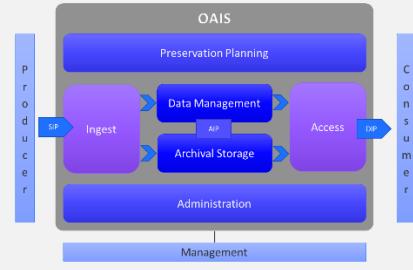


Data terms glossary

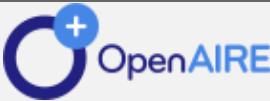
Terms		Definition
English	Français	
Archiving	Archivage	Data archiving is the action to preserve data no longer actively used for the long term and ideally conserved in an OAIS-compliant system to guarantee preservation. The archive must be self-descriptive (metadata), searchable and retrievable. An archival system is based on the concept of storage and can have backups .
Backup	Sauvegarde	Data backup is the action to preserve data for the short term . Generally, the backup process is automatic to avoid human errors and can be incremental. Backups are not searchable, except based on file names. A backup system is based on the concept of storage .
Storage	Stockage	Data storage is the action to save data on media or on specific technologies. A medium can be hard disk drives (HDD), optical and molecular materials, solid-state drives (SSD) or tapes (LTO). Technologies can be network-attached storage (NAS), storage area network (SAN) or object storage (S3).
Long Term Preservation	Conservation à long terme	Long-term preservation is concerned with the actions to be taken to archive data.
Digital Preservation	Conservation numérique	Digital preservation refers to the action of preserving data in an archiving system, ideally compliant with OAIS.
OAIS	OAIS	<p>OAIS refers to the ISO standard 14721, which specifies the requirements for designing a digital preservation system. The OAIS model is the most well-known standard, usually represented as follows:</p>  <p>The OAIS system guarantees the preservation of the archive for the long term based on replication, fixity check, migration, etc. One example of OAIS-compliant system is Yareta.</p>
Data transfer	Transfert de données	Data transfer is the process of moving data from one location to another, either within a single device or between devices and networks.
Data quality control	Contrôle qualité des données	Data quality control includes processes and measures aimed at maintaining data accuracy, completeness, consistency, and reliability. It involves a systematic approach to identifying, rectifying, and preventing errors or discrepancies in data sets, ensuring they remain fit for purpose.
Dark archive	Archive froide	A dark archive is an archiving system in which the archive is not retrievable. In the OAIS model, a dark archive has no access module.
Dataset	Jeux de données	A dataset is a set of data, which includes metadata and possibly para-data. A research dataset can represent the result of research works or scientific experiences. The dataset granularity should authorize the access of individual data.
Data	Données	In the research domain, data refers to research data. The data artefacts could be a file, multiples files, a structure of folders and files, etc. Data can be associated with the concept of dataset .

Metadata	Métadonnées	Metadata is data of data: it represents the information that describes the data. Generally, the metadata must respect a schema, typically represented in XML or JSON. The DataCite schema is one of the most recognized formats for citations in the research domain.
Para-data	Para-données	Para-data is metadata of processes. They can for instance describe the context of a research experience.
Active data	Données actives	Active data is data in the stages of the lifecycle in which data are still alive (<i>updatable</i>), that is, before to be archived (and put in the <i>immutable status</i>).
Cold data	Données froides	Cold data refers to data that are rarely accessed, if ever.
Hot data	Données chaudes	In contrast to cold data, hot data is frequently accessed.
Raw data	Données brutes	Raw data are data that have been created or collected from research sources or experience instruments before being processed. This is a fuzzy concept as processing is often applied at the level of sources and instruments.
Processed data	Données traités	The research process transforms <i>raw data</i> into <i>processed data</i> . Processing can be data cleaning, data quality check, etc., the goal being to yield qualitative or quantitative data for further uses in the research work.
Analyzed data	Données analysées	The research process can include further analyses to give meaning to the <i>processed data</i> . Specific algorithms, such as those based on statistic tools, are often involved at this stage.

Data acronyms glossary

Acronym	Definition		
FAIR	FAIR principles for Findable Accessible Interoperable Reusable https://www.go-fair.org/fair-principles/		
OAIS	Open Archival Information System https://en.wikipedia.org/wiki/Open_Archival_Information_System		
SAE	Système d'Archivage Electronique <i>OAIS in French</i>		
DataCite	Metadata for data citation https://schema.datacite.org/		
DOI	Digital Object Identifier https://www.doi.org/		
ORCID	Open Researcher and Contributor Identifier https://orcid.org/		
ROR ID			
RDF	Resource Description Framework https://www.w3.org/RDF/		
IIIF	International Image Interoperability Framework https://iiif.io/		
DP	Digital Preservation		
DMP	Data Management Plan		
ORD	Open Research Data		
DTPA	Data Transfer and Processing Agreement		
DTUA	Data Transfer and Use Agreement		
ELN	Electronic Laboratory Notebook		
LIMS	Laboratory Information Management System		
HDD	Hard Disk Drive		
SSD	Solid State Drive		
LTO	Linear Tape-Open		
NAS	Network-Attached storage		
ROPA	Register Of Processing Activities https://www.europarl.europa.eu/data-protection/fr/data-protection-register		
SAN	Storage Area Network		
S3	Amazon Simple Storage Service		
NFS	Network File System https://en.wikipedia.org/wiki/Network_File_System		
SMB	Server Message Block https://en.wikipedia.org/wiki/Server_Message_Block		
CTS	Core https://www.coretrustseal.org/	Trust	Seal

Data acronyms glossary

Institution	Website
 UNIVERSITÉ DE GENÈVE	https://www.unige.ch/researchdata/en/presentation/glossaire/ https://www.unige.ch/researchdata/fr/propos/glossaire/
 ETH zürich	https://unlimited.ethz.ch/spaces/DD/pages/194127817/Glossary+-+Research+Data+Management
 OpenAIRE	https://www.openaire.eu/rdm-glossary
 RESEARCH DATA OXFORD	https://researchdata.ox.ac.uk/rdm-glossary
 IMPERIAL	https://www.imperial.ac.uk/research-and-innovation/support-for-staff/scholarly-communication/research-data-management/training-and-resources/glossary/
 DigitalPreservationCoalition	https://www.dpconline.org/handbook/glossary
 DCC	https://www.dcc.ac.uk/about/digital-curation/glossary
 SPHN Swiss Personalized Health Network	https://sphn.ch/