14721:2012
	Acknowledge OAIS Archive responsibilities	OAIS Archive Responsibilities: Section 3.1 of CCSDS 650.0-M-2 / ISO 14721:2012
2. Initial self-assessment, PLATTER/DRAMBORA	Self-assessment through PLATTER and DRAMBORA	PLATTER Key Self-assessment questions
		DRAMBORA Key Self-assessment questions
3. Peer-reviewed self-assessment I, DSA	Peer-reviewed self-assessment I, DSA	Data Seal of Approval Guidelines
		Support: NESTOR criteria
4. Peer-reviewed self-assessment II, ISO 16363/DIN 31644	Conformance to the OAIS Detailed Functional Model	OAIS Detailed Functional Model: Section 4.1 of CCSDS 650.0-M-2 / ISO 14721:2012
	Self-audit with the ISO 16363	CCSDS 652.0-M-1 / ISO 16363:2012
	Alternatively, self-audit with DIN 31644	DIN 31644
5. Certification and Optimization	External review and formal certification in conformance with the ISO 16363	CCSDS 652.0-M-1 / ISO 16363:2012
	Alternatively, with DIN 31644	DIN 31644

DASISH