



## Core Trustworthy Data Repositories Requirements: Glossary

---

*This Glossary of Terms supports the Core Trustworthy Data Repository Requirements:*  
<https://doi.org/10.17026/dans-22n-gk35>.

(\* Taken from OAIS<sup>1</sup>)

**Access Rights Information\*:** The information that identifies the access restrictions pertaining to the Content Information, including the legal framework, licensing terms, and access control. It contains the access and distribution conditions stated within the Submission Agreement, related to both preservation (by the repository) and final usage (by the Consumer). It also includes the specifications for the application of rights enforcement measures.

**Archive\*:** An organization that intends to preserve information for access and use by a Designated Community.

**Authenticity\*:** The degree to which a person (or system) regards an object as what it is purported to be. Authenticity is judged on the basis of evidence.

**Consumer\*:** The role played by those persons, or client systems, who interact with repository services to find preserved information of interest and to access that information in detail. This can include other repositories, as well as internal repository persons or systems.

**Data\*:** A reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing. Examples of data include a sequence of bits, a table of numbers, the characters on a page, the recording of sounds made by a person speaking, or a moon rock specimen.

**Designated Community\*:** An identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities. A Designated Community is defined by the Archive and this definition may change over time.

**Digital Migration\*:** The transfer of digital information, while intending to preserve it, within the repository. It is distinguished from transfers in general by three attributes: a focus on the preservation of the full information content that needs preservation; a perspective that the new archival implementation of the information is a replacement for the old; and an understanding that full control and responsibility over all aspects of the transfer resides with the repository.

**Digital Object\*:** An object composed of a set of bit sequences.

**Domain or subject-based repository:** Specializes in a specific research field or data type.

**Ingest:** The process of entering data and associated metadata into a data repository.

**Integrity:** Internal consistency or lack of corruption of digital objects. Integrity can be compromised by hardware errors even when digital objects are not touched, or by software or human errors when they are transferred or processed.

**Institutional repository:** Generic, multi-subject repository serving a research performing institution.

**Long Term\*:** A period of time long enough for there to be concern about the impacts of changing technologies, including support for new media and data formats, and of a changing Designated Community, on the information being held in a repository. This period extends into the indefinite future.

**Long-term Preservation\*:** The act of maintaining information, Independently Understandable by a Designated Community, and with evidence supporting its Authenticity, over the Long Term.

**National repository system, including governmental:** Multidisciplinary, national infrastructure. Has a legal mandate for certain (public or governmental) data types.

**Open Archival Information System (OAIS)\*:** An Archive, consisting of an organization, which may be part

<sup>1</sup> Consultative Committee for Space Data Systems. Reference Model for an Open Archival Information System (OAIS). Recommended Practice -- CCSDS 650.0-M-2. Magenta Book, June 2012. <http://public.ccsds.org/publications/archive/650x0m2.pdf>



of a larger organization, of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community. It meets a set of responsibilities that allows an OAIS Archive to be distinguished from other uses of the term 'Archive'. The term 'Open' in OAIS is used to imply that this Recommendation and future related Recommendations and standards are developed in open forums, and it does not imply that access to the Archive is unrestricted.

**Preferred Formats:** Formats that a repository can reasonably assure will remain readable and usable. Typically, these are the de facto standards employed by a particular discipline.

**Producer\*:** The role played by those persons or client systems that provide the information to be preserved. This can include other repositories or internal repository persons or systems.

**Provenance Information\*:** The information that documents the history of the Content Information. This information tells the origin or source of the Content Information, any changes that may have taken place since it was originated, and who has had custody of it since it was originated. The Archive is responsible for creating and preserving Provenance Information from the point of Ingest; however, earlier Provenance Information should be provided by the Producer. Provenance Information adds to the evidence to support Authenticity.

**Publication repository:** Generic, multidisciplinary repository, focussing on data linked to publications.

**Reference Model\*:** A framework for understanding significant relationships among the entities of some environment, and for the development of consistent standards or specifications supporting that environment. A reference model is based on a small number of unifying concepts and may be used as a basis for education and explaining standards to a non-specialist.

**Research project repository:** Capture research results that require a deposit mandate by a funder or organization as a 'record of science'. Often tied to a specific (multi)disciplinary project.

**Reuse:** The use of data collected for one purpose to study a new problem or to verify the conclusions of the data producer.

**Succession Plan\*:** The plan of how and when the management, ownership and/or control of the repository holdings will be transferred to a subsequent repository in order to ensure the continued effective preservation of those holdings.