DESCRIPTION OF THE INFRASTRUCTURE AND ITS ACTIVITIES

The Swedish National Data Service (SND) supports and promotes efforts to make research data produced in Sweden open and FAIR (findable, accessible, interoperable, reusable). The infrastructure is a knowledge hub for research data management at the centre of a system for disseminating and exchanging know-how, expertise, and experience. We initiate, create, and participate in networks and collaborations with stakeholders from the national and international research data community. Through SND, users gain access to technical tools and services that facilitate describing, sharing, and reusing FAIR data.

1. A SWEDISH SYSTEM FOR FAIR AND OPEN DATA

In collaboration with Swedish research organisations, SND will be the leading national provider of FAIR and open research data to the scientific community. Achieving the national goal of free access to publicly funded research data, as set by the Swedish government, requires local research data support functions at Swedish research organisations. These functions, referred to in this application as Data Access Units (DAUs), have been established with SND support during the current funding period (2018–22). Over the next period (2023–26) the majority of DAUs are expected to reach maturity, and SND will continue to support their further development and collaboration.

Although not an organisational part of SND, the DAUs are our closest partners in supporting researchers in data-related issues. The DAUs have joined the SND Network, a central part of the national system that promotes and facilitates research data sharing and preservation, as well as collaboration and mutual assistance. This system will give Swedish researchers access to the tools and services they need to describe data and make them findable and accessible, as well as to information and training resources. The collaboration between DAUs and SND in reviewing data and metadata before publication is unparalleled internationally and builds on a tradition of using distributed competence and federation, as well as forming national consortia. Technical and procedural solutions within the collaboration allow for handling of personal data and data with other restrictions. Through the SND Network, DAUs and other research infra-structures will continue to be able to exchange knowledge and receive training. Together, the SND consortium and the DAUs in the SND Network will constitute an essential data access service for the Swedish research community.
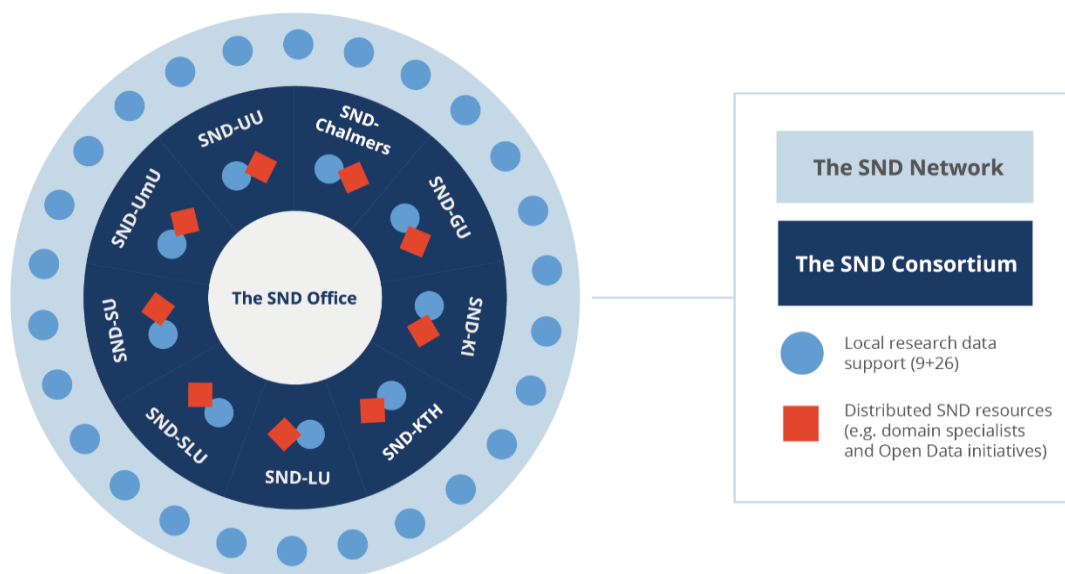

Figure 1.1: Overview of the SND Consortium and the SND Network.

2. ORGANISATION AND LEADERSHIP OF THE INFRASTRUCTURE

2.1. THE SND CONSORTIUM: A MODEL OF DISTRIBUTED EXPERTISE

The partners of the SND consortium are University of Gothenburg, Chalmers University of Technology, Karolinska Institutet, KTH Royal Institute of Technology, Lund University, Stockholm University, Swedish University of Agricultural Sciences, Umeå University, and Uppsala University. During the current funding period, the partners provided experts in a range of scientific domains as their co-financing of the infrastructure. These *domain specialists* are a national resource for Swedish universities and offer expertise on specialised aspects of research data and Open Science for the research community. They widen the scope of SND's expertise and address the "lack of attention to the specificity of research practice, processes and data collection" [30: p.3][1] that had been identified as a key overarching issue in achieving open access to research data.

In order to address the rapidly changing landscape of Open Science in general and access to research data in particular, the SND domain specialist model will be further developed. While some research domains move rapidly toward data sharing, other domains face technical or legal obstacles due to the types of data they use. Support and solutions need to be adapted to the particular challenges that each research field faces as the national 2026 deadline for free access to research data "as a rule" approaches. The domain specialist model will be augmented with a greater possibility for targeted, time-limited, *Open Data* initiatives. Such initiatives can involve the development of national guidelines, best practices, or processes necessary to provide access to particular data types, for example for interview data, video material, or data with sensitive information. They can assist a particular research group, centre, or other research environment, national or local, to progress toward producing FAIR and open data, and use their experience as a template for similar research communities on a national level. This kind of strategic initiative will be called *Open Data Flagship*, and will provide a way for researchers to set good examples on how to make data FAIR and open. The need for necessary metadata schema and other tools and services can be communicated directly to SND from the Flagships, and the Swedish research community will benefit from their experiences. The Open Data Flagships will be managed locally or across universities and will receive data management support from the SND office and from local DAUs.

The activities of this augmented domain specialist model will be divided over all modules but with a particular focus in the *Knowledge Hub* and *Networks and Collaborations* modules (described in sections 3.4.1 and 3.4.2). The distributed SND organisation will maintain close ties to the various domains and to scientific communities of practice.

2.2. CONSORTIUM ORGANISATION AND LEADERSHIP (MODULE 1)

Operational management, administration, IT and technical development, and expertise in data curation are mainly located at the SND office at the University of Gothenburg. Another office, staffed by local domain specialists, may be established in the Stockholm–Uppsala region during the next funding period, to promote closer cooperation between domain specialists.

2.2.1. Strategic Management

Strategic management is the purview of the steering committee, constituted by one representative from each consortium partner and the director of SND. The steering committee is chaired by the host university representative. It is guided by an annual general assembly, representing the leadership of the consortium universities. The general assembly provides input on matters such as the strategic plan, annual budgets, and significant changes to operations (see figure 2.1 for an organisational overview).

Two further groups advise the steering committee. The Research Data Support Council links the steering committee to the national activities at the DAUs. The council strengthens the collaboration in the consortium and advises on data curation requirements at a local level. An

---

[1] For references, see *Appendix E: Key References*.

Infrastructure Advisory Group will be established, with the main purpose of advising the steering committee on data requirements for the research community. The group will consist of representatives from publically funded, national research infrastructures. This stakeholder representation will ensure that functions, tools, and standards developed within SND are apposite to researchers who use such infrastructures in Sweden.

The consortium structure puts SND at great advantage, as nine of Sweden's top ten research universities are part of it. In 2019, the consortium universities accounted for more than 80 per cent of the total revenues for research and postgraduate education, and almost 80 per cent of the doctoral students were active at these universities (full-time equivalents).

SND consortium members are driving forces in several national collaborations on research data and Open Science, and belong to the consortia behind URFI (the Universities' Reference Group for Research Infrastructures) and SNIC (Swedish National Infrastructure for Computing). Through the SND consortium, we are able to share and learn from each other's top-level expertise, research, personnel, and technical skills. We have a close-knit, tight cooperation and are able to speak with one voice through SND. Members of the consortium pool resources and work together to more efficiently meet the needs of their faculties and research staff with regards to Open Data and data management issues.
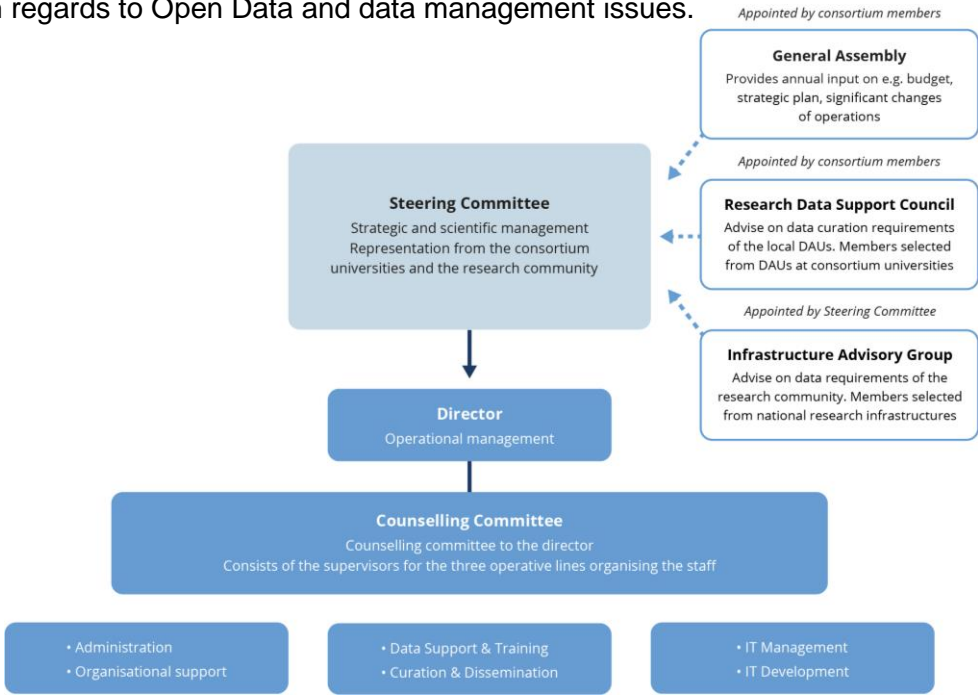

Figure 2.1: Operative Organisational Structure for SND.

2.2.2. Operative Management

The operative management of SND is arranged into a line and staff organisation, under the executive management of the director. The consortium agreement will establish the procedure for how the director is appointed by the host university. The three lines, each with a staff supervisor, are (i) Administration and Organisational Support; (ii) Data Support and Training, Curation and Dissemination; and (iii) IT Management and Development (see figure 2.1).

The director appoints the staff supervisors, who supervise the daily work in the operation lines and provide specialist advice to the director as members of the counselling committee.

2.2.3. Administrative Support

Most administrative support in human resource management, finances, communication, legal issues, and project management is provided within the SND organisation.

Required Resources: A total of 100% director, 120% administrators, and 180% communication officers are allocated for the administration.

3. INFRASTRUCTURE OPERATIONS
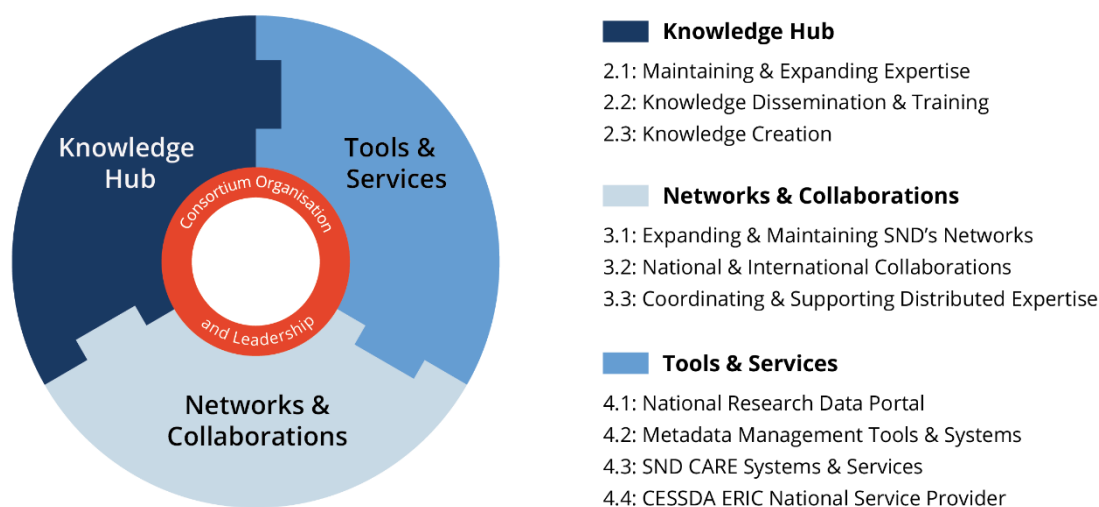
3.1. GRAPHIC OVERVIEW

Knowledge Hub

2.1: Maintaining & Expanding Expertise
2.2: Knowledge Dissemination & Training
2.3: Knowledge Creation

Networks & Collaborations

3.1: Expanding & Maintaining SND's Networks
3.2: National & International Collaborations
3.3: Coordinating & Supporting Distributed Expertise

Tools & Services

4.1: National Research Data Portal
4.2: Metadata Management Tools & Systems
4.3: SND CARE Systems & Services
4.4: CESSDA ERIC National Service Provider

Figure 3.1: Graphic Overview: SND Modules and Submodules.

3.2. TIME SCHEDULE

Please see attached Gantt chart in Appendix A: Time Plan.

3.3. USERS

SND has three main categories of users: **research data producers** (mainly researchers), **research data users** (researchers, businesses, public authorities, etc.), and **research data support** (DAU staff and other data professionals). The next funding period will see an increase in activities towards the first two categories and a clearer separation of tools, services, and resources. The separation is prompted by the expected increase in the number of users, the greater number of collaborative initiatives to meet the needs of these users, and the rising complexity of research data issues. The transition to two websites will be the most obvious outward sign of this separation (see 3.3.1: Communication and submodule 4.1).

SND's relation with research data producers centres on the tools, training, and support they need in order to produce well-documented data that can be shared in a FAIR and open manner. This includes measures that tie in with the mandated work of funders and universities to create incentives for Open Science [4]. Certain SND resources for the producer category are also valuable to research data users, such as training in research data management and data management plans. SND will further meet the needs of research data users through the various functions and tools incorporated in the *researchdata.se* (*forskningsdata.se*) portal, including instructions on how to use them. Research data support will continue to receive assistance through the tools, training, and coordinated networks that also facilitate exchange of experience and expertise, and which provides SND with information about user needs. More SND resources will be dedicated to work with data support staff in data-producing research infrastructures, with the goal of making their data FAIR and open.

Some infrastructures have already contacted SND for assistance, support, and collaborations. They require different services, but there are certain tools (e.g. a search function for

4

survey variables or questions) that SND can develop cost-efficiently for several of these small research infrastructures that are in contact with us, such as (by February 2021):[2]

- NEAR (National E-Infrastructure for Aging Research)
- The Oden Research Platform
- REWHARD (RElations, Work and Health across the life-course – A Research Data infrastructure)
- ETF (Evaluation Through Follow-up)[3]
- SITES (Swedish Infrastructure for Ecosystem Science)
- CORS (Comparative Research Center Sweden)

3.3.1. Communication

The main communication portals for SND users will be *researchdata.se* and *snd.gu.se*. The *researchdata.se* (starting in 2023) portal will contain information, training resources, specific search tools, and search functionality for the SND data catalogue, as well as for a number of other research infrastructures with data repositories. Its primary users will be research data users and producers. The current *snd.gu.se* site will become a website focused on information, resources, and tools for research data support users and the SND consortium. It will also include links to platforms for interactive collaboration, such as the internal communication platform Basecamp and the "DAU Handbook," a wiki (a community-edited web platform) for knowledge exchange in the SND Network.[4] Through these, research data support staff can find, and contribute to, resources on, for example, data review processes and guides to establishing a DAU.

SND will also continue to use social media (Twitter and LinkedIn), newsletters, network meetings, and national and international conferences to communicate with its user categories, other stakeholders in the research data community, and with the public. Other forms of communication are included in DORIS (the SND Data ORganisation and Information System), support activities (via online drop-in sessions, e-mail, and phone calls), and teaching and training (e.g. courses, webinars, hands-on training) (see submodules 4.1 and 4.2 for details).

3.3.2. Training and User Support

SND currently provides training and user support to both researchers and DAU staff through workshops on data management plans, instruction videos, drop-in support webinars on using DORIS and other systems, and presentations covering a range of aspects on research data management and related information. Online resources for basic data management for researchers[5] have been developed, as well as an introduction to research data management and curation for DAU staff.[6] Together with University of Borås, SND offers a 120-hour professional development course for DAU staff.[7]

During the next funding period, SND will intensify collaborations with DAUs from the consortium universities to create training and education in data management for junior researchers and PhD students. These pedagogical resources will then be shared within the SND Network. Another aim for the next funding period is to provide training resources for all of the tools and services that SND will develop during the period, including a comprehensive training guide for DORIS and *researchdata.se* (see submodule 2.3).

---

[2] NEAR https://www.near-aging.se/, Oden https://polar.se/en/about-polar-research/icebreaker-oden/, REWHARD https://www.su.se/rewhard/, ETF https://www.gu.se/utvardering-genom-uppfoljning-ugu/, SITES https://www.fieldsites.se/, CORS https://cors.se/en/widget-en/
[3] They use their Swedish acronym, UGU (Utvärdering Genom Uppföljning), for their application to the Swedish Research Council.
[4] The SND Wiki Handbook is at https://dhb.snd.gu.se/
[5] Our online resources on data management can be found at https://snd.gu.se/en/manage-data
[6] The BAS Online learning material is on https://snd.gu.se/sv/hantera-data/utbildning/bas-online
[7] For details: https://www.hb.se/Bibliotekshogskolan/Utbildning/Uppdragsutbildning/Forskningsdatahantering/

3.3.3. User Fees

Researchers use the SND tools and services free of charge. Universities with DAUs that are members of the SND Network pay an annual user fee to SND. In return, SND offers e-systems (DORIS), meetings, training and information on data management and curation. Additional services, such as data curation, will be charged at cost. Infrastructures and other parties that use SND, including universities which have not yet established a DAU, will be charged hourly fees for the specific services that they require (see attached budget for figures).

3.4. CONSTRUCTION, DEVELOPMENT, AND OPERATION OF THE INFRASTRUCTURE

SND works to allow researchers in Sweden to share and reuse FAIR and open data. Our goals (detailed in the *Scientific Plan*, section 1.1) are to

1. **facilitate the sharing of research data** in a way that is simple, secure, and trust-worthy, and that complies with requirements for data preservation, access, and sharing;
2. **provide data visibility and usage metrics** by exposing research data both nationally and internationally, and offering information on sharing and citation;
3. **facilitate trustworthy data access** to findable, high-quality, well-documented data, along with information on how those data can be accessed and reused.

SND's objective is to be a driving force in building a system that enables these goals, using our data management experience and expertise.

SND is described through four modules. Module 1 was addressed under section 2.2: Consortium Organisation and Leadership, above. Below (under sections 3.4.1–3.4.3) are addressed the other three, highly interdependent, modules, divided into ten submodules The submodules in each module describe the activities related to a particular aspect of SND's operations.

Module 2 *Knowledge Hub* includes activities that cover the building, maintenance, creation, sharing, and dissemination of expertise concerning research data management and curation generally and within particular scientific domains. Module 3 *Networks and Collaborations* comprises coordination and initiation of professional networks related to data management support and research infrastructures, participation in national and international collaborations, and project support and management. Finally, Module 4 *Tools and Services* includes development and maintenance of tools and services related to the new national research data portal, to metadata management, to the SND certified repository systems (SND CARE), and to SND's role as the national CESSDA ERIC Service Provider.

As the modules are activities-based, they are largely interdependent – the activities in one module often presupposes activities in the other two, or activities from two or all modules intermesh. The modules thus represent a high-level overview of the activities in SND, not a division of these activities into independent work packages. For example, creating a webinar for SND Network members on how to use new DORIS functionality involves activities from all three modules: it requires network activities (submodule 3.1), knowledge dissemination (submodule 2.2), and DORIS development (submodule 4.2). And DORIS development, for example adding new controlled vocabularies, is in itself a result of knowledge created in collaboration with, for example, the DDI Alliance (submodules 2.3 and 3.2), and technical skills acquired through professional development (submodule 2.1). Most SND activities are the result of similar chains of activities connecting the modules.

3.4.1. Module 2: Knowledge Hub

To be able to continue to offer solutions to the many challenges of data accessibility, SND will maintain a high level of expertise and know-how within areas such as research data management, data curation, and data management planning. By gathering and analysing information about the multitude of factors and actors in the research data ecosystem, SND will be able to anticipate and deal with potential problems, in the same way that we foresaw the need for a distributed curation model and a national storage solution for the current funding period. The

networks established during the current period will remain important venues for stakeholders to bring pressing questions and issues within research data management to SND's attention, as will the two reference groups presented in this application (see section 2.2.1: Strategic Management). Open Data Flagships, initiatives that are part of the consortium's in-kind contributions (see section 2.1: The SND Consortium), will provide an additional channel for researcher communities to communicate their data management and sharing concerns to SND.
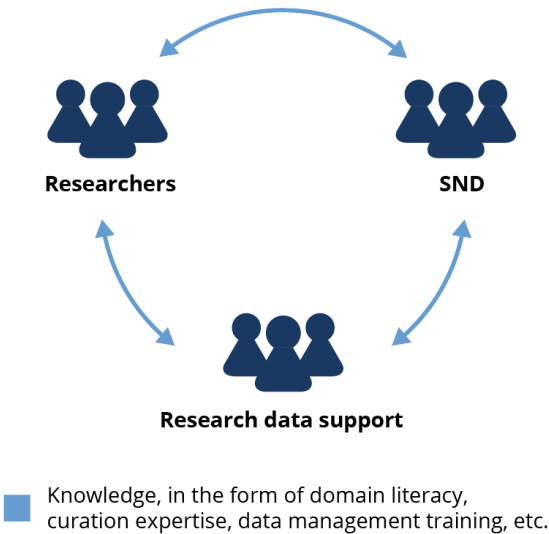
The submodules under the *Knowledge Hub* module capture activities related to (1) the maintaining and building of expertise and know-how; (2) the sharing and dissemination of such knowledge; and (3) the creation of new knowledge.

Required Resources: A total of 500% research data advisors, 40% legal officer, 200% IT, and 300% domain specialists are allocated to the Knowledge Hub.

Submodule 2.1: Maintaining and Expanding Expertise

The landscape of research data management, curation, and sharing will continue to evolve rapidly. New data types, metadata requirements, means of access, and new technologies will require continuous activities aimed to maintain and increase relevant expertise and know-how.

Contribution to goals*: By maintaining and expanding expertise, the activities in this submodule contribute to *facilitating the sharing of research data*. To provide a system and services that make data sharing as simple and secure as possible, while in compliance with stakeholders' requirements, requires maintaining appropriate expertise.

Knowledge, in the form of domain literacy, curation expertise, data management training, etc.

Figure 3.2: Knowledge Exchange. *SND operates in a context where all actors learn from each other's experience and expertise. Knowledge flows are always bidirectional.*

For SND, efforts will include professional development through courses, seminars, and conferences, as well as literature reviews of reports, articles, guidelines, best practices, and other publications on administrative methods of data management and curation, and discipline-specific data management methodology. Collaborations with other organisations, such as DAUs and other national and international research infrastructures, will provide opportunities for hands-on experience and knowledge exchange through visits and projects. Such collaborations can offer new perspectives and ideas on how to improve processes and workflows in an SND or Swedish context. In-house development, not least of methods for managing data types previously not encountered, will provide valuable experience.

Submodule 2.2: Knowledge Dissemination and Training

SND is a repository of know-how related to data management, curation, and sharing. Crucially, this role involves coordinating the sharing of expertise within SND's various networks, in particular, the SND Network of DAUs, and among researchers and data professionals.

Contribution to goals*: By knowledge dissemination and training, the activities in this submodule contribute to *providing visibility* of, and *facilitating trustworthy access* to, shared data. The collaboration and sharing of expertise enables a common and equal service for all

researchers when describing, depositing, sharing, and finding quality research data. The SND Network plays a key role in the day-to-day support for researchers across Sweden.

SND will carry out a broader dissemination of knowledge through physical and virtual courses, seminars/webinars, workshops, and learning material. User training of DAU staff will continue as we also increase activities aimed at other data professionals, junior researchers, and doctoral students. SND will also offer advice and curation support for DAUs and other research infrastructures.

Submodule 2.3: Knowledge Creation

In order to fulfil the role as a Swedish Knowledge Hub, SND will initiate and participate in activities that aim to contribute with new knowledge regarding data curation and management. This engagement is necessary in order to develop solutions to future challenges in the rapidly changing research data environment.

Contribution to goals: By knowledge creation, the activities in this submodule contribute to *facilitate the sharing of research data*. To make it possible for researchers at Swedish universities to comply with the requirements of access to FAIR data, it will not be enough just to follow developments in research data management. SND must be active in creating solutions, especially solutions that are compliant with Swedish regulations.

Collaborations, networks, and projects with national as well as international partners will yield new knowledge regarding data curation and management. SND will, for example, contribute to the development of guidelines, best practices, standards, and relevant terminology and controlled vocabularies; and thus, help create the necessary knowledge for managing, curating, and sharing FAIR high-quality research data in an uncertain future.

3.4.2. Module 3: Networks and Collaborations

Through a long history of collaborations and project participations as well as through its consortium partners, SND is a solid part of a national and global research infrastructure. Active in the Swedish research data community, SND has a central role in several national networks. The *Networks and Collaborations* module covers many current and future activities necessary to manage memberships in, and collaborations with, national and international organisations, authorities, and infrastructures. The submodules aim to ensure that Swedish research data practices are developed in a wide national framework and in an international context. At the same time, these methods and processes will contribute to the national aim of making Sweden one of the leading countries in research quality, with access to research data that promotes greater scientific transparency.

A vast majority of the *national cooperation* is connected to the SND networks. SND's *international commitments*, on the other hand, are dominated by other forms of collaborations, e.g. EOSC and organisations that facilitate FAIR research data. The *Networks and Collaborations* module captures activities related to (1) maintaining the existing networks that make up the Swedish research data infrastructure, and initiating and building new networks as required; (2) the collaborations required to provide a national and international context and the essential services required for FAIR data sharing; and (3) the coordination and support of collaborative activities, targeted initiatives, and domain expertise distributed across the consortium partners.

Required Resources: A total of 500% research data advisors, 100% IT, 300% domain specialists, and 50% domain specialist coordinator are allocated to Networks and Collaborations.

<u>Submodule 3.1: Expanding and Maintaining SND's Networks</u>
SND's vision is to be part of a global network through which researchers can easily share, find, access, and reuse high-quality research data. To realise this vision requires several networks on different levels and with different purposes and types of expertise.

Contribution to goals*:* By expanding and maintaining SND's networks, the activities in this submodule contribute to *facilitating the sharing of, and trustworthy access to, research data*. They also help *providing data visibility*. Through the various SND networks and the SND consortium, a frictionless and high-performing national research data infrastructure can enable researchers to preserve, share, and reuse data in a simple and secure way.

At the heart of the SND networks is the SND consortium, a partnership of nine major research universities dedicated to facilitating the Swedish objective to make all publicly funded research data accessible from 2026. During the current funding period, SND has initiated networks on archival, legal, and IT issues, as well as networks dedicated to establishing a common terminology and controlled vocabularies, and to develop training programs and material related to data management and data sharing. These networks will be expanded and developed further in the next funding period, and new will be added.

The SND Network of DAUs is vital to the SND operations. By 2023 this network is expected to have over forty members at Swedish universities and research organisations, trained in research data management and curation, as well as in data management plans and metadata review. During the next funding period, SND will intensify support for the network by additional training and professional development, researcher training, and knowledge exchange among the members. Through the Research Data Support Council, the SND Network members provide input to SND on a strategic level; on an operational level, contacts with SND are already well-established through SND's research data advisors.

New networks are planned for the next funding period. A network for research infrastructures that receive funding from the Swedish Research Council is part of SND's strategy to broaden its domain expertise. This network will provide strategic input to SND through the Infrastructure Advisory Group (see 2.2.1: Strategic Management), and regular contacts within various domains will be established to facilitate knowledge exchange. To support individual data professionals (managers, curators, and stewards of research data) who work in projects and infrastructures, a national forum and network will be initiated for exchange of experience and expertise, including SND-run workshops and training sessions. This network will be slotted in with the SND Network in order to bring DAU staff and other data professionals together for the advancement of data-management work.

<u>Submodule 3.2: National and International Collaborations</u>
In the next funding period, SND will continue its wide-ranging collaborations and maintain its close connections to other research data infrastructures and e-infrastructures. While the Infrastructure Advisory Group will advise the consortium management on matters of strategic importance, national and international cooperation is necessary for solving practical problems.

Contribution to goals*:* By national and international collaboration, the activities in this sub-module contribute to *facilitating the sharing of, and trustworthy access to, research data*. They also help with *providing data visibility*. These relationships will ensure that Swedish researchers have access to national and international datasets for reuse, and their research data will gain international exposure, leading to a broader awareness of Swedish research.

SND's cooperation with SUNET and SNIC is essential to providing secure and efficient storage solutions for research data. These three national infrastructures aim to provide Swedish researchers with accessible, trustworthy, and easy-to-use tools and services regarding

research data. With DiVA (the Academic Archive Online)[8] and RUT (Register Utiliser Tool),[9] SND maintains important collaborations to make research data findable (see submodule 3.1: *Expanding and Maintaining SND's Networks*.) SND's domain expertise will increase by intensifying our contacts with Swedish ERIC (European Research Infrastructure Consortia) nodes and other similar consortia.

SND will continue to facilitate easy access to international data for Swedish researchers, as well as to promote the use and citation of Swedish research data internationally. The partners needed for these tasks are mainly the large-scale EU initiative the European Open Science Cloud (EOSC) and related initiatives funded by the EU; and various organisations that provide essential services or expertise related to data management and repositories.

In 2020, with the University of Gothenburg as host university, SND was accepted as member of the **EOSC Association** along with eight other Swedish organisations. EOSC was created to promote access to European and global research data. SND's major contributions to EOSC have been through participation in the projects *EOSC-Nordic* (focus on the Nordic and Baltic countries) and *SSHOC* (Social Sciences and Humanities Open Cloud). SND communicates a Nordic-Swedish perspective on data management, e.g. the implications of Swedish personal identification numbers. In the next funding period, SND will also provide a national entry point to the distributed research data system through the *researchdata.se* portal.

Additional international collaborations will be organisations central to providing trustworthy FAIR data or necessary for data publication. Developers of metadata standards, international promotors of data-sharing, and contributors to findable and citable data will remain essential to SND's international network, adding a global context for Swedish research data.

- Membership in **DataCite** allows SND and members of the SND Network to issue research data with Digital Object Identifiers (DOIs), a type of persistent identifier for finding, reusing, and correctly citing data.
- Membership in **World Data Systems** means that data shared through the SND research data catalogue can attain greater international impact. Their standards and guidelines concern sustainable and secure data storage, management, and access.
- **The Inter-university Consortium for Political and Social Research** (ICPSR) is a research infrastructure with almost sixty years' experience in curating and providing access to research data. SND's membership allows Swedish researchers free access to the ICPSR data collections and the opportunity to take part in their summer school for a reduced fee.
- The global, member-driven **Research Data Alliance** (RDA) promotes easy sharing and reuse of data for researchers and innovators. RDA Sweden, the Swedish node, is coordinated by SND, and promotes common solutions and the FAIR principles.
- By collaborating with **the DDI Alliance**, SND contributes to the Data Documentation Initiative (DDI), an international standard for describing data produced, for example, in the social, behavioural, and health sciences.

Submodule 3.3: Coordinating and Supporting Distributed Expertise

The collaborative activities, targeted initiatives, and domain expertise distributed across the consortium partners will need national coordination to provide maximum benefit to the infrastructure, and these various activities will also require support from the SND office.

Contribution to goals: By coordinating and supporting the distributed expertise of the SND consortium, the activities in this submodule contribute to *facilitating the sharing of research data*. The products and distributed expertise will help establish a research data system that is simple to use, secure, and trustworthy for all disciplines, and this submodule is essential in maximising, facilitating, and directing the outcomes from that work. Some contributions will also be made to the other goals, for example by products that enhance visibility of data for

---

[8] To visit the DiVA website, follow this link: https://info.diva-portal.org/
[9] Learn more about RUT at https://rut.registerforskning.se/

particular domains and by the Open Data Flagships that lead to tools that will improve access to data.

As detailed under section 2.1: The SND Consortium, the consortium partners will contribute resources to a distributed organisation. While most of their activities, including the Open Data Flagships, will fall under other submodules, this submodule contains the activities necessary to coordinate the many people involved in activities, locally or as collaborations with consortium partners. Such coordination will ensure that initiatives do not duplicate each other's efforts, and that the results from activities are of national relevance. This submodule also encompasses support activities from the SND office, as resources and expertise are contributed to the distributed activities. This includes a national coordinator for the distributed initiative.

3.4.3. Module 4: Tools and Services

An important part of SND's support to universities and researchers in their creation, management, and sharing of data is the development of tools and provision of services for various data management tasks. Such tasks range from creating data management plans at the beginning of a project to documenting and describing datasets for future use. The support also includes e.g. means to track how much data are shared, discovery and searchability assistance, and tools and best practices for data curation and review.

Required Resources: A total of 500% research data advisors, 600% IT, and 200% domain specialists are allocated to Tools and Services.

Submodule 4.1: National Research Data Portal
The main SND service aimed directly at researchers will be a national research data portal, *researchdata.se*.

Contribution to goals: By the national research data portal, the activities in this submodule contribute to *facilitating trustworthy data access*. In the portal, researchers can find high-quality, well-documented data for secondary use and research validation. The portal will also include clear instructions on how data can be accessed and reused.

The *researchdata.se* portal will contain tools, services, and information primarily designated for researchers (see section 3.3.1: Communication). The use of a new web address for this portal will be part of establishing a "neutral" space, not tied specifically to a given university, infrastructure, or consortium, where the focus instead lies with research data and the researchers who use it. At the heart of the portal will be its search functionality. It will be possible to search for research data from the SND research data catalogue, but also from many other research infrastructures. The SND catalogue contains metadata that describe data produced by researchers who work at universities or research organisations in the SND Network. These metadata are created through DORIS (see submodule 4.2) or harvested via a metadata ingest platform and harmonised. A steering group for *researchdata.se* will include representation from SND, the SND Network, and the Infrastructure Advisory Group.

Other tools will be added to the platform. A public search API will allow other organisations to build applications based on the existing metadata, thus maximising exposure for Swedish research. A public harvesting service will permit external services to harvest metadata. One example of an external service is *dataportal.se*, the national metadata portal for public data run by DIGG, the agency for digital government. Specialised search tools will also be developed, beginning with a survey question bank and a database with variables from quantitative studies. (Tools development will be based on the needs arising in the Open Data initiatives, in particular the Open Data Flagships; see section 2.1: The SND Consortium).

The *researchdata.se* platform will also contain data management training material and instruction material for how to use the various tools.
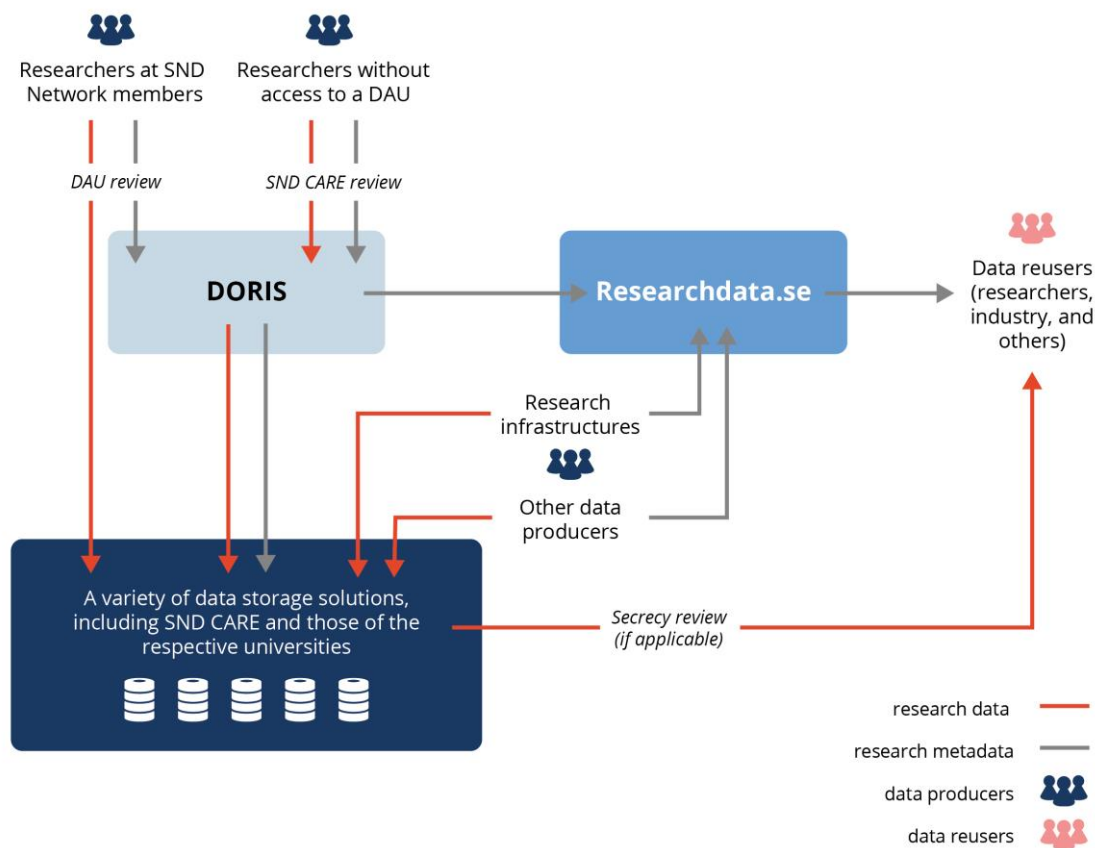
Figure 3.3: Key Services for Sharing Research Data. *Data are kept on a storage solution controlled by a university, infrastructure, or research organisation. Researchers describe data in DORIS. Data descriptions from DORIS and other infrastructures can be found through* researchdata.se, *and data can be downloaded directly or ordered. For sensitive data, the data owner performs a secrecy review before granting access.*

Submodule 4.2: Metadata Management Tools and Services

During the next funding period, SND will continue to develop user-friendly and interoperable tools and services for research data management, and maintain the tools developed during the current period. Foremost among these tools is *DORIS*, a national workspace for researchers, DAUs, and SND staff.

Contribution to goals: By metadata management tools and services, the activities in this submodule contribute to *providing data visibility and usage metrics*. The system for metadata management enables researchers to make their data visible, and allow their data to be used by others. The system will also provide tools for data-sharing and -citation metrics.

DORIS is used for creating, editing, and reviewing research data descriptions before they are added to the data catalogue, and it will be augmented with a metadata extraction service. This service, which will be developed by SND but run where the data are stored, will extract metadata directly from compatible data files in order to make data description quicker and easier for researchers. DORIS also contains SND's entry-level, DataCite-compliant, metadata profile for registration of research data, along with functionality to improve the FAIRness of entry-level data descriptions. This profile allows rapid descriptions of data as a means to provide data owners with an overview of their data, and it is made available to researchers as requested by their university. Entry-level data descriptions are clearly marked as such until sufficiently augmented. (This process was introduced to meet researcher and university

demand for quick initial data registration and to maximise access to research data, while promoting FAIR data.)

Other tools that will be added include:

- a research impact service with connection to citation databases;
- a restricted write API to allow machine-actionable updating of metadata in DORIS (for example, to support continuous publication of metadata from scientific instruments);
- a controlled vocabulary management tool that enables a more effective handling of pre-defined lists of terms used in the services provided by SND.

Integration with other external resources has been proposed, e.g. with storage providers, RUT, DiVA, publication databases, SweCRIS, and local services at universities. How such integrations will be prioritised will depend on the forthcoming SND evaluation (2021) and will ultimately be decided by the SND steering committee under the advice of stakeholders, advisory groups, and consortium members. The development of the services will be coordinated on a national level, and with an international connection to EOSC.

Submodule 4.3: SND CARE Systems and Services

SND CARE (CentrAl REpository) is an OAIS-compliant and certified repository with support for FAIR data and metadata, including the systems and services connected to SND's repository function.

Contribution to goals: By the SND CARE systems and services, the activities in this submodule contribute to *facilitating the sharing of research data*. SND CARE enables researchers from research organisations that have not yet established a DAU to share data simply, securely, and in compliance with data access requirements.

SND will continue to develop its repository function with the aim to preserve research data, and to facilitate their reuse. The repository has secure research data storage (currently contracted from SNIC) and includes a dissemination service for datasets deposited with SND. Such datasets can be stored in SND CARE for legacy reasons, or can be deposited with SND from research organisations that do not yet have their own research data support function. For such organisations, SND CARE functions as a "default DAU," offering advice and support to individual researchers for a fee (see also 3.3.3: User fees). It will serve as an example for other repositories in the SND Network, and provide assistance as research organisations certify their own repository functions. The aim is to create a nation-wide network of the certified repositories in the SND Network.

Submodule 4.4: CESSDA ERIC National Service Provider

SND will remain a Service Provider for the Swedish membership in CESSDA ERIC in the next funding period. CESSDA is a European Research Infrastructure Consortium with some 20 European data archives that work together to improve access to social science research data.

Contribution to goals: By the activities as a CESSDA ERIC National Service Provider, this submodule contributes to *facilitating the sharing of, and trustworthy access to, research data*. It also helps *providing data visibility*. The CESSDA membership grants Swedish researchers access to well-documented and high-quality European research data, expert knowledge on data preservation and access, and training resources in several fields. The CESSDA data catalogue and the vocabulary service allow researchers to make their data visible and reused internationally.

The CESSDA Statutes call upon its Service Providers to fulfil a number of obligations listed in Annex 2 of the Statutes [29]. This includes enabling metadata harvesting for inclusion in the CESSDA data portal; contributing to cross-national data and metadata harmonisation activities; sharing of data repository tools and expertise with other CESSDA members and

prospective members, especially in countries with immature or fragile national infrastructures; and adaptation to the DDI metadata standard.

SND is active in many of the CESSDA areas, and at present involved in the work with certification of repositories, a guide to data archiving, the creation of a common technological platform, and in providing support to aspiring and new CESSDA members. SND will also play a crucial role in the further development of the European Question Bank. Previous involvement has been in, for instance, the work with persistent identifiers and the development of a common metadata profile for CESSDA. SND is also responsible for the Swedish language translations of the multilingual ELSST thesaurus and the common controlled vocabularies in CESSDA.

3.5. COOPERATION WITH OTHER INFRASTRUCTURES

SND is the only Swedish infrastructure dedicated to providing access to research data from all scientific domains, along with a broad expertise in data management and curation. Some Swedish research infrastructures provide access to research data but only in specific disciplines. Examples include Nationella Språkbanken Text (for mainly Swedish text corpora), SEAD (environmental archaeological data), and ICOS-Sweden (data related to the carbon-balance). Such infrastructures are natural partners for SND, rather than competitors, and there are long-established collaborations in place. At present (January 2021), SND harvests meta-data from Språkbanken Text in order to make their data findable in the SND catalogue. For the next funding period, metadata will also be harvested from other Swedish research data infrastructures, including ICOS, SEAD, and Bolin Centre for Climate Research.[10]

But far from all scientific disciplines have dedicated data infrastructures, nationally or internationally. Nor are there any other Swedish research infrastructures that take a com-prehensive approach to the research-data landscape in the way that the SND consortium does together with the members of the SND Network. There are, however, international infra-structures similar to SND. For example, SND has a well-established relationship with national data repositories in the Netherlands (DANS), Germany (GESIS), and the UK (UKDS).[11] Our collaborations with national data repositories in the Nordic countries have been of particular value due to the similarities in research-data-related legislation. European national data repositories have much to learn from each other and are very important partners to SND. Being national repositories, however, they do not, as a rule, accept Swedish research data and are not alternatives for most researchers in Sweden.

Other international options include digital, open-access repositories such as Figshare and Zenodo, which are accessible to data producers worldwide. While many of these appear to function well on a technical level and their depositing process is fairly swift, SND has a globally recognised trust seal, superior metadata schemas through domain-specific metadata profiles, and, importantly, functionality to describe and share datasets containing personal data. The personal data issue is complex, and from a legal and institutional perspective, certified Swedish data repositories would always be required. The fact that SND can offer domain-specific support for researchers and DAU users, including data and metadata review, will remain important as Sweden moves towards its 2026 target.

3.6. RISK ANALYSIS

The overall risk level in SND is estimated as low. The consortium partners have substantial scientific and economic interests in a long-term commitment to a national multidisciplinary research data infrastructure. The possibilities for a financially sustainable expansion during the funding period and beyond are also considered favourable, thanks to the dedication of the universities and research organisations in the SND Network. SND will systematically analyse risk areas during the funding period in order to detect possible problems early on.

---

[10] Visit https://spraakbanken.gu.se/; https://www.icos-sweden.se/; https://www.sead.se/; https://bolin.su.se/
[11] DANS https://dans.knaw.nl/en, GESIS https://www.gesis.org/home, UKDS https://ukdataservice.ac.uk/

Table 3.2: Risk Analysis

Risk description	Likelihood (1–5)	Impact (1–5)	Risk* (1–25)	Risk management actions
Difficulties in recruiting personnel with relevant expertise	4	4	16	Strategic human resource management; development of training programmes; active recruitment through established contacts
Loss of key personnel	3	4	12	Strategic human resource management; well-defined and documented workflows not dependent on single staff members; improved career opportunities within the organisation
Difficulties in ensuring access to data in case of terminated funding from 2026	2	4	8	Strategic plan for sustainable funding; succession plan including actions for access and preservation of research data
Problems related to the implementation of a distributed organisation	2	3	6	Well-defined and documented workflows and structures for collaboration; short and well-defined decision-making processes on all levels of the organisation
Underestimation of necessary resources	2	3	6	Options for additional funding will be explored; resource allocation will be reviewed
Delays of deliverables and milestones	2	3	6	Regular monitoring of time plan; rescheduling and revision of time plan if necessary
Management failure (unfortunate decisions)	1	5	5	Well-defined organisational structures; reporting lines and responsibilities will be reconsidered; problems will be referred upwards in the management chain
Conflicts of interest between consortium member organisations	2	2	4	Steering committee actively resolves potential conflicts of interest among the consortium members
Changes in the legal framework	1	3	3	Monitoring of legal framework development; cooperation with other stakeholders to adapt to changes

* Risk is estimated by multiplying likelihood by impact. Low risk = 1–8; Medium risk = 9–15; High risk = 16–25.

4. DATA MANAGEMENT AND THE NEED FOR A SUPPORTIVE E-INFRASTRUCTURE

The SND infrastructure facilitates the describing, sharing, and preserving of research data. Within the framework of these core activities, no data are currently generated, only metadata. Were data to be generated within collaborations and projects, they would be managed in accordance with appropriate best practices and the FAIR principles, with data management plans in place. Unless prevented by legal or ethical reasons, data to which SND will contribute will be made openly accessible in the SND CARE repository (see submodule 4.3).

In 2020, SND introduced DORIS, an integrated online service for describing and reviewing data (see *Appendix C: Activities during the Current Funding Period*). SND's previous data submission infrastructure allowed any researcher to submit data, but metadata were reviewed using separate offline software with data files uploaded to SND servers. However, DORIS provides functionality that permits both the DAU and SND to review metadata and data, as well as functionality for university-specific data storage. For researchers, DORIS provides multiple metadata schema to enable detailed, domain-specific data descriptions. For users searching for data, tools will be added that make it possible to, for example, find the same variables in several datasets, or to search for particular questions in multiple surveys.

Although SND can manage large amounts of data of various types, mainly metadata will be stored in the SND systems. Normal office equipment and existing data networks will suffice for the majority of SND's activities. SND services are hosted on virtual servers or as individual containers run on an Openshift cluster, both rented from and managed by the University of Gothenburg. SND CARE requires storage space, which will be provided by either SNIC Swestore or SUNET STaaS. Data stored as part of SND CARE will not include any personal data, nor will SND manage any personal or other sensitive data.

The SND data system requires vast data storage for its operations. Members of the SND Network need to be able to store data that their researchers describe and deposit using DORIS. With SUNET, SND has developed an API that allows users to publish data, stored at a university-controlled storage allocation, for instance SUNET STaaS, via DORIS. Where applicable, a link to the file location will be added to the metadata record to enable direct download. Should a research organisation prefer another storage solution, they can connect that storage to the API.