

Strategic Plan 2022–2026

Including Work Plan for 2022

version 27 September 2022

Contents

| | |
|--|----|
| Introduction..... | 3 |
| Aim | 3 |
| Vision and Mission | 3 |
| Background..... | 4 |
| Overview of NBIS activities | 5 |
| Overall strategy | 6 |
| Infrastructure..... | 6 |
| Data management..... | 6 |
| Vision (Desired future position) | 6 |
| Mission (Overall objectives and the approach to reach those objectives)..... | 6 |
| Systems Development & Tools..... | 7 |
| Compute and Storage..... | 9 |
| Interactions with other research infrastructures | 10 |
| Local interactions..... | 12 |
| Support..... | 13 |
| Prioritisation..... | 14 |
| Training..... | 15 |
| Information and Outreach | 17 |
| Collaboration with industry..... | 18 |
| Organisation | 19 |
| Electronic Meetings and Retreats..... | 20 |
| Topical meetings..... | 20 |
| Professional development of staff | 20 |
| Affiliated NBIS persons | 21 |
| Funding | 21 |
| International | 22 |
| ELIXIR node | 22 |
| Nordic collaboration..... | 22 |
| Appendix 1 – Abbreviations | 24 |

Introduction

NBIS (National Bioinformatics Infrastructure Sweden) is a distributed national research infrastructure with support from Science for Life Laboratory (SciLifeLab), the Swedish Research Council, Knut and Alice Wallenberg Foundation, and Swedish universities. NBIS is hosted by Uppsala University. NBIS constitutes the SciLifeLab Bioinformatics Platform.

NBIS is formed by 11 partners: Uppsala universitet (hosting NBIS), Chalmers tekniska högskola AB, Göteborgs universitet, Karolinska Institutet, Kungl. Tekniska högskolan, Linköpings universitet, Lunds universitet, Naturhistoriska riksmuseet, Stockholms universitet, Sveriges lantbruksuniversitet, and Umeå universitet

The NBIS Strategic Plan serves to formulate the long-term general goals for NBIS. It is discussed at the NBIS Board (Styrgrupp) and decided by the NBIS Annual General Meeting (Stämman). The current version is covering the period from 2022 to 2026 and was approved by the Annual General Meeting on 23 March 2022. The organogram on page 19 was updated to reflect the current support organisation and approved by the Board on 27 September 2022. The strategic plan will be revised at least annually, considering input from NBIS partners, the International Advisory Board and the Reference Group. The plan describes the major activities foreseen within the areas of Infrastructure, Support, Outreach, Training, Organisation and International.

The detailed goals to be achieved during 2022 are highlighted in green boxes.

Aim

NBIS provides excellence in bioinformatics support to researchers in Sweden, enabling world-class life science by offering support, infrastructure and advanced training. NBIS constitutes the SciLifeLab Bioinformatics platform and forms the Swedish node in ELIXIR (the European infrastructure for biological information). The organisational structure allows for changes in support needs over time as new techniques are developed and utilised.

NBIS supports both research groups not having their own bioinformatics resources and large established research groups with their own bioinformaticians needing specialised expertise. As life scientists get successively more educated in bioinformatics, the scope of NBIS will be shifted towards a focus on more advanced expertise.

Vision and Mission

Enable world-class life science research and maximise scientific and societal impact of collected data by:

- Providing expert knowledge, innovative data integration, advanced training, efficient data publication for open science, and access to high-performance data analysis methods
- Coordinating bioinformatics support within Sweden and making bioinformatics easily accessible for life science researchers
- Swiftly responding to changes in support needs as new techniques are developed and utilised

- Forming the Swedish ELIXIR node and participating in relevant international projects

Background

NBIS (National Bioinformatics Infrastructure Sweden) with its extensive experience in large-scale data analysis is in a unique position to be a key driver for data-driven life science. NBIS was established in 2016 after fusion of four infrastructures to form **a single point of contact** for all users needing bioinformatics support. NBIS thereby has all types of bioinformatics support within one organisation, facilitating user contacts and enabling efficient service provision. NBIS provides excellence in bioinformatics support to researchers in Sweden, **enabling world-class life science** by offering expertise, infrastructure and training. The topics covered by our experts include genome assembly, genome annotation, genetic variation, comparative genomics, phylogenomics, transcriptomics, proteomics, metabolomics, systems biology, single-cell biology, biostatistics, systems development, data management, image analysis and multi-omics integration. Furthermore, NBIS forms the Swedish node in ELIXIR (the European infrastructure for biological information). The organisational structure allows for changes in support needs over time as new techniques are developed and utilised.

NBIS is a **distributed infrastructure** with staff placed at all major universities in Sweden, creating access points to facilitate contacts with our users. Nevertheless, all projects are **nationally prioritised** and allocated to achieve best possible efficiency and competence matching. All NBIS staff are placed in excellent scientific environments in order to keep up-to-date with front-line achievements in bioinformatics and to create a critical mass at each site.

NBIS supports both research groups without bioinformatics resources and large established research groups with their own bioinformaticians needing specialised expertise. As life scientists are getting successively more educated in bioinformatics, the scope of NBIS is continuously shifting towards a focus on very advanced expertise. The advantages of an infrastructure like NBIS are:

- 1. Guarantee for excellence** – NBIS can provide a multitude of expertise, which is difficult to match by a locally employed bioinformatician.
- 2. Long-term stability** – NBIS is a sustainable resource, in contrast to PhD students or post-docs who need to move on when the time-limited position is ending.
- 3. Effective use of resources** – The large scope of many projects puts an increased demand on the ability to analyse the data effectively, using considerable programming/scripting skills, needed to automate much of the analysis. NBIS provides data handling competence beyond what is common for most biological/medical researchers. NBIS generalizes solutions built for specific research groups such as they can be re-used for new projects.
- 4. Critical mass** – NBIS provides a critical mass of bioinformaticians giving synergies from collective learning, which is hard to reach in an individual research group.
- 5. Expertise** – For research groups already having skilled bioinformaticians NBIS can, due to our breadth, provide complementary expertise.

Overview of NBIS activities

NBIS provides specialised competence having **experts** in multiple essential domains of large-scale analyses, but also has **generalists** able to integrate data from different omics areas. One of the strengths is our ability to assign **several experts from different domains** when required. NBIS keeps up-to-date with scientific developments, and for the period 2022–2026, we foresee a number of new directions:

- increased diversity in projects
- expansions in areas like single cell and spatially resolved transcriptomics
- increased collaboration with new SciLifeLab units, e.g. ancient DNA
- formation of a federated EGA landscape including efficient routes to access and analyse increasingly large amounts of genomics and imaging data
- supporting users with data management and FAIRification of data
- AI increasingly used in tools, affecting both the development process and the provision of the infrastructure; and increased needs of imaging informatics.

We also see new areas emerging, e.g. continued rapid development in single cell omics, advances in artificial intelligence (AI) techniques, and high-content imaging informatics. Furthermore, we see more projects combining imaging with genomics, proteomics and other omics techniques, and it is therefore logical that from 1 Jan 2021 the BioImage Informatics Facility (BIIF) and AIDA Data Hub are valuable additions to NBIS providing important competence.

NBIS will interact with the recently launched programme in DDLS (Data-Driven Life Science) funded by KAW 2021–2032 and coordinated by SciLifeLab. Of special importance, our efforts to help users to FAIRify data and provision of important data resources in our function as the Swedish ELIXIR node will be of value to enable the DDLS effort to be successful. NBIS will also provide support to the increased number of researchers in this area. In 2022, DDLS will establish Data Area Nodes (DANs) in the four areas of DDLS. For the users, it is important that the DAN activities are well coordinated with those of NBIS, and we will strive to achieve optimal coordination between NBIS and Data Centre in these matters.

NBIS will 2022–2026 continue our activities that were favourably reviewed by the Swedish Research Council (VR) recently: *“The NBIS approach is the only way forward to enable statisticians and informaticians to use optimal methods for data access and data analysis. These aspects are well documented in the proposal. The opportunity to support emerging new fields of application is of great interest.”*

NBIS provides Sweden with the necessary means to allow for simultaneous exploitation of vast amounts of biological data coming from different research fields and derived at different scales, from the molecule to the organism, and even to the population. It also allows for integration of these data with information from other disciplines, such as chemical, medical and environmental data. A national bioinformatics infrastructure **enables advanced user support at a level that single research groups (or even single institutions) cannot reach**. We provide specialised expertise in a number of areas, and our staff can simultaneously participate in multiple projects. NBIS thereby enables our users to benefit from the data-driven life science.

NBIS enables world-class life science research to **maximise scientific and societal impact** of publicly and privately funded research by providing expert knowledge, creative data integration, advanced training, efficient data publication and access to high-performance data analysis methods. NBIS coordinates bioinformatics support within Sweden and **makes bioinformatics easily accessible** for life science researchers.

Overall strategy

In order to fulfil our vision and mission, NBIS provides a number of activities, as detailed in the following. This includes provision of expertise and infrastructure to facilitate bioinformatics analyses and having access to necessary computational and storage resources. NBIS also provides routes for data publishing and engages in numerous advanced training activities.

NBIS coordinates its activities with other relevant infrastructures, such as other platforms at SciLifeLab, SNIC, SUNET and SND. NBIS main user groups are at academic institutions, but NBIS will also interface with the healthcare sector, governmental agencies, and commercial companies for mutual benefits. Internationally, NBIS constitutes the Swedish ELIXIR node and engages in Nordic and European collaborations.

Infrastructure

A fundamental part of NBIS is the formation of a sustainable bioinformatics infrastructure for life sciences, consisting of access to tools and data. The NBIS infrastructure is typically constructed as domain-specific supporting layers utilising resources from e-infrastructure providers such as SNIC and SUNET. These computational and storage needs are set up in close collaboration with SNIC/SUNET, in order to avoid duplication of efforts and to benefit from their competences.

Data management

As science in general, and life sciences in particular, is facing an unprecedented growth of data output due to technological advances in instrumentation, NBIS will strive to be positioned to be able to handle this data onslaught for Swedish life science. The following strategic Vision and Mission has been established to meet these demands.

Vision (Desired future position)

Swedish life science researchers apply good data management practices so that the research outputs produced are available to the global research community, and to society at large, according to the principles of Open Science, Reproducible Research, and FAIR.

Mission (Overall objectives and the approach to reach those objectives)

Support: Swedish life science researchers shall have access to data management support to be able to manage their research data according to international best practice.

- Develop **procedures** for providing efficient data management support to research projects throughout the research data lifecycle.
- Provide data management **planning** knowledge and resources/tools for research projects.
- Provide support and resources/tools for **data submission** to recommended community deposition databases.
- Provide support and resources/tools to enable **reproducible research**.

Training: Swedish Life Science researchers shall have access to training in efficient data management practices that can be applied in their daily work.

- Provide Data Management **training** for researchers.
- Provide **guidelines and documentation** for data management “best practises”.
- Promote awareness of the importance of Open Science, Reproducible Research, and FAIR in the Swedish life science community.

Collaboration: Data Management practices, solutions and training for Swedish Life Science shall be established in collaboration with NBIS internally, as well as with relevant national and international stakeholders

- Collaborate with **NBIS staff** to develop data management aspects of support procedures and policies.
- Contribute to Data Management activities and capacity building within **ELIXIR**.
- Collaborate with **SciLifeLab Data Centre** to develop data management policies, procedures and solutions to support researchers.
- Interact with **other national and international stakeholders** to harmonise data management policies, procedures and solutions to support researchers.

Work plan for 2022:

- Launch procedures for providing efficient data management support to NBIS analysis support projects.
- Collaborate with the SciLifeLab Data Centre to provide Data management support to prioritised SciLifeLab initiatives, such as DDLS, Pandemic preparedness, and Biodiversity.
- Define and start developing DDLS Data Platform services for Data management, e.g. for biological sample metadata management (together with SciLifeLab Data Centre).
- Launch the helpdesk function for data submissions to EGA-SE, to ensure that EGA-SE becomes operational and starts hosting datasets.
- Define and establish a submission brokering service to the European Nucleotide Archive (together with SciLifeLab Data Centre).
- Create a consolidated resource for SciLifeLab Data management guidelines and templates for Swedish life science researchers (together with SciLifeLab Data Centre).
- Continue to provide assistance and training to researchers in creating data management plans, including providing suitable templates and tools (together with SciLifeLab Data Centre).
- Hold the Introduction to Data Management Practices course at least twice. Extend the data management training portfolio in line with the training needs of the DDLS programme (together with SciLifeLab Data Centre).
- Continue to contribute to the EU ELIXIR-CONVERGE project, including to lead Work Package 1 and the ELIXIR Data Management Network.
- Follow and engage in the European developments in Data management, including developments in ELIXIR efforts in FAIRification of data, EOSC Association's work on semantic interoperability, and the NeIC Affiliate project Pandemic Research Infrastructure.

Systems Development & Tools

Access to professional and usable software tools is a cornerstone in current day research projects. This is particularly the case in fast-changing, data-intensive fields like next generation sequencing and proteomics. Easy-to-use tools that enable researchers to carry out more of the data analysis

themselves will also help to alleviate the demands for bioinformatics support in the form of consultancy.

While the development of new methods and algorithms is typically carried out in research projects, making these tools usable to the research community and keeping them maintained is often neglected due to lack of resources, expertise or incentives.

The NBIS systems development team creates user interfaces and provides support in deploying tools so that they can be used by the entire life science community, and not just by bioinformaticians. The combination of bioinformatics experts, who can appraise the scientific value and usefulness of tools, with developers who have the capability of making tools stable and accessible, allows for development projects driven by user needs.

The development efforts will entail creating user interfaces (e.g. web interfaces), providing assistance in programming best practices (documentation, source code management, bug tracking), and deployment. The development team will also assist in internal development projects, e.g. for integrating data services with other initiatives.

Bioinformatics has since long used machine learning for sequence pattern recognition and structure prediction. With recent advances in computational resources, these techniques have improved considerably and are summarised under terms such as Deep Learning and Artificial Intelligence (AI). NBIS will participate in national AI initiatives and benefit from the cross-disciplinary research on for example feature selection/extraction and pattern recognition, which are highly relevant for advancing NBIS capability on Big Data analysis. We will provide systems development support to users creating AI-based tools, and we will, together with our eInfrastructure partners, provide the computational requirements necessary for new AI-based tools. Further, NBIS has since long supported users in sharing data in a FAIR fashion, which is a prerequisite for data to be used in AI.

NBIS will continue to develop components in the HPA ([Human Protein Atlas](#)) which now is an ELIXIR core data resource. We will also make the [Metabolic Atlas](#) (genome-scale metabolic models) integrated with other ELIXIR resources and hopefully also an ELIXIR core data resource in the future. Already in 2019, [curated metabolic maps were integrated in HPA](#).

Prioritisation of new systems development projects proposed by staff or the user community are done regularly (typically every 2–3 months) by the management, according to the general principles:

- Highest prioritised are international agreements and national services of large interest.
- Next level are projects of large impact for Swedish users and of central importance for NBIS functions.
- Medium-prioritised are individual support projects for systems development that are treated as other support cases, including charging user fees.

Methods and software developed within NBIS are made publicly available through Open Access publication and Open Source licensing of software. NBIS has developed and published [coding guidelines](#) to aid in developing better software, and our staff also participates in the ELIXIR task Software [best practices](#). We maintain **public repositories** (GitHub) for codes and scripts that are developed by NBIS.

Computational resources to our users will predominantly be provided by SNIC. For infrastructure services, SUNET will be used for EGA-SE encryption, and EGI (European Grid Infrastructure)

will be used for prediction servers. In the future, we expect that emerging computational services within ELIXIR and EOSC (European Open Science Cloud) to be utilised.

Within the Systems Biology unit, NBIS will implement tools for analysis of high-throughput data and visualisation of results. Based on functionality provided by the Metabolic Atlas framework we can currently support researchers on omics data integration. Ongoing academic research also indicate that genome-scale metabolic models can be converted to “knowledge enriched” neural networks. This opens up opportunities for future development of Metabolic Atlas into a research infrastructure service for data-driven precision medicine.

Starting 2021, NBIS is engaged in the 6-year EU IMI project BigPicture aiming at creation of an infrastructure for digital pathology. NBIS contributes with expertise in systems development and with bioimage informatics via AIDA Data Hub.

NBIS continuously enters relevant Swedish tools in the ELIXIR Tools Registry.

Work plan for 2022:

- Maintain prediction servers and bioinformatics tools of importance for the Swedish life science community, of which several are expected to also be of interest for ELIXIR.
- Engage in the development of tools for genome analyses.
- Develop Nextflow workflows for data processing and analysis together with NGI.
- The management makes regular prioritisations on new systems development projects.
- Enter relevant Swedish tools in the ELIXIR Tools Registry.
- Participate in European collaborations on systems development to enable efficient bioinformatics tools for large-scale analyses and management of data – both human and non-human.
- Continue with our engagement in 1+MG and B1MG.
- Engage in the recently launched EU IMI project BigPicture.

Compute and Storage

High throughput biomedical science nowadays depends on high-performance computers for bioinformatics analysis. The hardware is maintained by SNIC but NBIS provides expertise needed to efficiently access the computational and storage resources. We expect a continuing high number of projects as well as increasing sizes due to decreasing costs for sequencing and the increased capacity of NGI. In order to preserve the frontline position of Swedish bioscience, we need to stay in the forefront of e-Infrastructure resources and related expertise. With growing numbers of project, increasing sample and data sizes, we will develop plans for how to accommodate for this scenario in the best way. We will explore the benefits of cloud computing and emerging data analytics methods in the realm of AI and machine learning. We will also investigate how scientists can improve the automation of bioinformatics analysis using scientific workflow tools. In addition, we will improve the support available regarding data management in relevant contexts.

From 2023, the advanced user support will not be provided by the new SNIC in accordance with the guidelines from VR. In order to keep the well-functioning user support for NGS analyses at Bianca and Rackham that the Compute & Storage unit currently provides with NBIS staff together with SNIC staff, NBIS will from 1 Jan 2023 finance the current SNIC staff providing this user

NBIS Strategic Plan 2022–2026 including Work Plan 2022

support at the Compute & Storage unit, i.e. 2 FTE at UPPMAX. NBIS will also collaborate with the other user support areas at UPPMAX and the new SNIC.

Support for users is currently focused on the environment present at UPPMAX. By leveraging new software deployment methods, we will make this environment more portable and vastly improve our ability to support users also on other SNIC systems and elsewhere.

The largest compute cluster provided by UPPMAX (Rackham) and the sensitive data system (Bianca) reach end-of-life during this period. We expect that at least 2000 users and their 12+ PB of data will need administrative and/or technical help moving to replacement systems.

Increased support for automatic workflows and digital research environments. Workflows in e.g. Galaxy are growing in power and popularity, and are quickly becoming an important tool also for experts and non-experts alike.

The Wallenberg Artificial Intelligence, Autonomous Systems and Software Program (WASP) has invested in two large-scale computer systems for artificial intelligence/machine learning (AI/ML) methods. We will support our community with AI/ML needs on these resources.

Furthermore, the recently launched KAW-funded initiative in Data-Driven Life Science (DDLS) will increase the demands on computational capacities on computational and storage resources, especially for sensitive data handling. NBIS will adapt to these increased demands, and will provide necessary services in collaboration with SNIC and UPPMAX. NBIS will also provide access to databases needed by the increased number of researchers in DDLS.

It is expected that long-term storage systems of various kinds will continue to spring up locally and nationally to meet the need for archival and publication of research data. We will increase our level of support towards the adoption of good practices with regards to long-term data storage and publication according to FAIR principles and legal requirements.

Work plan for 2022:

- Continue providing user support within compute and storage.
- Follow usage patterns and facilitate efficient utilisation of the computational and storage resources.
- Coordinate with NGI.
- Follow the development in the compute and storage area and make suitable pilot studies.

Interactions with other research infrastructures

NBIS is increasing the interactions with other SciLifeLab platforms as the needs of users' of other platforms for bioinformatics support becomes more pronounced. NBIS already has **regular meetings** with NGI (National Genomics Infrastructure) and Clinical Genomics. NGI also participates in the weekly virtual drop-in sessions that NBIS host where users can ask bioinformatics or sequencing related questions. We have **joint staff** with CBCS (Chemical Biology Consortium Sweden) in chemoinformatics, and **embedded NBIS staff** in other SciLifeLab platforms, as NBIS is running a pilot within proteomics facilities and providing image- and data

analysis support for the Single Cell and Spatial Biology Platform. Further details are provided below.

During 2022, NBIS will continue to assign dedicated persons to be primary contacts to the SciLifeLab platforms/facilities.

NBIS has regular contacts with SNIC and its centres in order to get optimal provision of computational and storage resources. Also coordination with the application experts in bioinformatics will be important.

NBIS, in close collaboration with data-producing infrastructures, e.g. Clinical Genomics, will provide expertise in systems development and access to ELIXIR-related databases and tools enabling improved and cost-efficient health care. We collaborate with Genomic Medicine Sweden (GMS), mainly on data and knowledge discoverability and sharing, in GA4GH Beacon and federated human genomic data technologies. We also coordinate our efforts in 1+MG, where we both are active. Furthermore, NBIS participates in an GMS-led Vinnova project to facilitate for patients to give their consent for secondary use of genomics data.

On the data publishing side, NBIS has ongoing collaborations with the SciLifeLab Data Centre, SUNET and SND, Swedish National Data Service. Furthermore, NBIS follows the progress with MAX-IV and ESS for potential new user communities.

On the international side, NBIS follows the work in EOSC-Life, coordinated by ELIXIR, and in which all European biomedical infrastructures are partners. UU as host university of NBIS is one of the founders of EOSC. NBIS is also actively engaged in RDA (Research Data Alliance). Furthermore, NBIS follows the work in GA4GH (Global Alliance for Genome and Health).

NBIS has several collaborations ongoing in the field of biodiversity. We will continue our engagement in ELIXIR Biodiversity Focus Group. Work in the group is currently focusing on efforts to increase usability of biodiversity data by enabling correct data management. Together with NGI, NBIS is engaged in a VR funded programme for increased interaction with Swedish governmental agencies in the area of biodiversity. This project also serves as a pilot for a Swedish Earth Biogenome Project (EBP), with genomes produced to EBP standards. A European level EBP is also being formed by the European Reference Genome Atlas (ERGA), and NBIS is involved in the ERGA assembly committee and as well as helping out in the collaboration of ELIXIR and ERGA. An application to fund ERGA was granted by Horizon Europe early 2022.

NBIS will during 2022 intensify our collaboration with SBDI – Swedish Biodiversity Data Infrastructure – for the benefit of all users in these scientific areas.

Work plan for 2022:

- Regular meetings with SNIC and SUNET to coordinate computational and storage and issues.
- Assign dedicated persons to be primary contacts to SciLifeLab platforms/facilities.
- Regular meetings with the SciLifeLab platforms Genomics, Clinical Genomics, Clinical Proteomics, and Single Cell and Spatial Biology.
- When motivated, coordinate NBIS activities with relevant platforms at SciLifeLab and relevant national infrastructures.
- Participate in relevant European and international infrastructure meetings.
- Engage in the ELIXIR Biodiversity Focus Group.
- Participate in the EU project and develop assembly standards and methods to be used in ERGA.

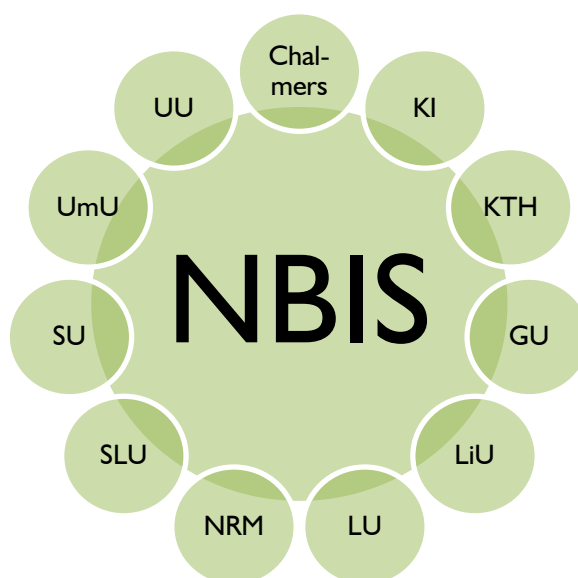
Local interactions

In order to optimise coordination between the national infrastructure NBIS and local bioinformatics activities, e.g. core facilities, NBIS has assigned site coordinators to guide outreach efforts. Bioinformatics is entering more and more research projects, and we expect local support facilities to be more involved in support needs.

Currently NBIS in Gothenburg participates in GOTBIN, a gathering of bioinformaticians and bioinformatics in Gothenburg. In Umeå work has begun to start a local bioinformatics network in collaboration with UmU. There are already events organised or co-organised by NBIS there. NBIS will expand the effort to have more site coordinators and strengthen contact with local bioinformaticians and life scientists.

The **NBIS site coordinators**, typically 20 %, are responsible for keeping contacts between the national activities of NBIS and local activities. This is important for knowledge transfer, where new techniques taken up by NBIS should be spread also to the local sites. Similarly, NBIS will faster become aware of needs from local facilities for national support. They will also serve as an important channel for effective information distribution, to facilitate access to bioinformatics for all life scientists.

Since bioinformatics is entering more and more research projects, we expect that local support facilities will successively take a larger part of the total support needs. We also expect a trend where common and standardised analyses are done at multiple local sites, where new state-of-the-art techniques are provided by NBIS. Furthermore, the local facilities will be tailored to meet local needs that are better catered for by a local facility than NBIS which is providing the national perspective. All NBIS site coordinators will be under the responsibility of the community coordinator that is part of the management group.



BioImage Informatics, a part of NBIS, is already in close contact with all major Swedish Microscopy Facilities.

Work plan for 2022:

- Our community coordinator will establish contacts between NBIS Management and Site coordinators and to administer NBIS affiliates.
- Continued and increased engagement in, or startup of, local bioinformatics groups.

Support

NBIS supports excellence in research. One of our major activities is support, where NBIS staff helps researchers with bioinformatics tasks in various projects. The time spent in each project varies from short (weeks) to long (months). The topics for NBIS experts are decided by the NBIS Board and Management, following suggestions from open NBIS calls, NBIS partners, the International Advisory board, the Reference Group, evaluations, or the Board itself. In order to be flexible and to test the needs for new topics, the Board can decide upon launching short-term (1–2 years) project-type activities. This will enable NBIS to provide expertise in the areas needed by the life science researchers.

Currently, NBIS provides expertise in many areas within bioinformatics: genome assembly, genome annotation, genetic variation, comparative genomics, phylogenomics, transcriptomics, proteomics, metabolomics, systems biology, single-cell biology, biostatistics, image analysis and multi-omics integration.

NBIS has users from **all Swedish major universities** and predominantly from the faculties of medicine, science, technology, and pharmacology. NBIS handles all support projects at the **national level** and assigns the **best expert** for each project, regardless of geographic location. Furthermore, NBIS has the possibility to assign **multiple experts, when needed** for providing expertise in multiple areas or for provision of long-term redundancy.

Support constitutes about half of NBIS activities, and our users are from all Swedish major universities. NBIS provides bioinformatics support in the form of different services, ranging in commitment from short meetings to extensive collaborations. On the lighter end of the spectrum, we arrange **bioinformatics drop-in sessions** at all sites in Sweden, providing guidance on experimental design, choice of analysis methods, software etc. They also give the opportunity to learn more about NBIS services and how to apply for these. In conjunction with large grant applications, we arrange extra drop-in sessions with focus on these applications. We will arrange **~250 drop-in sessions** annually, which are much appreciated by the community. During the corona pandemics, these have been held online. In the future, we will enable online connections at the drop-in sessions in order to facilitate participation by experts at other sites. We also offer **bioinformatics consultations**, where we discuss research projects with users. This service is provided for free, but the NBIS staff do not perform any work on the users' data. Many consultations result in a support project later on.

We will continue the major activity of **bioinformatics hands-on support service**, where NBIS staff work actively in the projects for shorter or longer time, allowing research groups access to

cutting-edge expertise that would otherwise be difficult to obtain. We provide three hands-on support tracks.

In the **Short- and Medium-term Support track**, the focus is on short and medium-sized projects (typically 40–500 h) under a user-fee based model (currently 800 SEK/h = 80 EUR/h). In the last few years, we have seen an increase in the time requested, reflecting more complex data and often of different types. Projects are accepted continuously with the aspiration of having short waiting times, with most analyses starting within a few weeks from signed contract. We make a technical evaluation of each project to assure its feasibility and that we have the specific competence within NBIS.

The **Partner Project track** is intended for projects with a large bioinformatics component, where NBIS can enter as a project partner based on cost coverage by the research project. This track is intended for projects requiring NBIS support of at least 12 person months over 2–5 years (e.g. 0.5 FTE over 2 years). Since 2019, the number of partner projects has increased, and we foresee further increases during the next years.

The **Long-term Support (WABI) track** provides extensive support to a limited number of scientifically outstanding projects that involve very large data sets and/or require extended, creative and customised analyses to accurately answer the scientific questions. The primary funding is from Knut and Alice Wallenberg Foundation (KAW) and no user fee is charged. Supported projects are selected in a rigorous scientific peer-review process in open national calls three times annually. Knowledge transfer is a key aspect of the support model, and dedicated researchers working hands-on alongside the support staff is required. The DDLs programme includes additional funding to the Long-term Support track from 2022 onwards, with an instruction to broaden the range of expertise to include support for Cryo-EM projects.

In order to facilitate contacts with NBIS experts, we have staff at all major university sites. They constitute easy-to-find local entry points into the NBIS infrastructure. These local contact points will have good knowledge about the NBIS staff at other sites. NBIS staff has a national responsibility and should serve the users' needs regardless of their affiliation.

The BioImage Informatics unit provides support in a fashion very similar to the general NBIS support, but also provides monthly digital 'call for help' session with support on image processing and analysis, including advice on all steps from optimisation of data collection to presentation of results. We focus on free and open-source tools, and always aim to train the users to do at least part of the analyses themselves. The need for support on image and data analysis connected to the SciLifeLab Single Cell and Spatial Biology platform is quickly increasing, especially when it comes to processing and visualising spatial omics data, and connecting it to morphological information, e.g. extracted by deep convolutional neural networks. We are currently discussing how to best create an efficient data flow between the platforms, also including the AIDA Data Hub.

Prioritisation

As a national infrastructure NBIS aims at providing bioinformatics support in all projects where our competence is asked for. However, when the available resources are not sufficient to match the total needs, a prioritisation has to be made according to the principles listed below.

Short- and Medium-term Support accepts projects continuously and prioritises according to the principles listed below. New support projects are assigned to the appropriate staff member(s) by a

project coordinator. Managers and project coordinators follow up that each project proceeds according to plan.

Long-term support has application rounds 3 times per year, where projects are scientifically ranked by the Proposals Evaluation Committee, an independent committee of scientists from Swedish universities. The managers prioritise the projects according to technical feasibility in agreement with this ranking.

Below are the current prioritisation principles, as decided by the NBIS Board and supported by the International Advisory Board. The prioritisation is done by the NBIS managers, based upon information from the NBIS staff.

- Technical feasibility
- Availability of data
- Projects which are judged excellent by VR or our external prioritisation committee are prioritised.
- Projects where the NBIS staff has appropriate competence are prioritised.
- Projects where the NBIS efforts make a large impact are prioritised.

In order to more efficiently be able to help more users, NBIS will provide guidance so that the users become able to do part of the bioinformatics analyses on their own. In line with this, NBIS staff also devotes part of their time to training activities. Furthermore, NBIS maintains a useful infrastructure, including tools and data handling, available for the users (cf. Infrastructure, above).

Work plan for 2022:

- Continue providing support according to the guidelines above.
- Increase staff in order to match increased needs
- Recruit for the DDLS-funded expansion of Long-term Support, including Cryo-EM experts.
- For 2022, the Board has set the academic user fee for NBIS support to 800 SEK per hour.
- For 2022, the Board has set the Partner Project user fee to 44 kSEK/month+social fees. The OH surcharge is 35%.
- Encourage researchers to include bioinformatics costs in project grant proposals. NBIS will assist in estimating these costs.
- NBIS continuously follows trends in life science research in order to be prepared for emerging technologies and new bioinformatics approaches.
- Establish a close collaboration with the Single Cell and Spatial Biology platform, and provide a clear dataflow between the platforms.

Training

In the ongoing transformation of biology and medicine into large-scale data driven research, **advanced training is a key factor to ensure Sweden's scientific competitiveness**. NBIS has the mission to provide advanced bioinformatics courses and NBIS staff is involved in a wide range of training activities targeted towards the Swedish Life Science community. Here, NBIS offers a course catalogue with ~20 courses in advanced bioinformatics and data science topics as well as collaborative efforts with other European institutes in state-of-the-art methodology and research

topics. In addition, NBIS offers commissioned training with approximately 3 training initiatives per year as part of the Bioinformatics Support. NBIS experts also are regularly invited as teachers/speakers in events arranged by others.

Part of the bioinformatics support, offered by NBIS in the various projects we undertake, is the individual training of researchers in order to teach them new bioinformatics tools and to help them utilise these tools more efficiently in their own projects. The training activities are also an efficient way to increase the flow of projects through the NBIS organisation by helping scientists to be able to perform parts of the bioinformatics analyses themselves.

There is an increasing need of training, both at an introductory level and at an advanced level, and this is seen throughout the international arena. NBIS also sponsors and co-organises different workshops. NBIS training events are announced via NBIS website, SciLifeLab website and the ELIXIR TeSS website in addition to announcement via NBIS Twitter and other bioinformatics forums. The aim is for the courses to be of a national spread, facilitating users at different universities to take our courses.

In the trace of the pandemic we have switched the training delivered to an online format and this is something we will partly continue with after the end of COVID-19 as delivery of courses in an online format is in line with the openness and FAIRness of training material NBIS intends to uphold as one of its mission. We are continuously evaluating the demand for training activities, and we are making sure to keep our content and topics state-of-the-art. NBIS is encouraging our experts to professionally develop in pedagogics and cognitive science. Currently, the demand for NBIS courses is larger than our resources permit.

NBIS has since 2015 run the highly appreciated and successful **Bioinformatics Advisory Mentorship** Programme, a mentor programme where PhD students (typically 15–20 new per year) get a senior NBIS expert as an advisor for up to two years of their PhD studies. Since NBIS as a platform in 2021 welcomed BioImage Informatics (BIIF), which is closely involved in the Network European Bioimage Analysis, [NEUBIAS](#), the NBIS course catalogue grew with training schools related to the Bioimaging topic. BIIF is involved in the recently launched on-line NEUBIAS academy.

NBIS engages in the **ELIXIR Training Programme** for delivery of courses in the life science sector such as bioinformatics, biocuration, data management and data science in addition to being part of the setting of training infrastructure across ELIXIR nodes. Since 2021, ELIXIR-SE has the role as one out of three Executive Committee members for the ELIXIR Training platform and hence has the role to strategically be part of developing training for the member nodes. Since the last years, ELIXIR training has successfully increased the training engagement across all nodes and with this been able to successfully integrate the infrastructure and outcomes from the joint node efforts into the NBIS training work.

Work plan for 2022:

- During 2022, we will arrange several training events on at least the following topics:
 - next generation sequencing data analysis, e.g. bulk RNA and single cell analysis
 - statistics and machine learning
 - epigenomics analysis
 - programming courses, e.g. in Python and R
 - Single-cell transcriptomics collaborative course with SIB (Switzerland)
 - Advanced Methods in Bioimage Analysis
- During 2022, we will in the areas where we see increased demand, set-up new or modify existing training events to meet the needs of our users e.g. neural networks and deep learning, data management
- Deliver NBIS training using the Learning Management System CANVAS/Studium (hosted by Uppsala University)
- Continue with our National Advisory Mentor Programme in Bioinformatics.
- Continue our engagement in the National Research School in Medical Bioinformatics.
- Provision of information on bioinformatics courses available in Sweden and within ELIXIR.
- All NBIS material and courses are Open and FAIR.
- Continue to collaborate with other SciLifeLab platforms when relevant for advanced courses and seminar series.
- Continue to collaborate with other ELIXIR nodes for advanced courses and training events.

Information and Outreach

NBIS has a web site (<https://nbis.se>) for providing up-to-date information both to our users and for internal purposes. It will be thoroughly updated in 2022 to deliver a more straightforward experience and a clearer visual design. NBIS also has a project management system facilitating tracking of projects and allowing for NBIS staff to easily share data and information with their customers.

Outreach activities have proven important to inform the scientific community about the support that NBIS can provide, making bioinformatics easily accessible for life science researchers. These consist of:

- Annual symposium and user meeting.
- Presentations at different universities, providing the possibility to meet NBIS staff representing our wide variety of competences.
- Additional presentations at various symposia and conferences.
- Weekly Bioinformatics Drop-in sessions enabling face-to-face contacts between researchers and NBIS experts, which many times is the first contact in a support case. These sessions are arranged online and, when there is no pandemic, at all sites (Uppsala, Stockholm, Göteborg, Linköping, Lund and Umeå).
- Monthly digital call-for-help sessions in bioimage informatics / image analysis and newsletter
- Joint thematic outreach events such as the NBIS/NGI VR Outreach, letting life scientists know that we can help improve their grant applications for VR funding.

An Outreach group has been created within NBIS that will create, update and maintain outreach materials for presentations, graphical assets and informational content. The outreach group will strive to create meaningful and effective channels of information distribution which should only reach recipients that might be interested in our services or community. They are also working on a coherent graphical profile for NBIS. NBIS has Twitter and LinkedIn accounts for additional interactions and information dispersal and are working with SciLifeLab communications office to aid in getting relevant information to the users.

NBIS aims to increase the collaboration with other platforms within SciLifeLab. In 2018, we launched the Seminar series *BiG Talks!* as a Bioinformatics and Genomics seminar series arranged in collaboration between NBIS, NGI and Clinical Genomics. The seminar series is financially supported by SciLifeLab and talks are given by internationally renowned speakers within the fields of Bioinformatics and Genomics. The seminar series aims to increase the interaction between bioinformaticians and researchers within the SciLifeLab community. The BiG Talks are broadcasted in order to reach a wider audience with possibilities for participants to ask questions and comment in real time. The idea is that the seminar series will rotate to different SciLifeLab sites during the four times per year these take place.

In 2021, NBIS community coordinator joined a task force to set up a new event, the “SciLifeLab Infrastructure Outreach Week”. The event hosts one topic per day over a week and presents case studies where SciLifeLab infrastructure platforms have been utilised. The event is planned to be recurring.

Work plan for 2022:

- Continue to keep the NBIS web site up-to-date and informative – both for our users and for our staff.
- Weekly (virtual) Drop-in sessions at the sites where NBIS personnel are located.
- Annual symposium and user meeting will be held in 2022.
- Continue the *BiG Talks!* seminar series during 2022.
- Increase collected feedback from all users of NBIS services to enable future improvements.
- Create efficient site-side information distribution channels.
- Increase presence in local bioinformatic communities.
- Continuous updates of information material, both web and print.

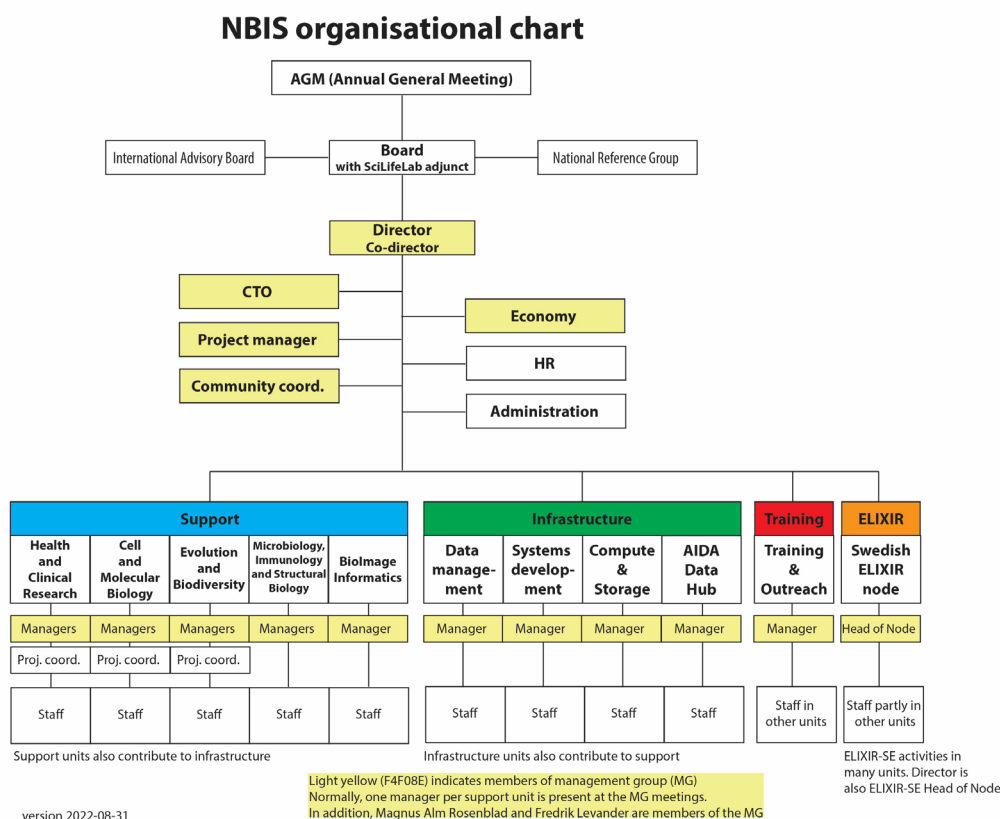
Collaboration with industry

NBIS will continue our efforts to increase contacts with life science companies to enable them access to NBIS-provided tools, databases, expertise, and training.

Work plan for 2022:

- Provide support to companies at a full-cost rate when we have appropriate expertise and capacity.
- Invite scientists at companies to NBIS training activities, to the annual symposium and user meeting, and to NBIS site visits.

Organisation



The organisational governance is regulated in the NBIS consortium agreement, updated in 2019. The organisational scheme is depicted above.

The Board is nominated by the Annual General Meeting and assigned by Uppsala University after consultation with VR and SciLifeLab. The Board meets 3–4 times per year, often via teleconference. The SciLifeLab infrastructure director and Reference group chair participate in the Board meetings.

For the coordination and leading the daily management of the infrastructure, a director is appointed by the Board. This is a 50% position, typically a professor in bioinformatics or other relevant field. The director has the executive responsibility for running the infrastructure, delegated from the Board. As the NBIS activities has grown, NBIS has since 2020 also a co-director in order to facilitate the running of the infrastructure. In addition, a technical manager (CTO; 100%) is appointed leading and coordinating the technical management of the infrastructure.

The activities of NBIS are divided into organisational units, as depicted in the figure above. The managers together with the director form the management team that meet using video conference system at least monthly to coordinate activities and discuss operational matters. Important questions of policy, strategy and economy are put forward to the Board for decision. When time-wise suitable these meetings will be physical.

Electronic Meetings and Retreats

In a distributed infrastructure like NBIS, it is important that all staff members are aware of the special competences of their colleagues at other sites. In order to achieve this, we use a text-based chat system (Slack) where we have set up many channels for discussing specific topics. Here, questions can be asked at any time and colleagues monitor their channels of interest as time allows. We also have a weekly short video meeting to inform about current NBIS activities of relevance to many in the organisation. The weekly meetings also provide opportunities for staff members to ask general questions, exchange ideas and socialise. In addition, we have annual retreats for all staff to increase interactions and give ample time for long-term planning and strategy discussions, bringing up new ideas, and develop the activities.

Topical meetings

NBIS organises internal topical meetings, focused on a particular bioinformatics sub-discipline, *e. g.* Next Generation Sequencing, proteomics, large-scale data management, training. The purpose of these meetings is to facilitate internal networking and knowledge transfer. The topical meetings will allow for more in-depth discussions on new papers or methods, and of current NBIS projects. In order to minimise travelling time, the topical meetings will predominately be held on-line. At these meetings, relevant NBIS-affiliated persons are invited.

Professional development of staff

In order to assure a continuous competence development of the NBIS staff, they should ideally have their basis in a bioinformatics research environment, giving them opportunities to keep up with progress at the research frontiers and attending lectures and seminars. Furthermore, they should be given time for own education and development, *e.g.* when involving in support tasks needing additional competence. The time available for competence development is up to 20% (including the above mentioned own education associated with support tasks). As bioinformatics is a rapidly evolving discipline and new areas emerge, over time NBIS staff might move between different areas, depending on user needs and their own interests. NBIS is engaged in creating career development for staff scientists.

Work plan for 2022:

- Continue with chat-based knowledge exchange and weekly meetings.
- Continue and further develop topical meetings.
- Organise at least one internal course for competence building within NBIS.
- Arrange two retreats – digital or physical.
- Participate in SciLifeLab working group on career development for staff scientists.

Affiliated NBIS persons

In order to increase the national bioinformatics networking, NBIS enables affiliation of bioinformaticians at other SciLifeLab platforms and at core facilities. This will give them access to NBIS topical meetings, provide valuable contacts with NBIS staff performing similar work and thereby increase their network. NBIS affiliations are decided by the NBIS management.

Work plan for 2022:

- Continue with affiliation of relevant bioinformaticians to NBIS.
- Invite affiliated NBIS persons to knowledge exchange meetings.

Funding

The major funding sources of national NBIS activities are the Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council (VR). NBIS also has financial support from the participating universities. In addition, users are contributing with user fees. These combined funding streams enable NBIS to grow successively as the demands for bioinformatics support increase. When suitable, NBIS will participate in national and international grant applications, predominantly in the infrastructure area.

Work plan for 2022:

- Increase user fee contributions.
- Continue expansion to meet the increased demands.
- Participate in relevant grant applications.

International

ELIXIR node

ELIXIR is the European infrastructure for biological information with currently 22 countries and more to join next years. Sweden is one of the founding members, and NBIS constitutes the Swedish node. Sweden is very active, e.g. with Human Protein Atlas (HPA), now an ELIXIR Core Data Resource, reflecting its fundamental importance to the life-science community and long-term data preservation; with the EU projects ELIXIR-CONVERGE 2020–2023, where we lead European efforts in data management; and B1MG 2020–2023, where we provide coordination support and outline the infrastructure for the European 1+ Million Genomes Initiative; and with multiple ELIXIR training efforts.



In the **European 1+ Million Genomes initiative (1+MG)**, 20+ countries will provide a cross-border federated network of national genome collections associated with relevant data for advancing data-driven health. NBIS is active in building and establishing components of the technical framework, e.g. the Federated EGA for secure storage of sensitive genome data, and the GA4GH Beacon for discoverability.

Since 2013, Sweden contributes with the Human Protein Atlas (HPA) and its integration into the ELIXIR landscape. HPA was declared ELIXIR Core Data Resource in 2018. Efforts will be made to get Metabolic Atlas to become an ELIXIR Core Data Resource in the future. Sweden will engage in multiple ELIXIR-related activities, including the federated EGA, the ELIXIR-CONVERGE project, RDConnect, and the 1+MG project.

To assure optimal coordination between ELIXIR user communities and Swedish researchers, NBIS will assign community liaisons for relevant user communities. This is already initiated for the ELIXIR communities Proteomics, Metabolomics, and Structure. Additional communities will be formed.

Nordic collaboration

NBIS aims at strengthening the Nordic collaborations with Norway, Denmark and Finland on computing, storage, training and on ELIXIR node activities. Since 2011, we have regular meetings between the Nordic ELIXIR heads of nodes. Travel costs for these meetings are kindly supported by a NordForsk grant. Since 2013, NeIC (Nordic eInfrastructure Collaboration) engages in supporting biomedical sciences, initiated by letters of interest from NBIS together with the Nordic ELIXIR nodes. The project Tryggve aims at constructing a federated solution across the Nordic

countries for secure data handling and analysis, allowing for exchange of data when the ethical permits so allows. The first and second phases were very successful, and a third phase NeIC Heilsa (Tryggve3) is granted 2021–2023.

Work plan for 2022:

- Continue the work on integration of Human Protein Atlas into the ELIXIR landscape and provide seamless integration with other important data sources.
- Contribute to ELIXIR 2019–2023 programme.
- Participate in ELIXIR-CONVERGE project activities according to the plans in the application.
- Participate in relevant ELIXIR implementation studies and other ELIXIR-related activities.
- Continue our work on Federated EGA.
- Engage in ELIXIR-related human data activities, including 1+MG, RDCConnect,
- Establish contacts between new ELIXIR user communities and relevant groups in Sweden
- Collaborate in relevant areas with the Nordic ELIXIR nodes.
- Plan for joint advanced training activities with the Nordic ELIXIR nodes.
- Continue our activities in Tryggve2.
- Engage and collaborate with NeIC concerning their activities on the biomedical arena.
- Engage in relevant international initiatives (e.g. EOSC, GA4GH, CINECA, RDA) that can benefit Swedish life science.

Appendix I – Abbreviations

1+MG – European 1+ Million Genome Initiative
CINECA – Common Infrastructure for National Cohorts in Europe, Canada and Africa (an international project led by EBI)
DDLS – Data-Driven Life Sciences
EBI – European Bioinformatics Institute
ECDS – Environment and Climate Data Sweden
EGA – European Genome-phenome Archive
ELIXIR – European Infrastructure for Biological Information
EMBL – European Molecular Biology Laboratory
ESS – European Spallation Source
FEGA – Federated European Genome-phenome Archive
FTE – Full time equivalent
GA4GH – Global Alliance for Genome and Health
GU – Göteborgs Universitet, University of Gothenburg
HPA – Human Protein Atlas
IAB – International Advisory Board
KI – Karolinska Institutet, Stockholm
KTH – Kungliga Tekniska Högskolan, Royal Institute of Technology, Stockholm
LiU – Linköping University
LU – Lund University
NBIS – National Bioinformatics Infrastructure Sweden
NeIC – Nordic eScience Infrastructure Collaboration
NGI – National Genomics Infrastructure
NGS – Next Generation Sequencing
NRM – Naturhistoriska Riksmuseet, Swedish Museum of Natural History, Stockholm
NSC – National Supercomputer Centre at Linköping University
PDC – PDC Centre for High Performance Computing at KTH
PI – Primary investigator
RDA – Research Data Alliance
SciLifeLab – Science for Life Laboratory
SLU – Sveriges Lantbruksuniversitet, Swedish University for Agricultural Sciences
SND – Svensk Nationell Datatjänst
SNIC – Swedish National Infrastructure for Computing
SU – Stockholm University
SUNET – Swedish University Network
UmU – Umeå University
UPPMAX – Uppsala Multidisciplinary Center for Advanced Computational Science
UU – Uppsala University
VR – Vetenskapsrådet, Swedish Research Council
WABI – Wallenberg Advanced Bioinformatics Infrastructure