

National data management strategy based on FAIR principles

Contents

Contents	1
Introduction	2
Purpose and vision	2
Background	3
Scope	4
The value of open and FAIR data	5
Target group and structure	5
Principles for a data management practice that supports FAIR	7
• Data management must support the FAIR principles and opportunities for data sharing and re-use.	7
• The relevance and weighting of the individual FAIR principles vary between research areas and must be defined within these	7
• It must be possible to store all types of research objects and file formats of relevance to the research in both the short term and the long term	7
• Data management is a continuous process throughout the data life cycle	8
• Relevant technical infrastructure and tools must be offered to all researchers, regardless of academic, professional, and institutional affiliation	8
• The necessary help and knowledge resources must be accessible regardless of academic, professional, and institutional affiliation	8
Actions for establishment of a data management practice that supports FAIR	9
Making data accessible and FAIR is a cultural change in many academic fields	9
Proposals for specific actions	9
Support for academic and professional groups in defining their own implementation of FAIR principles	10
Proposals for specific actions	10
Supporting preservation of research data in both the short term and the long term	11
Proposals for specific actions	12

Supporting that the research can provide continuous project documentation throughout its life cycle	12
Proposals for specific actions	13
Establishment of technical infrastructure that supports data management and accessibility of FAIR data	13
Proposals for specific actions	13
Establishment of knowledge infrastructure that provides necessary help and knowledge resources for data management and FAIR	14
Proposals for specific actions	15
Funding	15
Principles for distribution of expenses on data management and physical data repositories in connection with foundation-funded research projects	15
Research funding	15
The institutions' obligations	16
Funding and responsibility for establishment of the basic e-Infrastructure	16
Expenses for Data Management and physical data repositories	16
Principles for establishment of cost level and allocation in the two categories	18
Proposals for specific actions	18
The research-funding foundations' requirements for handling of research data	18
Governance and implementation	19
Glossary	19

Introduction

Purpose and vision

The purpose of the strategy is to establish a basis for decisions on implementation and financing of a data management practice based on the FAIR principles¹ (*Findable, Accessible, Interoperable, Reusable*, i.e. the data must be findable, accessible, and reusable, and the conditions for use of data must have been clearly formulated). The strategy must be in

¹ Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3:160018 doi: 10.1038/sdata.2016.18 (2016)

accordance with the EU's Open Data Directive² and the Danish implementation of this in the PSI Act³, and contribute to meeting research and societal expectations for increasing accessibility of research results.

The strategy must contribute to significant steps being taken to ensure that research data become more FAIR and that research data can in themselves be processed as recognized as research results that are shared in national and global data infrastructures in ways that stimulate open research and open innovation—with due consideration for necessary safeguards for both personal data and rights, commercial interests, and national security interests.

The overall vision for research data management is that researchers continuously make qualified decisions and actions in relation to maintenance of and access to their data (and any other results such as software, protocols or the like)—with due consideration for the promotion of research, the needs of society, and the existing legislation.

Solutions must be internationally anchored and meet international standards. Where recognized international solutions already exist—for example open source software—such implementation solutions should be preferred over own-developed solutions. This applies to the implementation of infrastructure and technical tools, as well as to the choice of metadata standards, vocabularies, etc.

Background

Society and research are increasingly using data as a decision-making basis and for analysis purposes, and it is therefore important to create opportunities to find, access, and re-use data, including that the conditions for re-using data are clearly formulated. The plan to develop a new national data management strategy to support that more research data become FAIR originates from the Strategy for National Cooperation on Digital Research Infrastructure⁴, and thus in an idea that data management based on the FAIR principles has its place in the national e-Infrastructure landscape.

In this context, the strategy also serves as the Danish accessibility policy under the EU's Open Data Directive (formerly the PSI Directive), which—in its latest version—also comprises research data and requires that the member states draw up policies for research data accessibility in accordance with the FAIR principles.

The objective is to make data as open as possible in accordance with the FAIR principles, so that data can be accessed and re-used by others to the greatest possible extent. However, it

² Directive (EU) 2019/1024 of the European Parliament and of the Council on open data and the re-use of public sector information (2019)

³ The current *Act on re-use of public sector information (Lov om videreanvendelse af den offentlige sektors informationer)* is being revised to implement the latest version of the Directive.

⁴ Strategy for National Cooperation on Digital Research Infrastructure, the Danish Agency for Science, Technology and Innovation (2018)

should be stressed that FAIR data are not the same as open data. There may be good reasons why data cannot be made openly accessible. However—with a few exceptions—data that cannot be accessed openly can still be made FAIR, as the researcher may publish metadata describing the actual existence of the dataset, as well as conditions for how others may be able to access the data, either the dataset in its entirety or derived versions thereof.

Denmark is a member of European Open Science Cloud (EOSC)⁵, and the expectation is that newly produced research data will be made accessible through EOSC. At the time of writing, an infrastructure for this is not yet in place, and the EOSC architecture is still being developed. The best way to achieve this goal is therefore to follow the recommendations from EOSC⁶ and leading international organizations such as Research Data Alliance (RDA)⁷ and the GO FAIR initiative⁸ regarding implementation of the FAIR principles. At the same time, it must be ensured that infrastructure and services are based on known standards and are continuously adjusted to any future requirements for architecture and interoperability in EOSC.

The strategy must also make recommendations on how to finance the handling of research data in accordance with the principles established and to make proposals for policies for FAIR data management for projects funded by the research-funding foundations. It can be stated overall that the strategy objective is to establish principles and the accompanying actions for a FAIR data management practice and its funding, as well as possible actions for the realization thereof.

The data management work under this strategy may involve a number of issues regulated by other legislation and local policies. This includes the Danish Data Protection Act (*Databeskyttelsesloven*), the General Data Protection Regulation (GDPR), the Danish Archives Act, safeguards, etc. It is especially stated in the explanatory notes to the revised PSI Act that the EU's General Data Protection Regulation and the Danish Data Protection Act will always take precedence over the provisions in the Open Data Directive and the PSI Act.

Scope

In accordance with Article 10 of the Open Data Directive and the Bill for the Danish PSI Act, the present strategy applies to publicly funded research data, meaning publicly funded research and research conducted by public institutions where there may be private funding. In accordance with this legal basis, the strategy is aimed at research institutions and research-funding organizations. In particular, the Directive stresses that research data which have been made accessible by researchers, research institutions, or research-funding foundations through an institutional or thematic repository must—as a general rule—be re-usable for both commercial and non-commercial purposes⁹.

⁵ <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

⁶ The report *Turning FAIR into reality*, the EU Commission, doi: 10.2777/1524 (2018), as well as policies and reports under publication from the current EOSC FAIR Working Group

⁷ <https://www.rd-alliance.org/>

⁸ <https://www.go-fair.org/>

⁹ The Open Data Directive, Article 10(2).

Research data are defined in the Open Data Directive¹⁰ as *documents in a digital form, other than scientific publications, which are collected or produced in the course of scientific research activities and are used as evidence in the research process, or are commonly accepted in the research community as necessary to validate research findings and results.*

In accordance with this definition, data in this strategy are perceived primarily in the conventional sense as the digital content that can be referred to in the form of datasets, files, and databases of various kinds. However, it should be stressed that the methods regarding FAIR can, to advantage, be used for all types of digital objects included in the research, be it logbooks, software, workflows and scripts, digital surrogates for physical objects, organizations, concepts, and much more. The strategy also includes the publication of data published in their own right, without necessarily being the basis of an article.

The strategy does not replace or set aside other applicable legislation in this area, for example the notification obligation under the Danish Archives Act.

The value of open and FAIR data

It is a fundamental principle of the sciences that researchers build on each other's results. In a world in which research is becoming increasingly digital, data must be easily accessible, so that it is possible for researchers to incorporate other researchers' data into their own research. A more extensive data basis provides new opportunities for analyses and models, e.g. climate models.

FAIR data will make research results more accessible to society in general, facilitate the research and development work of companies, and be of significance to citizens' performance of citizen science activities. It will also contribute to research transparency and thus legitimacy.

One of the prerequisites for these visions being realizable and benefiting research is that data are findable and identifiable, and that the conditions for re-use are known. These are some of the objectives of the FAIR principles, which are fully designed to enable re-use of data in a responsible manner, for the maximum benefit of research and society in general.

Target group and structure

The document lays down a number of principles for data management practices that support FAIR, followed by recommendations for actions aimed at strengthening the implementation of the individual principles. An attempt has been made to structure these actions in accordance with who are the primary players in the implementation thereof. The recommended actions often support more than one principle.

¹⁰ The Open Data Directive, Article 2(9).

The principles are broadly aimed at all stakeholders in such practice: researchers, data managers/data stewards, research managers, and research-funding foundations. The proposals for action areas are, however, primarily aimed at decision-makers in research-conducting and research-funding institutions that have the capacity and means to implement and finance the measures. This means persons who can assume responsibility for the implementation of the strategy, including for how it is organized and how the responsibility is distributed in the individual organizations.

As an actual implementation plan is expected under the terms of reference, the strategy does not comprise a proposal for specific implementation, but only proposals for *action areas*.

Principles for a data management practice that supports FAIR

- Data management must support the FAIR principles and opportunities for data sharing and re-use.
 - Data that are made accessible must be made FAIR by, as a minimum, being identified unambiguously by data reference (*Persistent Identifier* (PID)) and with documentation (metadata) added.
 - Implementation of FAIR is a step-by-step process, not all principles need to be met from the outset.
 - When data are assessed as having permanent value and are preserved, the data must be made accessible if there are no specific grounds against this.
 - As a general rule, data (metadata) descriptions can be made open, also in cases in which data are not openly accessible.
- The relevance and weighting of the individual FAIR principles vary between research areas and must be defined within these
 - The implementation of FAIR—in addition to the minimum requirements described above—is defined by the individual research areas at national level in relation to international standards.
 - The research areas must consider how far they want to go with the implementation and which FAIR principles they weight the highest.
- It must be possible to store all types of research objects and file formats of relevance to the research in both the short term and the long term
 - Data that are regarded as having a potential future value must be stored and made accessible in a technically and organizationally secure data infrastructure in accordance with the FAIR principles;
 - Particularly valuable data for continued storage for an extended number of years are identified and handled continuously with a view to any special steps for long-term storage.
 - In cases in which data that have been made accessible are erased, the PID should remain accessible and be able to generate metadata about the erased data as well as information that they are now inaccessible.

- **Data management is a continuous process throughout the data life cycle**
 - Methods and tools should be accessible and be able to ensure that data and documentation are collected and stored during the process.
 - The conditions (licensing) for re-use must be accessible as part of metadata, in either human or machine-readable form.

- **Relevant technical infrastructure and tools must be offered to all researchers, regardless of academic, professional, and institutional affiliation**
 - Data storage and accessibility in accordance with the FAIR principles must be available to all researchers
 - Data in relevant infrastructures must be able to serve as performance of the research institutions' obligations in relation to the Danish National Archives.
 - Selected technical tools must make data retrievable and FAIR in an open, standardized form. The use of data must not be prevented by the data being accessible using licensed software only.

- **The necessary help and knowledge resources must be accessible regardless of academic, professional, and institutional affiliation**
 - It must be possible for researchers to work to make data FAIR, without having to be experts in all areas of data management, by these researchers being offered help from the institutions.
 - The necessary knowledge about data management and FAIR must be accessible to researchers at all research institutions.

Actions for establishment of a data management practice that supports FAIR

Making data accessible and FAIR is a cultural change in many academic fields

A research world in which researchers are not only open about their methods and data—but also make them accessible in ways that make the data re-usable—will help broaden the perception of what is useful and usable research output. This has implications not only for the researchers' working methods, but also for a number of the surrounding structures, including:

- for the credit transfer system, which is to recognize these alternative outputs
- for the grants system, which is to recognize that making metadata and data re-usable entails an additional cost
- for the recruitment system, which is to recognize the need for new competences.

This will require action by the research institutions, which will contribute to a cultural change in several academic fields in relation to digital research practice and work actively in relation to the metrics used in evaluations and appointments.

Research councils and foundations will need to incorporate not only support for this type of activity, but also recognize the activities as credit-bearing activities.

Research institutions should stimulate and facilitate this change in culture, particularly in academic fields which do not currently have a tradition for sharing data outside a closed group, including

- That the research strategic importance of good data management is recognized. For an increasing number of sciences, research is today entirely dependent on data being collected, systematized, categorized, analysed (e.g. by computers), stored (e.g. in repositories or databases), and re-accessed continuously by international research colleagues, or their computers.
- Change the attitude to what is publishable and shareable,
- to facilitate an understanding that the FAIR principles do not necessarily mean open access to data, but that also closed data can be made FAIR by being linked to an identifier (PID) as well as descriptive metadata.

Research institutions should work nationally and internationally with the credit transfer system so that researchers' work with publishing, maintaining, and supporting the sharing of FAIR data is recognized and credited.

Proposals for specific actions

Strengthening of communication and dialogue—regarding Open Science and including FAIR.

Work to make data FAIR is made credit-bearing.

Support for academic and professional groups in defining their own implementation of FAIR principles

The individual academic fields have different traditions for how data are shared and structured. In addition, different academic fields have very different views on the importance and practical implementation of the individual FAIR principles. It is therefore important that the academic environments are those that define how and how far FAIR can be implemented in their field.

The importance and structure of data can vary widely both within and between academic fields. In some cases, this is about personal data, in others it concerns data that may have been created in a private-public partnership—which can be subject to confidentiality agreements—or data may be copyrighted, to name some examples. However, most research data can be made FAIR without major legal and ethical challenges.

Research institutions must support the different academic fields in developing/clarifying their profile in relation to international activities and—taking into account the costs—preparing a plan for how far they will come with the FAIR process within the next three years.

Research institutions and DeIC¹¹ must ensure that the right competences and tools are available to supplement the researchers in their formulation and definition of their profile.

Proposals for specific actions

The research institutions must ensure at national level that descriptions of good data management practices are prepared in different academic fields. As a minimum, these must contain a description of how metadata can be collected as an integral part of the work process and be a requirement in connection with conclusion of projects/activities. This may entail:

- *Preparation of short-term goals for making data accessible and sharing of data, for example based on methods regarding maturity models¹²*
- *Helping academic groups identify, develop, and/or publish standards, methods, tools, and infrastructure to make their data FAIR.*
- *Acceptance of the assessment of existing datasets in relation to re-use (source criticism) is an important discipline.*

¹¹ DeIC has been established as a collaboration between the eight Danish universities and the Danish Ministry of Higher Education and Science. The basis for DeIC has been described in 'A National Strategy for Research e-Infrastructure'

¹² For example CMMI Institutes *Data Management Maturity (DMM)*, (<https://cmmiinstitute.com/data-management-maturity>) or DCC's *RISE Model*, (<https://www.dcc.ac.uk/guidance/how-guides/RISE>)

Supporting preservation of research data in both the short term and the long term

Preservation of data after the conclusion of a project requires both human and technical resources. The best way to preserve data is to keep them alive, thus ensuring that they are relevant and accessible in accordance with the current standards. This requires stable funding and organization.

The primary aim of this strategy is to make research data accessible in accordance with the FAIR principles. In the nature of things, this requires that data be preserved, often in an infrastructure or data system operated by a research institution, DeIC, data preservation institutions, or by research groups. In the Open Data Directive, such data systems are called institutional or subject-based repositories. In addition to technical infrastructure and IT systems, the operation of such repositories requires data-literate staffing and therefore stable governance and funding.

As a general rule, short-term preservation (5-10 years) will consist in the operation of a repository with the necessary level of governance and security to ensure confidence in the contents as research objects. In some—but not all—cases, this will involve certification in accordance with an international standard.

Many data will lose their relevance and be erasable after some years, but some data with particularly high scientific or societal value will need to be preserved in the long term. In some cases, this may require special measures such as assessment of formats and conversions, as well as the need for stable long-term organization. Over time, more stringent requirements will presumably be made for data to remain FAIR and understandable outside the original framework of understanding.

Where—after consultation with the researchers—the Danish National Archives receives data for long-term preservation under the Danish Archives Act and makes these data accessible in accordance with the FAIR principles, the requirements of this strategy can still be said to have been met.

In addition, the recommendation in the first five years after the strategy has entered into force is to focus on short-term preservation in order to take significant steps to preserve data on a large scale and generally make such data accessible. The issue of specific requirements for infrastructure, organization, and funding for long-term preservation should be re-addressed before 2025.

Research institutions should work with the various academic and professional groups to map existing stable data storage solutions and identify where there are no such solutions. In addition, principles and methods should be developed to decide which data are to be stored and which can be erased.

Research institutions must contribute to the establishment of a practice for continuous data management, also after projects have been concluded.

Research institutions and DeiC must ensure that the operation of institutional and national data repositories has sufficient long-term funding and continuous ongoing protection of data, including identification of data for long-term preservation.

The Danish National Archives must contribute to ensuring that long-term preserved and FAIRified research data can be included in a future joint national data infrastructure.

Proposals for specific actions

Collaboration—including with international players—on the development of methods and infrastructure for long-term preservation of research data adapted to research requirements for diversity in object and file types, as well as scaling in relation to data volumes.

Supporting that the research can provide continuous project documentation throughout its life cycle

All information, structures, metadata, etc. necessary for data and other research outputs to be sharable with others are created as an integral part of the research process. The relevant considerations are implemented at the start of a project, and more and more research institutions and foundations now demand that a data management plan be prepared, for example in connection with applications. Good data management also helps ensure reproducibility of the research.

On conclusion of a project, there should be follow-up on the data management plan, where it is reconsidered what should be stored, where, and for how long. If data are submitted to the Danish National Archives, this is regarded as being in compliance with the strategy requirements. Are there scripts and methods that can—to advantage—be shared with others, and how?

Research institutions and foundations must consider the requirements they make for documentation of the research process.

The Danish National Archives must develop procedures so that—by providing standardized FAIR metadata—research institutions can meet their notification obligation.

Research institutions must ensure that all researchers recognize their responsibility in connection with data management and support the performance thereof.

Research institutions and DeiC must facilitate the widespread use of data licences.

Proposals for specific actions

Recommendations must be prepared for methods to ensure continuous documentation of research data in line with the data being collected and generated, so that all necessary information is present when results are to be made accessible.

In connection with handling of confidential data, it must be supported that it is necessary to perform a risk assessment in relation to the choice of data storage solutions.

Universities should stimulate and facilitate a discussion of the risk of open data—for example in relation to intellectual property rights, patents, or unauthorized use.

National guidelines on the use of data licences must be prepared and disseminated, on the basis of and in collaboration with international players.

Establishment of technical infrastructure that supports data management and accessibility of FAIR data

Researchers at universities and other public research institutions must have access to the necessary technical infrastructure for collection, storage, and accessibility of data. What is relevant will depend on the project, the size of the project, and the level of internationalization. As research is often inter-institutional and international, this infrastructure can—to advantage—be seen as part of a global ecosystem, including in EOSC.

DeiC must identify and make proposals for the parts of the technical infrastructure which can—to advantage—be solved nationally, taking into account the international development as well as national and local needs and experiences. Calls for solutions should subsequently be implemented as soon as possible.

The Danish National Archives, research institutions and DeiC will collaborate on ensuring coherence and interaction between their respective technical infrastructures and processes in connection with notification and possible submission of data to the Danish National Archives.

Proposals for specific actions

Specification and construction of Danish parts of the technical infrastructure:

- *Establishment of a generally accessible nationally trusted data repository. By depositing data in this system, the researcher can store and expose metadata and data, link a PID to data and metadata, and attach a licence¹³.*
- *Specification and offer of any other components nationally that are requested by the researchers, for example PID systems, ontology database, or other components that can facilitate the work with making data FAIR.*

¹³ This action has been proposed in the memo to DeiC's Board of Directors, Data Management in Denmark, 2019

- *Extending the national service for data management plans with academically customized templates as well as with machine readability, thus making data management plans FAIR and possibly also usable in connection with notification to the Danish National Archives.*
- *Ensuring that recommended services and infrastructures are certified in accordance with international standards to the necessary extent.*

Establishment of knowledge infrastructure that provides necessary help and knowledge resources for data management and FAIR

The changes towards the research becoming more data-driven and more international affect research methods and require new expertise and specialized support functions in addition to new technical tools. These traditions are already established in many fields of research. For others, it will involve a big step from recognizing the importance of data management to being an active part of infrastructures and continually creating the necessary metadata. In many contexts, there have been calls for the establishment of a new profession, ‘data stewards’, who contribute to facilitating data management processes and making data FAIR.

‘Data stewards’ with in-depth knowledge of data management issues, methods, and technical tools must be available and enter into research projects—where appropriate—in line with, for example, laboratory technicians, programmers, engineers, or librarians. This field of competence can be found in variants ranging from very general competence at a more advisory level, to subject-specific competence at the highest academic level, which is fully integrated in the research.

Research institutions and DeIC must ensure that the necessary competences for FAIR data management are developed and made available during the research process.

Research institutions and DeIC must ensure through national and international collaboration that the right competences and tools are present to support researchers in their implementation of FAIR data. Competences in Data Stewardship have several dimensions and can:

- A. be understood as level-divided academic and professional competence, by ranging from qualified IT staff or librarians, to specialized BSc or MSc graduates, to also being seen as academic competence that is integrated in and by the research groups as research competence on an equal footing with researchers’ other academic and professional competences (i.e. integrated at peer level as co-author of articles, etc.),
- B. be understood as having different content, covering a continuum from general (generic) data management to the more subject-specific,
- C. be understood as fundamentally different—in both form and content—in the various main areas (technical sciences, natural sciences, health insurance, social science, and humanities).

DeiC must ensure national critical mass and coordination, and facilitate knowledge exchange across universities and academic fields.

Research institutions and DeiC must support the development and use of maturity models and metrics in relation to FAIR, in accordance with emerging international standardization, primarily in RDA and EOSC.

Proposals for specific actions

As part of DeiC's proposed¹⁴ back office function, a national data stewardship competence centre is to be established for coordination of local initiatives, and to support competence development and large-scale interdisciplinary projects.

As part of the universities' front office function, local data stewardship competence centres are established to support institutional researchers.

Relevant continuing education and training in FAIR, data management and data stewardship must be developed and offered.

Interdisciplinary collaboration on tools, methods, and exchange of experience must be coordinated nationally.

Knowledge of data management and the FAIR principles must be made available at relevant stages of the researchers' study programmes, for example in the PhD schools.

Funding

Principles for distribution of expenses on data management and physical data repositories in connection with foundation-funded research projects

Research funding

Preparing and maintaining data in accordance with FAIR principles and then making data openly accessible—if there are no arguments to the contrary—make requirements for new research processes, new technical infrastructure and tools, new competences, and support functions. By using the principles for FAIR and open data, it is expected that research investments can be utilized more effectively. But there are costs connected with the implementation thereof. The additional costs must be covered by those that already pay for the research projects, i.e. primarily research institutions and private and public research-funding foundations.

¹⁴ Proposed in the memo *Data Management in Denmark*, 2019

The institutions' obligations

The institutions are obliged to service the researchers with a basic e-Infrastructure, which, for example, makes it possible to create, use, and store research data securely. The basic infrastructure must be able to handle all types of data and comply with national and international acts and regulations. Competent user guidance and competence building are a fundamental part of the e-Infrastructure.

Funding and responsibility for establishment of the basic e-Infrastructure

In future, a significant part of the e-Infrastructure will be supplied by Danish e-Infrastructure Cooperation (DeiC), as a national infrastructure with data repositories and large-scale computer systems. The whole thing is bound together by the Research Network— including connection to European Open Science Cloud (EOSC) and access to processing power at EuroHPC— including the LUMI centre, in which Denmark has a special interest.

The costs are paid by the eight universities and the Ministry of Higher Education and Science in accordance with a further agreed distribution scale.

Locally, the universities have full responsibility for the establishment of IT solutions to ensure access to national and international resources for researchers and students, as well as to establish support functions—front office—where users can get help.

The costs for handling data according to the principles described in this strategy will vary considerably, depending on the size of the projects and the area of focus. Some can naturally be handled in the basic infrastructure, others will—due to scale and complexity—necessitate a significant drain on resources that exceeds what one can expect from the basic e-Infrastructure. It is a prerequisite for the strategy that the basic infrastructure with adequate supply of resources can be expanded and thus supply the necessary e-Service for large and complex data projects. Only in extremely rare cases should it be necessary to build up new systems from scratch and to incur expenses for maintenance and operation of these by the institutions.

Expenses for major extensions of the basic infrastructure—in relation to hardware, data management, and increased drain on support functions—must be included in the project budgets. In addition, expenses that the institutions incur for the establishment of e-Infrastructure must be included as any co-funding of projects by the institutions.

Expenses for Data Management and physical data repositories

Expenses for data management infrastructure which the host institutions must establish and maintain to support the projects, can roughly be divided into two categories:

1. Basic IT infrastructure that the University must in any case establish, for example networks, print facilities, basic software, security, and the FTEs necessary to keep the basic IT

infrastructure functioning and secure. The basic infrastructure also includes data management and general data storage as exemplified in Table 1.

Expenses are paid by the institutions, as it must be regarded as a necessity for the maintenance of study programmes and research on a level with international partners that there is a basic e-Infrastructure and is therefore part of the basic budget of the institutions.

Some projects will draw on the basic IT infrastructures to such an extent that it will be necessary to inject additional resources to cover the e-Infrastructure needs of the project. These may, for example, be general support functions and licences, as well as upward adjustment of storage capacity.

Expenses related to the project’s increased use of the basic e-Infrastructure **can be included in the budget for the project application**.

2. Parts of the e-Infrastructure that are solely connected to a given project—for example data stewards attached to the project—and the establishment of large long-term repositories (archives) that preserve data beyond the project period.

The expenses **are included in the project application budget**.

Table 1. describes in further detail the individual elements of the two categories.

Category	Software	Storage space	FTEs
1.a General IT infrastructure at the institutions	Basic	Basic EOSC infrastructure EMBL network Storage capacity after end of the grant period (Data from the project)	Maintenance Development Data managers GDPR Cross-institutional harmonization Control of FAIR principles
1.b Additional IT infrastructure at the institutions connected to research project or grant	Specific licences Special software	Very large draw on data repositories	Greater draw on the institution’s staff to develop data management and make data FAIR

2. IT linked to a grant or project	Specific licences Special software	Very special needs for long-term archives	Own data stewards
------------------------------------	---------------------------------------	---	-------------------

Principles for establishment of cost level and allocation in the two categories

It must be both transparent and easy for applicants to determine which expense items can be included in the application budget. There may be a difference in what can be applied for from the various foundations, for example due to different practices regarding overheads.

Proposals for specific actions

Establishment of a list of expenses for e-Infrastructure, which is updated on an ongoing basis. The list is directional, but is accepted by foundations and universities as a starting point for contractual negotiations.

The list is drawn up by a committee with representatives from Universities Denmark and the research foundations.

The research-funding foundations' requirements for handling of research data

Data management requirements are equated with other statutory and contractual rules applicable to foundation-funded projects. The host institutions are responsible for compliance with the rules. The researchers thus cannot set aside rules that data must be made FAIR on the grounds that data are produced in connection with a foundation-funded project. This applies to projects supported by funding from both private and public foundations.

The reporting to the foundations usually comprises a list of publications related to the project. In future, the list must also include a statement of the datasets established under the project identified by a PID (Persistent Identifier) and with a statement of whether data have been made accessible to other researchers after the results have been published, i.e. whether they comply with the Act on Open Data.

Governance and implementation

The steering committee for the drafting of this strategy will prepare an implementation plan as soon as the strategy has entered into force. As a first item in this plan, a steering committee will be set up to handle the implementation, with representation from the institution types and ministerial areas covered by the strategy. This group will be responsible for the ongoing follow-up and reporting of the implementation. The implementation plan lays down the composition of the steering committee as well as the frequency of follow-up and reporting.

The implementation plan must also contain guidelines for how to measure the implementation process for follow-up purposes.

Glossary

FAIR	F indable, A ccessible, I nteroperable, R eusable
FAIR Data Point	Data system that exposes data and metadata in accordance with the FAIR principles
Metadata	Data describing other data: identification, description/documentation, history of creation, licence, etc.
PID (Persistent Identifier)	Unique identification of digital resource. Must normally be translated into a specific website address via a PID service (e.g. DOI, ORCID)
Repository	Database of 'digital objects', extensive data, metadata, and PID, often with a searchable user interface and computer interface (API)
Trusted repository	Repository certified in accordance with a recognized standard—for example Core Trust Seal—which deals with organizational, technical, financial, and legal matters, etc. in relation to the repository.
API	Application Programming Interface (API) is a technical interface directly connected to a computer's software program. It defines the types of calls or requests that can be made, how to make them, the data formats to be used, which data are to be returned, etc.
FAIR metrics	Methods for examining the extent to which the FAIR principles have been met. Work is being done internationally—including in EOSC—to establish joint standards for FAIR metrics. FAIR metrics must primarily be seen as a tool available to the researchers to make data more FAIR.
Maturity models	Maturity models can be used by a research project or subject area to define the desired objective of FAIRification. FAIR metrics can then be used to evaluate compliance, so that the results are weighted in relation to the targets set.

Confidential data	Data which cannot be made accessible under an open licence, but which must be kept confidential in respect of intellectual property rights, protection of personal data, privacy, security, and/or legitimate commercial interests.
-------------------	---