IDOC Executive summary

IAS-IDOC-EX-001

	Name and Function	Date
Authors	Gilles Poulleau, Karin Dassas	
Checked	Marian Douspis	
Approved	***OSUPS head, IAS head***	

Diffusion List

Name	Function	Company

Change log

Issue	Date	Modifications
1.0	08/02/2016	1st draft in French
2.0	13/12/2016	translated in English
2.1	05/05/2017	Entirely rewritten
2.2	08/12/2017	Updated following the creation of the class of service “data management”
2.3	26/01/2108	Updated following the creation of the Instructions for data management document
2.4	16/01/2019	Some precisions
2.5	02/04/2020	Paris-Saclay and some other details
2.6	12/06/2020	Standards
2.7	23/06/2020	Particulars and denominations
2.8	03/06/2021	Updates
3.0	01/03/2022	IDOC, IDOC–INSTR, IDOC-OPE, IDOC-DATA clarification

SUMMARY

1 SCOPE OF THE DOCUMENT

The present document aims at describing IDOC (Integrated Data & Operation Center), how it is organized and how it operates. It is the “cover” of other documents:

- **IDOC-OD-002** presents the risk analysis and management.
- **IDOC-INS-003** details the general principles applicable to project design
- **IDOC-INS-008** is the baseline reference for implementing new services. This document covers five other documents as described in the chart below.
- **IDOC-INF-010** presents the IDOC organization chart.

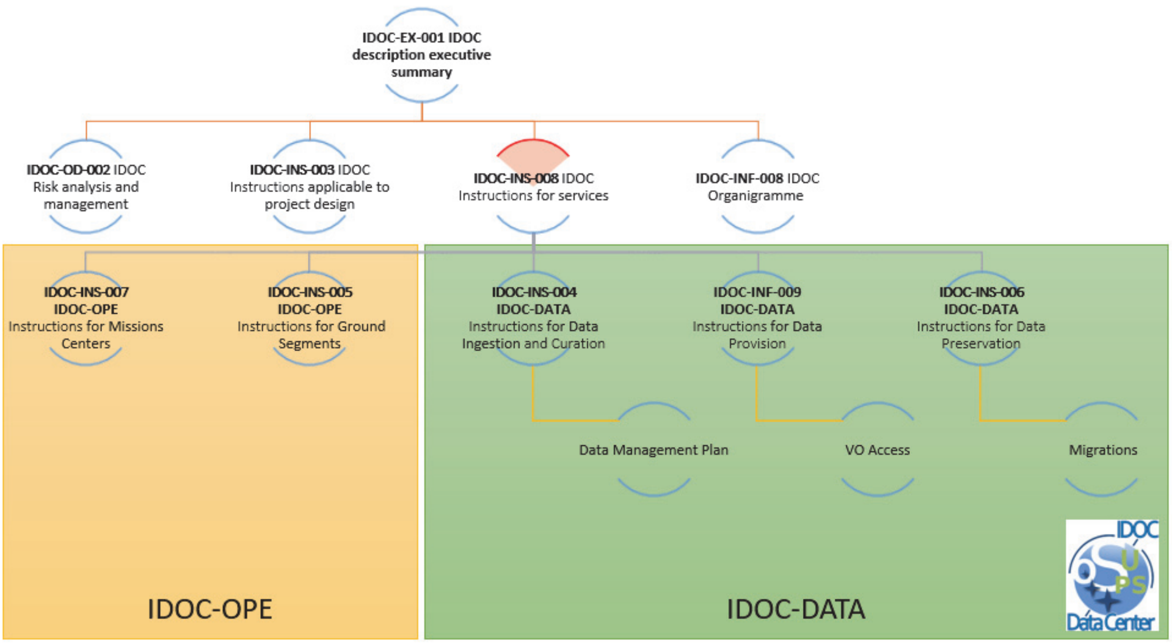


Figure 1 IDOC references repository

All together, these documents constitute the IDOC references repository.

2 REFERENCE DOCUMENTS

Acronym	Reference of the document	Document full name
RD1	IDOC-EX-001	IDOC-EX-001 IDOC executive summary
RD2	IDOC-OD-002	IDOC-OD-002 IDOC Risk analysis and management
RD3	IDOC-INS-003	IDOC-INS-003 IDOC Instructions applicable to project design

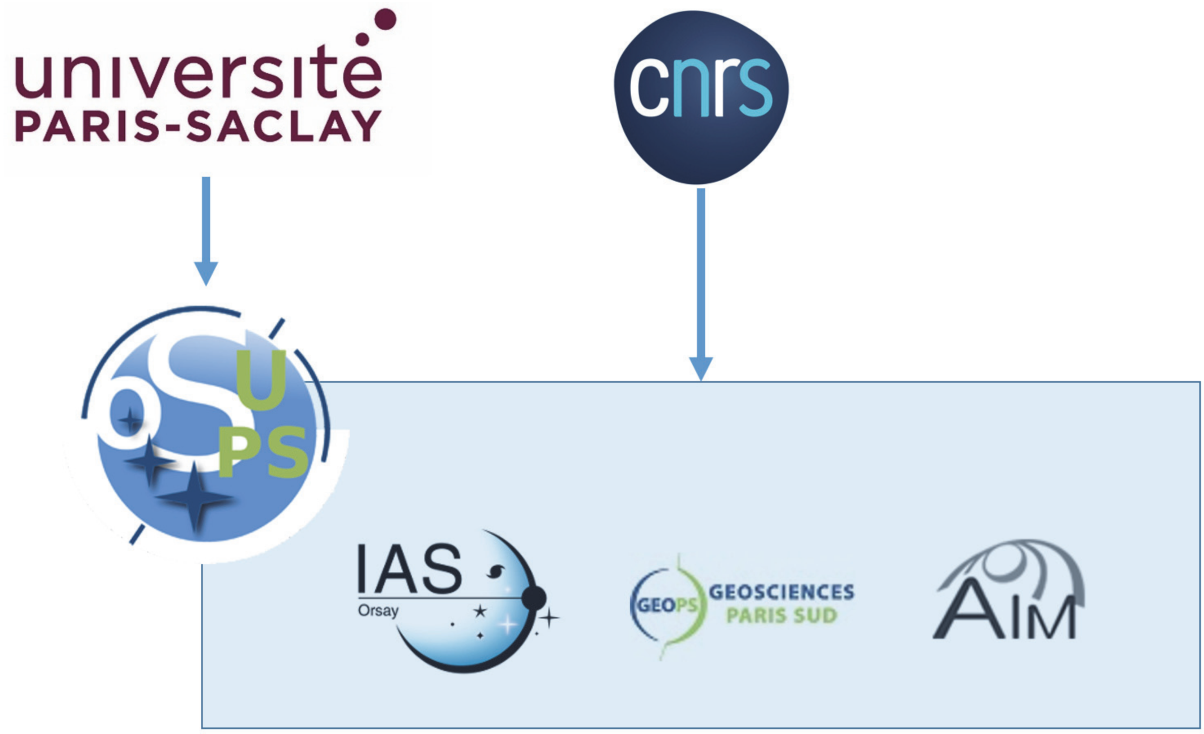
RD4	**IDOC-INS-004**	**IDOC-INS-004 IDOC-DATA Instructions for Data Ingestion and Curation**
RD5	**IDOC-INS-005**	**IDOC-INS-005 IDOC-OPE Instructions for Ground Segments**
RD6	**IDOC-INS-006**	**IDOC-INS-006 IDOC-DATA Instructions for Data Preservation**
RD7	**IDOC-INS-007**	**IDOC-INS-007 IDOC-OPE Instructions for Instrument Operations**
RD8	**IDOC-INS-008**	**IDOC-INS-008 IDOC Instructions for Services**
RD9	**IDOC-INS-009**	**IDOC-INS-009 IDOC-DATA Instructions for Data Provision**
RD10	**IDOC-INF-010**	**IDOC-INF-010 IDOC Organigrammes**
RD11	**IDOC-DW-011**	**IDOC-DW-011 Diverses schemas for documentation**
RD12	**IDOC-INS-012**	**IDOC-INS-012 IDOC instructions for architecture and coding practices**
RD16	**IDOC-EX-016**	**IDOC-EX-016 OSUPS Schéma Stratégique Numérique**
RD17	**IDOC-OD-017**	**IDOC-OD-017 Services offerts par IDOC**
RD30	**IDOC-HO-030**	**IDOC-HO-030 Presentation IDOC-public-english**
RD31	**IDOC-HO-031**	**IDOC-HO-031 Presentation IDOC Français**

3 IDOC DESCRIPTION

CONTEXT

IDOC (Integrated Data & Operation Center) has existed since 2003 as a satellite operations center and data center for the Institute of Space Astrophysics (IAS) in Orsay, France. Since then, it has operated within the OSUPS (Observatoire des Sciences de l'Univers de l'Université Paris-Saclay - first French university in shanghai ranking), which includes three institutes: IAS, AIM (Astrophysique, Interprétation, Modélisation - IRFU, CEA) and GEOPS (Geosciences Paris-Saclay) .
IDOC participates in the space missions of OSUPS and its partners, from mission design to long-term scientific data archiving. For each phase of the missions, IDOC offers services in the scientific themes of OSUPS to ensuring data integrity, facilitating data sharing, and providing data stewardship.
In order to meet its specific missions, IDOC's activities are divided into three departments:

- IDOC-INSTR: instrument design and testing,
- IDOC-OPE: instrument operations,
- IDOC-DATA: data management and data value chain.



The OSUPS context

MISSIONS AND RESPONSIBILITIES

IDOC depends on an institute of the National Scientific Research Centre (CNRS) that organises the whole French research, the first research organization in the world with more than 30 000 employees.

IDOC depends on the Université Paris-Saclay, 16th in the Shanghai ranking

IDOC-DATA is labelled as a CNRS " regional center of expertise" by CNRS since 2013 (https://www.insu.cnrs.fr/fr/les-centres-et-poles-de-donnees)

IDOC-DATA is labelled as a CNRS platform since 2018. (https://www.insu.cnrs.fr/plateformes-insu)

IDOC-DATA is recognized as a long-term archive by CNES, the french space agency since 2016.(https://medoc.cnes.fr/en/medoc-0)

IDOC-DATA is part of an OSU, a perennial structure in charge precisely of observation services (see the official assignments here:

https://www.legifrance.gouv.fr/loda/article_lc/LEGIARTI000027866130/2020...).

IDOC's two supervisory authorities depend on the Ministry of Research, which has made commitments in terms of open science:

https://www.ouvrirlascience.fr/wp-content/uploads/2019/08/National-Plan-for-Open-Science_A4_20180704.pdf

The CNRS also in particular with its Roadmap for Open Science and its Plan for Research Data 2020: https://www.science-ouverte.cnrs.fr
The INSU certified data centers provide data, tools or services in order to bring together or crossreference information from different production sources. The INSU CERs ensure the development and operational maintenance of National Observation Services (SNO) in astronomy and astrophysics.

https://www.insu.cnrs.fr/fr/les-centres-et-poles-de-donnees
https://www.insu.cnrs.fr/fr/les-services-nationaux-dobservation

The commitments of the University of Paris-Saclay to open science are described here: https://www.universite-paris-saclay.fr/recherche/science-ouverte/les-donnees-de-la-recherche
The overall objective of IDOC is to support the generation of instruments and then to enable them to produce scientific data. IDOC must then provide an environment that facilitates the appraisal of these data, their curation, wide reuse, mid & long-term preservation. This environment must be completely appropriate for the uses of the different communities

All these missions must be covered while respecting the principles that allow the data to comply with the FAIR principles. IDOC must therefore behave as a TRUST data repository by being attentive to the CARE aspects linked to the user communities (see glossary), even if the origin of the data makes this last point less sensitive than in other research fields.

The themes currently covered by IDOC are as follows:

- interstellar medium physics,
- cosmology,
- stellar physics,
- planetary surfaces,
- thematics of GEOPS
- thematics of AIM.

GOVERNANCE AND MANAGEMENT

IDOC is structured from a governance and a management structure, a technical center, three departments and seven scientific themes. It provides also training and communication services

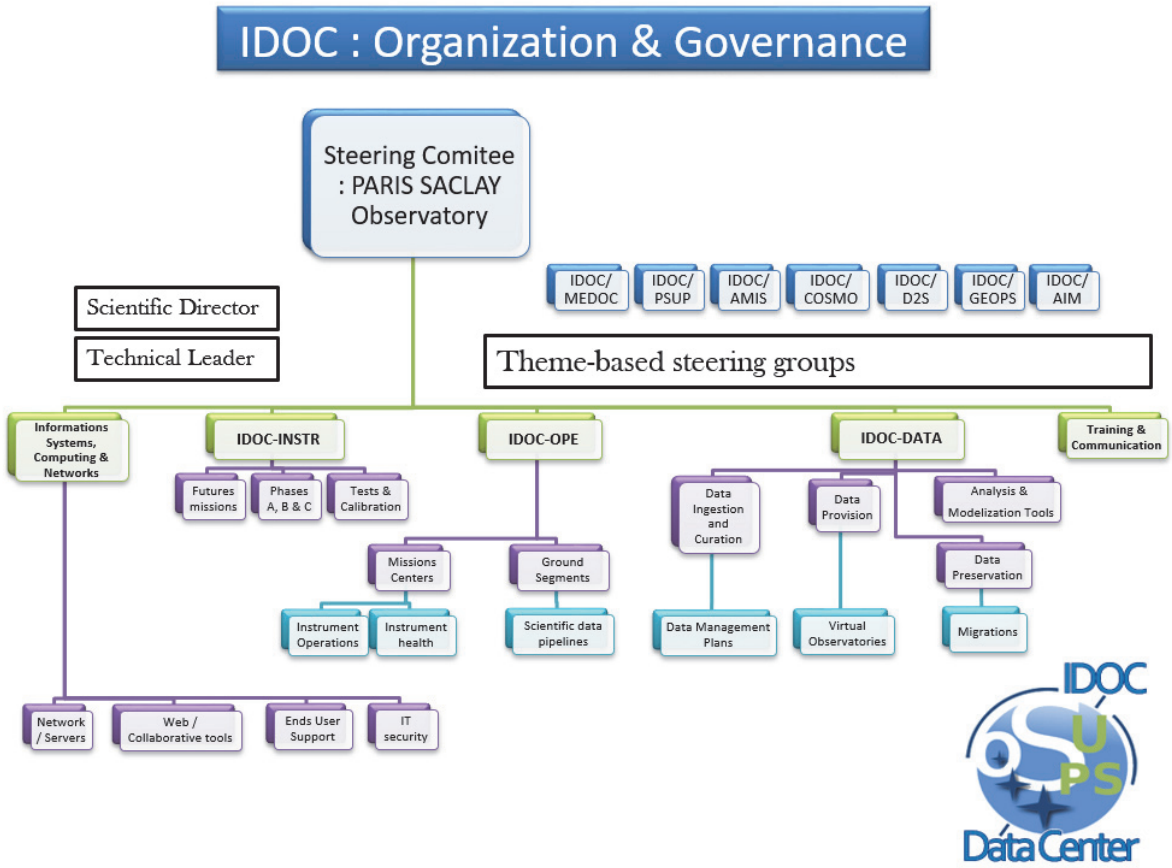


Figure 2 IDOC organization

- IDOC governance is assured by the IDOC steering committee. It is designated by the OSUPS Governing Board which gives its recommendations. The IDOC steering committee nominates both the IDOC technical and scientific leaders who are invited to attend its meetings.

- The scientific themes are:
    - MEDOC: Multi Experiment Data & Operation Centre for Solar Physics,
    - AMIS: Astrophysics of Interstellar Matter
    - COSMO: Cosmology
    - DS2 (Données Systèmes Stellaires) Stellar physics
    - PSUP: Planetary Surfaces Portal
    - Other themes of GEOPS
    - Other themes of AIM

Each of themes is organized on an ad hoc manner: a theme can have its own Directorial Committee (e.g., MEDOC) or can have a simpler structure. Nevertheless, each theme has one scientific coordinator with respect to IDOC. The scientific coordinator acts as representative of the scientific domain he represents. He has an overall view of the domain, projects, problems, and advises both IDOC and its requesters.

- The technical center provides a set of technical and pooling services allowing users to work with data and products, produced by spatial missions, and other products from OSUPS, available at IDOC. It includes both an operation center and a data center.

The people currently in charge are:

- IDOC steering committee
- IDOC technical leader: Gilles Poulleau
- IDOC scientific leader: Marian Douspis
- MEDOC scientific coordinator: Eric Buchlin
- AMIS scientific coordinator: Emilie Habart
- COSMO scientific coordinator: Julien Grain
- DS2 scientific coordinator: Frédéric Baudin
- PSUP scientific coordinator: Mathieu Vincendon
- GEOPS scientific coordinator: Christophe Colin
- AIM scientific coordinator: Pierre Olivier Lagage

The IDOC organization chart is detailed in [RD9].

Organizational strategy

To support its activities and its commitments, IDOC has limited human resources and proper financial fund. In order to optimize human resources and funds, IDOC has adopted a strategy based on two major principles: "pooling of resources" and "compliance with standards".
This strategy encompasses without distinction the three departments of IDOC: IDOC-INSTR, IDOC-OPE, IDOC-DATA

Pooling of resources

This part of the strategy includes hardware and software purchase, and software development. This pooling strategy is expanded by developing partnerships (i.e., « virtualdata platform of P2IO labex»).
IDOC activities being significantly pooled, it allows some skill redundancy and a more efficient resource management. It also allows a repartition of responsibilities in the information transit from hardware elements and the network to « high level » applications. Projects participate to this pooling strategy and profit from the best return on investment.
This approach is beneficial for both IDOC and its partners. It allows a better use of hardware resources and skills. This approach enables IDOC to meet its commitments. The global return on each investment is greatly increased.
The integration of a new project at IDOC requires from the project leader understanding of this pooling strategy and his commitment to it.

Human resources

Pooling strategy has many benefits:

- Pooling of IDOC staff skills allows each project to benefit from redundancy of skills required in the project development.
- Staff expertise increases due to pluralism of implementations using collective tools and software items.
- Tool mastery is strengthened by using common procedures and rules.

- This strategy brings a reactivity which would have been difficult to produce otherwise.

The pooling strategy of mastery/expertise/availability enables permanent or non-permanent staff to participate to different projects. The implication of the staff participation can be small but very efficient. Although gross percentages in term of FTE (full-time equivalent) might seem low, staff contribution to the progress of projects might be multiplied by factors 5 to 10.
Document [RD10] gives the current staff organization and composition.
Document [RD12] gives some figures about human resources and provides details of the expertise and their distribution.

Procurement and use of hardware resources.
IDOC sets up the most adapted hardware technologies to process, store, and transfer data, under project requirement constraints. The different hardware infrastructures are based on standard and modular elements, allowing a wide adaptability to the requirements and a reusability.
The pooling of hardware items allows to purchase in large quantities and with a higher quality within a constant budget (it might even occur that those items would not have been affordable to projects otherwise).
IT technologies evolving rapidly, a dedicated purchase strategy is used:

- Purchase of hardware platforms with five-year warranty after several tests and validations considering a lower cost for a project requirement.
- Networking of this new platform (usually more efficient than the previous version) for significant activities and big consumers.
- Purchase of a new platform after a 7-year usage maximum. The limit of seven years is usually enough to ensure a gain performance, a running and supply cost, which are significant to ensure the profitability of the new equipment. There is no maintenance performed on the hardware, except in the case of an urgent need for the availability of a new platform.
- The former platform is then re-used for less demanding usage (redundancy, long term archive.) Its lifetime is then the hardware lifetime, i.e., as long as it works, and as long as its maintenance at low cost is possible. Up to additional two to three years, the use of the hardware should be reserved to temporary troubleshooting, to serve as a buffer during IT migrations, and to non-critical long-term archive. After six or seven years, the high energy cost of hardware should lead to its abandonment.
- The acquisition of the computer equipment is done each year with some exceptional purchases if necessary. At IDOC, four groups of platforms are coherent with the procedure described above. This global quantity allows the necessary flexibility to ensure all the points requested with performance and security levels adapted.
- When possible, purchases are anticipated or delayed alleviating cost variation with time and/or with quantities.

Compliance with standards

This approach aims to have IDOC adopt methodologies, tools, software, procedures, etc. with the following characteristics: widely recognized and used, adapted to IDOC's themes, with excellent visibility and sustainability, and having limited implementation costs.
IDOC closely follows the evolution of these social and technical bridges that enable for instance open sharing and re-use of data.

Software development

Although IDOC might temporarily have enough resources to develop dedicated software, experience has shown that specific developments are difficult to maintain in the long term: they become quickly out of date and their potential updates are no longer supported.
The strategy is thus to systematically use or reuse existing solutions available from partners or the open-source community. IDOC thus aims to use commonly known tools and software for all developments.[1]. This set of tools is shared by the whole team and is enriched along each new activity.

Participation to standards, technological monitoring, and inclusion in citizens' objectives

In addition to its obvious interest in established standards, IDOC also participates in the development of new standards or new ways of proceeding, particularly in the field of data management or computer resources:

- HPC, GPU computing, machine learning, deep learning
- Clouds (openstack), big data
- Standards of format for data or metadata (FITS, PDS, HDF, NETCDF...)
- Data access (Virtual Observatories with protocol TAP/SIAP/CS, OGC with WMS, WCS, WFS …)
- Methodology / Infrastructure (ITIL / OAIS / FAIR data)
- Organization of open data: IDOC is involved in the Resource Data Alliance (RDA) organization
- Eco-responsibility: considering energy and environmental aspects in the design and maintenance of infrastructures and various actions. Involvement of personnel in these approaches (French national GDS Ecoinfo)

IDOC technological monitoring aims to explore future technical developments that seems promising for its present and future requirements. It allows to be prepared to integrate the right tools at the right time when their benefits are validated. The participation of IDOC members in regional and national networks allows to reinforce the effectiveness of this technology monitoring.

4 IDOC DEPARTMENTS AND SERVICE INTEGRATION

DESCRIPTION OF THE IDOC DEPARTMENTS

To meet its specific missions, IDOC's activities are divided into three departments:

- **IDOC-INSTR: instrument design and testing,**
- **IDOC-OPE: instrument operations,**
- **IDOC-DATA: data management.**

Two more entities of services are transversal to the three IDOC departments:

- **Infrastructure Computer Systems, Networks**
- **Communication & end-user support**

IDOC-INSTR

Instrument design and testing: accompany the genesis of instruments from their conception, feasibility studies, during their manufacturing and calibration, until their delivery. IDOC-INSTR also involved to design all supporting computing environments (conception, tests, calibration, ground segments...).

[1] As an example, IDOC has chosen the SITools framework developed by CNES. It is used for all interfaces via small necessary adjustments.

IDOC-OPE

Instrument operations: operate the instruments and ensure their technical follow-up and the adequate execution of their scientific mission. This task is integrated in the IDOC infrastructure considering the constraints of the instrument.

IDOC-DATA

Data management and data value chain: Three types of services are available to support the life of the data within IDOC

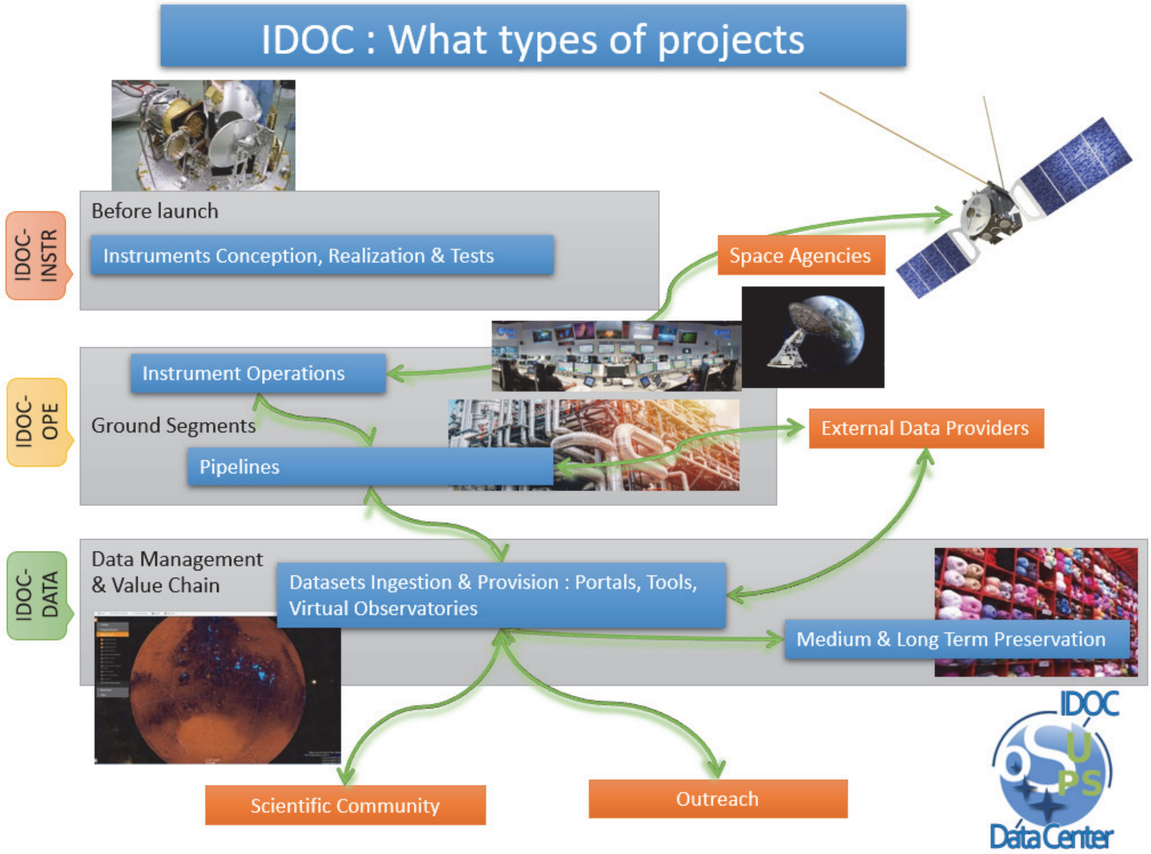
a) Data integration and curation: when a dataset (which may have been produced by IDOC-OPE or outside) needs to be integrated into the IDOC-DATA infrastructure. After the first ingestion process, as this dataset may not be complete, organized and documented, curation processes are applied to fill or mitigate these gaps, contextualize, and complete metadata and validate the whole.

b) Data provision: appropriate environments are built to allow users to search, discover and reuse any data hosted at IDOC-DATA. Various interfaces make it possible to query datasets, but it can also allow users to trigger value-added processing on part of the data using various tools. It also allows to register and browse the datasets in the appropriate virtual observatories for the designated community. These communities are solicited either formally through user or steering committees or more informally through feedback on the web interfaces.

c) Data preservation: refers to the medium and long-term preservation of a dataset hosted by IDOC-DATA and its ability to continue to be used by all audiences.

Infrastructure Computer Systems, Networks:

Assisting users in their use of the above services inside the three departments.



Procedure for a IDOC service request

Define the new service

When IDOC is requested to integrate a new service and implement it, the request is analyzed by the IDOC technical manager with the help of the relevant team members to verify if it can be effectively deployed and what measures need to be taken to allow its establishment in compliance with IDOC instructions.

Build the appropriate documents

Given the requested service, the needed inputs are different. In order to help each request, specific instructions describe the services and an appendix set the specific questions to be answered, checked, and validated for performing the expected implementation. This allows writing the requirement document (: “the plan”) which will be the basis of the mutual understanding between the IDOC Technical leader and the decision-maker. For further details, see [RD4], [RD5], [RD6], [RD7], [RD9] and [RD10]. In each of the documents, a "how-to" appendix called "procedure to prepare" allows to define the context for the creation of a new service in accordance with the procedures in force at IDOC. These procedures are designed to enable hosted services to achieve the levels of compliance expected by international standards (ITIL, OAIS, Fairness of the data, Open Science etc.) and space agencies. The heads-up document “Instructions for new services” describes the general principles that apply.

Figure 3 IDOC services (in blue)

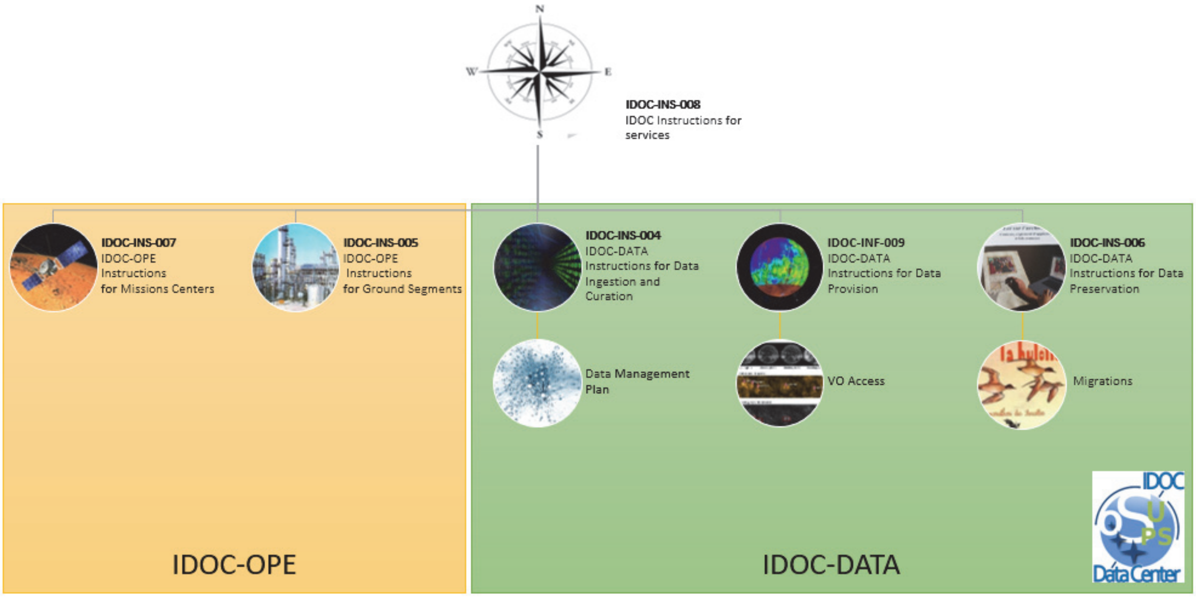


Figure 4 IDOC Instructions for new services

IDOC brings its technical skills to users and assists them for this response procedure. This process is iterative and aims to build a satisfying convergence between resources and requirements. The formalization of this convergence is set by the writing of the above documents.

5 IMPLEMENTATION OF IDOC COMMITMENTS

IDOC COMMITMENTS

Along with the services provided by IDOC, also go several commitments.

Continuity and preservation of service

IDOC ensures the continuity of its services. For this purpose, IDOC is organized for confronting incidents or unexpected events which can disturb its functioning. Its organization enables to manage the infrastructure with an effective 24/7 monitoring supported by a team best effort basis.
The organization considers the availability and security required by the services, establishes their monitoring, and ensures their traceability.

Design and implementation

[RD3] describes the general principles applicable to project design. The key components of this strategy are highlighted below:

- Make it simple (Kiss: Keep It Simple Stupid).
- Clearly define tasks of each element and force the simplest possible interfaces.
- Ensure the redundancy of key resources.
- Ensure independence of software developments with respect to the hardware.
- Use of the most popular standards.
- Develop software by keeping in mind the automatization of the processes.

Quality approach

The quality approach is based on the ITIL approach combined with the requirements of quality plans for spatial projects, and with the calls for proposals (Horizon H2020). A certification process is undertaken with independent agencies (Coretrustseal). The approach also aims to follow the OAIS norm.
The methodology of all quality activities (project management, development, monitoring, and communication) is supported by a suit of collaborative tools which allow all participants (project managers, developers, system and network administrators, technicians) to build and access the necessary documentation.
The integration of IDOC-DATA services and the means implemented in the OAIS context are as follows:

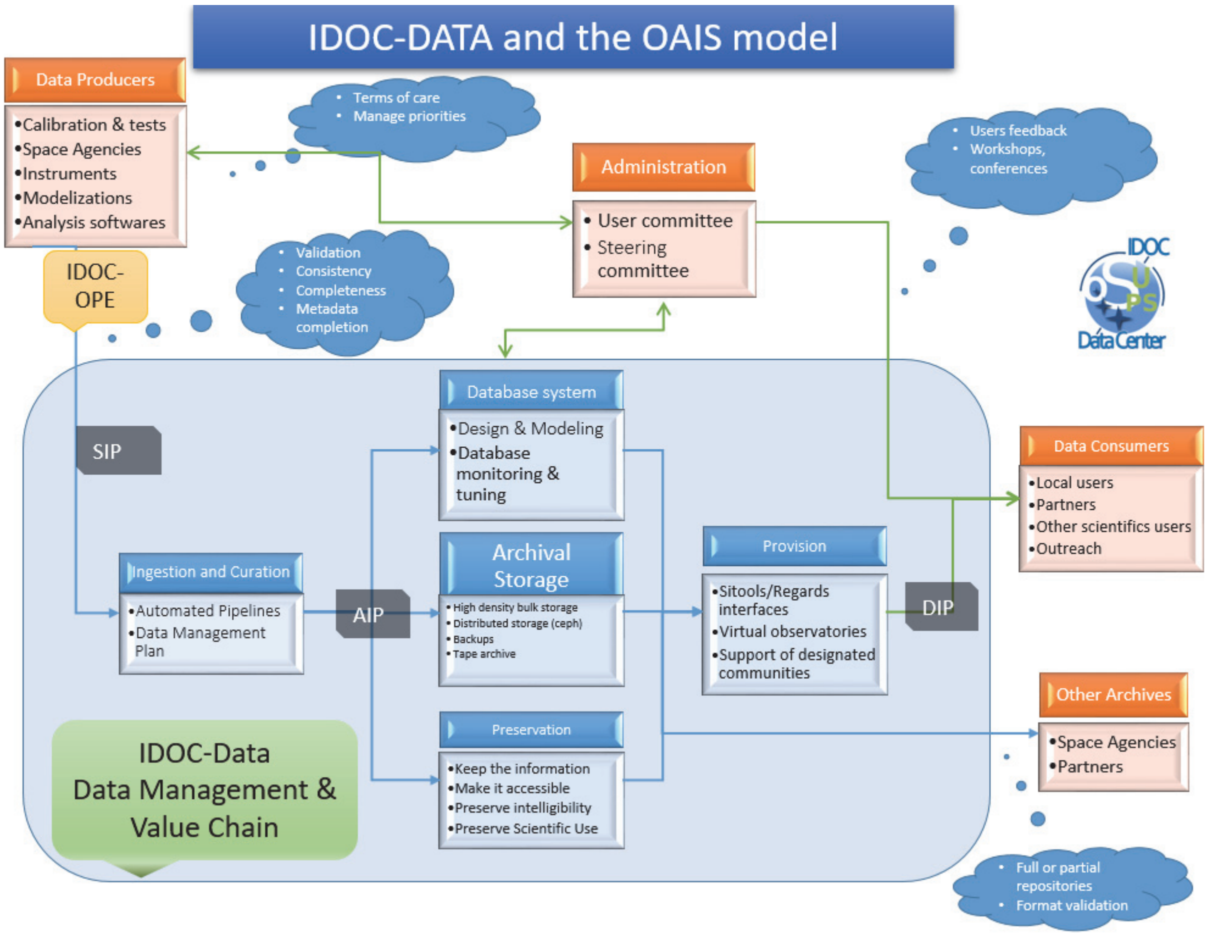


Figure 5 IDOC and OAIS

IDOC production monitoring

The adopted strategy aims towards the most possible automated monitoring of hardware and software elements, together with a qualitative analysis of products and services. This automated monitoring alerts IDOC team in case of problem. Those alerts are broadcasted by a mean which depends on its criticality and emergency: SMS, mail, ticket, etc.
The documentation explains the adequate way of solving the problem. This investment in writing articles describing the actions to be engaged in the event of a problem is a guarantee that the right decisions and actions will be taken, especially during critical periods.
This electronic documentation is also available on an autonomous hardware to ensure its availability in case of power shortage for example.

Monitoring of quality, performances, and evolution

To assure a given performance and quality for the delivered services, IDOC uses tools to monitor the production and to allow the analysis of the evolution of different performance criteria of the infrastructure. This applies to software development (e.g. : use of Jenkins, Icescrum, Sonarqube), production (running processing pipelines), and infrastructure management. Should a potential saturation of a resource be detected, solutions are elaborated, costed, and proposed to the project.
IDOC activities including a natural technology watch, innovative solutions can also be raised. After validation, extensions of resources are scheduled in agreement with the project.

Rights on data and information

IDOC-DATA makes data available to restricted communities and to public. Datasets are divided into 2 categories:

- Data from an instrument under the responsibility of a member of the OSUPS. These data can be “raw” or “n+1 level” processed (see 0) or data produced by scientists from OSUPS. IDOC is then the legitimate “owner” of these data. It owns the right to carry out operations and modifications of these datasets.
- Public data redistributed by IDOC-DATA. IDOC-DATA has then to be granted the authorization to make those data available. In specific cases, IDOC-DATA might request the data producer the authorization to modify metadata to improve the data management and its access.

Note that IDOC-DATA might also advertise scientific articles published by OSUPS scientist (e.g., The “news” sections on websites). In this situation, the related references to publications are clearly stated.

OVERALL DESCRIPTION OF IDOC “MODUS OPERANDI”

Since its creation, IDOC has chosen to build its hardware and software infrastructure on the most standard, open-source elements possible and seeks to limit the use of proprietary formats in all areas. These implies for instance the use of widely used open-source systems, platforms, or software packages:

- Linux (mostly Debian), Proxmox, Ceph, Mysql, Postgresql, Apache, ...
- System and network administrators are using the Redmine project management platform (restricted user access) to manage version upgrades and evolutions.
- IDOC employees receive continuous training and are active in professional networks that enable them to keep pace with the latest technologies and to deduce from shared experiences those that will be most useful for IDOC.
- IDOC's development plans are prepared with the help of user committees and amended and validated by the steering committees.
- All the information relating to the measures to be implemented in the event of incidents or accidents (disaster plan) is described on a dedicated website (for the sole use of IDOC operators for security reasons) in the form of clear procedures to be followed by those involved.
- An internal website is maintained, where each infrastructure element involved in IDOC services is described. Its implementation methodology is also detailed, as well as its maintenance actions, contact points (suppliers, experts, etc.) and incident resolution sheets describing the verification, first aid and resolution actions.

It should be remembered that in order to preserve IDOC's business continuity as much as possible, its resources are distributed over three sites, each provided with a separate, backed-up power supply, air conditioning, firewall and network. For the applications and data for which IDOC has availability commitments, the survival of only one of these sites enables the delivery of the associated services.

IDOC uses many tools, for example:

- for transmission, creation, and backups of datasets,
- for infrastructure monitoring, control, and operation,
- for service monitoring,
- for the supervising of software development,

with a permanent concern for the automation of tasks to make them as stable as possible.

IDOC is also monitoring the interface usage to report to user and steering committees on the cost-effectiveness of the efforts invested

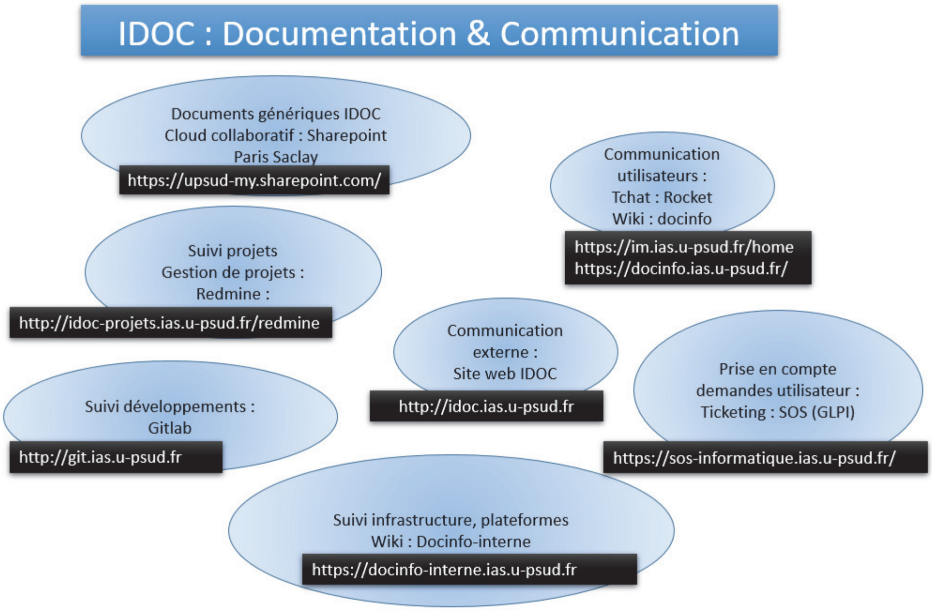


Figure 6 IDOC Documentation organization

IDOC obviously uses many more tools, for example,

- for transmission, creation, and backups of datasets,
- for infrastructure monitoring, control, and operation,
- for the supervising of software development.

These tools are described in [RD2].

RISK MANAGEMENT

In 2016, a global risk analysis has been performed at IDOC along various categories and is regularly updated (last update 2022):

- Risks related to the very existence of IDOC
- Risks related to the organization of IDOC

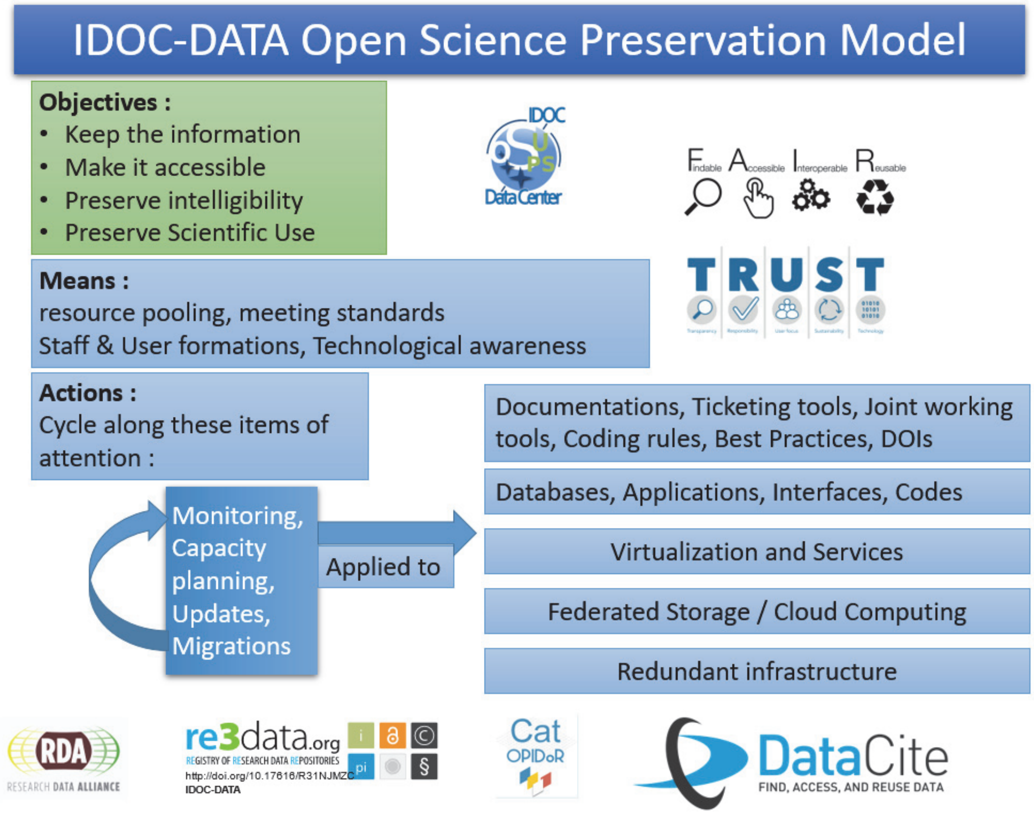
- Risks related to the technical facility
- Risks related to the IDOC network (physical and security)
- Risks inherent to operating hardware and software systems
- Risks raised through the design and the development of a new service
- Risks about the protection of the datasets and of the services
- Risks about the long-term preservation, evolution of standards and technologies.

Those risks are listed and for each, mitigations are description in [RD2].

IDOC-DATA PRESERVATION MODEL AND ITS STRATEGY

The general goal is to preserve the FAIR aspect of the data by placing IDOC-DATA in a TRUST approach for all its actions aiming at perpetuating the data itself as well as the accesses and the services around these data.

Therefore, the actions described below for the implementation of a new dataset for which IDOC-DATA is responsible for the long-term visibility, are subject to a recurrent evaluation that leads to actions updated by the evolution of the technical or scientific context of these data.
The information in this chapter 5 can therefore be summarized as follows:



The IDOC preservation model is detailed in [RD6]

6 ACRONYMS AND GLOSSARY

ACRONYMS

AIM	Astrophysique, Interprétation, Modélisation (AIM) (IRFU, CEA)
AMIS	Astrophysique de la matière Interstellaire (Astrophysics of Interstellar Matter)
COSMO	Cosmologie
DS2	Stellar System Data: one of IDOC's theme
FAIR	Findable, Accessible, Interoperable, Reusable,
GEOPS	Geosciences Paris Saclay

IAS	Institut d'Astrophysique Spatiale (CNRS, Paris-Saclay laboratory)
IDOC	Integrated Data & Operations Center
IRFU	Institut de Recherche sur les lois fondamentales de l'Univers
ITIL	Information Technology Infrastructure Library
IVOA	International Virtual Observatory Alliance
MEDOC	Multi Experiment Data & Operation Center
OAIS	Open Archival Information System
OSUPS	Observatoire des Sciences de l'Univers Paris-Saclay
P2IO	Physique des 2 Infinis et des Origines (French LABEX –EXcellence LABoratory)
PSUP	Planetary SUrfaces Portal: one of IDOC's theme

Glossary

Access interface

An access interface allows to select among the available datasets, those which are relevant to the interests of the user. The efficiency of the user interface should make this selection easy, which can concern data originating from several instruments.

Once the selection is done, tools can be used to improve the selection or to start to treat the information. Data can then be accessed or downloaded by the user for processing.

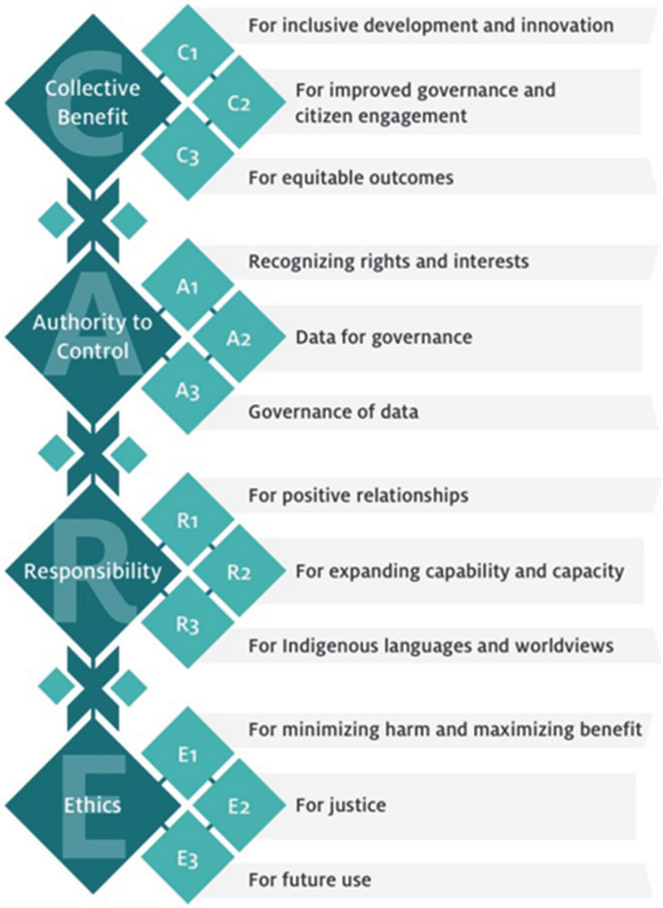
Features of the access interfaces are aimed to be available either interactively or from sytem-to-system (e.g., through web services).

Archival service (OAIS Archive)

Archival service aims to conserve information, to make them available to a user community in a comprehensive and usable form, and to enhance these data through added value services. This archival service is responsible for upload, management, maintenance of the data and their availability depending on the access level to be applied. This definition of archival service is a specialization as defined in the OAIS model (DA3) for this document.

CARE

Whereas the FAIR principles are data focused, the CARE principles are people and purpose oriented.



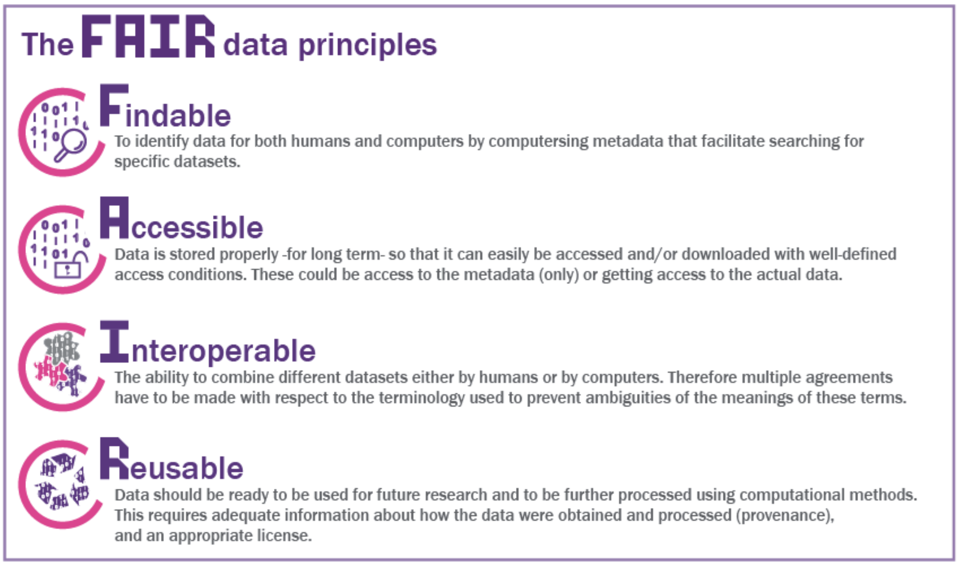
Dataset

A dataset gathers consistent data come from an instrument and combines the information required for scientific interpretation of observations made by the instrument. The dataset includes data to be distributed and/or archived, as well as the set of related elements: format, metadata, documents, tools, access interfaces, data bases, access rules for sharing and utilization, etc. The dataset should allow in a consistent and autonomous way the scientific use of its content by a person having the topical knowledge of the concerned field.

FAIR

The FAIR Guiding Principles for scientific data management and stewardship were first published in Scientific Data in 2016. They were developed to help address common obstacles to data discovery and reuse – long recognized as an issue within scholarly research and beyond.

The principles provide guidance for making data F indable, A ccessible, I nteroperable, and R eusable.



Pipeline

A data pipeline allows the production of a dataset or a new level of dataset. This pipeline must not be considered as a « pipe » of data transfer, but it must be considered as a «refinery » allowing the transformation from a set of source products of various origins to high level products.

TRUST

TRUST principles offer guidance for maintaining the trustworthiness of digital repositories, especially those responsible for the stewardship of research data. Guidance for each of the TRUST Principles is reproduced below.

Principle	Guidance for Repositories
Transparency	To be transparent about specific repository services and data holdings that are verifiable by publicly accessible evidence.
Responsibility	To be responsible for ensuring the authenticity and integrity of data holdings and for the reliability and persistence of its service.
User Focus	To ensure that the data management norms and expectations of target user communities are met.
Sustainability	To sustain services and preserve data holdings for the long-term.
Technology	To provide infrastructure and capabilities to support secure, persistent, and reliable services.

User community (present and future users)

User community is an identified group of potential users who can understand a given set of information. The target user community can be composed of several user communities.

Virtual observatory

Concerning the increase of data sources and their volume, the obvious interest of being able to cross this information leads to the concept of virtual observatory. This concept allows a user to specify his subject of interest in order to be able to collect and to work easily on all the available data concerning this subject, whatever their origin and their original form.

The technical and organizational framework of this concept is developed within the framework of IVOA (International Virtual Observatory Alliance), where expertise is shared, and the VO (Virtual Observatory) standards are discussed and formalized.

IDOC participates in these discussions and in the distribution and extension of these services.