2020

Strategic paper “Advanced Methods, Data and Analyses to understand Living Systems”

Version taskforce December 2019

Strategic paper NWO research community

“Advanced Methods, Data and Analyses to understand Living Systems”

Table of contents

1. Scope of the life science research community “Advanced Methods, Data and Analyses to understand Living Systems” 2
2. Current scientific strengths/themes/challenges and (inter-)national position 2
3. Focus areas for the next ten years 3
4. Impact on society and contribution to topsectors (and related missions), national science agenda-routes and SDG’s 3
5. Infrastructure in national & international perspective 4
6. Involvement of Dutch organisations in your research community 5
7. Specific challenges for the community 6

The overarching theme “Living Systems: Methods, Data and Analyses” focuses on the development and innovation of research methodology and tools for life science in the broadest sense. Modern life science has to explore ways for state-of-the-art methods that expand our possibilities to investigate life and its ecosystems. This theme brings together scientists at the forefront of these technologies and advances data and modelling expertise. With the interconnection of established methods and the new prospects, these techniques impact all levels of life science. This working group represents the research community that drives the developments of Methods, Data and Analyses in biology across life science fields.

1. Scope of the life science research community “Advanced Methods, Data and Analyses to understand Living Systems”

- Molecular methods, technologies and instrumentation to contextually study living systems at different scales and understand the dynamic complexity of the genotype-phenotype connection.
- Data, quantitative modelling and other digital innovations in life science through novel methodology, alignment of existing methodology (standardization), capacity building and training.
- Societal innovations from biology through collaboration among and across disciplines that connect fundamental research outcomes and multilevel data integration, from molecule to human, to urgent societal health and environmental questions.

2. Current scientific strengths/themes/challenges and (inter-)national position

Research in advanced methods, data and analysis of living systems in the Netherlands is traditionally strong in the development of approaches to study the dynamical contextual complexity of life. This interdisciplinary community brings together researchers with backgrounds in physics, chemistry, mathematics and informatics required to advance biology. Internationally this Dutch research community is well known and recognized for the innovations in molecular imaging technologies such as EM, MRI, Optical and correlative microscopy and recently mass spectrometry based imaging. In particular, single-molecule and single-cell molecular microscopy are hot-topics in the method development and instrumentation innovation. Innovative technologies to study larger biological ecosystems and their effect on society, transport, food and biodiversity are conducted at an internationally competitive level. The improvement of soil quality through the development of a better understanding of the rhizosphere biology has huge impact on crop sciences.

Research opportunities are created through access to large-scale research infrastructures and data, that, in addition to imaging and “omics” platforms (from molecule to earth observation), offer unique opportunities to conduct life science research across the boundaries of disciplines. Combined, they allow a constant development of new innovative technology and methodology to answer complex questions in the life sciences. Analytical challenges exist in resolution (temporal, spatial and molecular), quantitation, throughput, data reduction and integration. Currently a huge opportunity presents itself on the integration of cell and systems biology with detailed molecular understanding. At the same time this is a challenge, not in the last part as a result of increasingly large and complex data sets that are being generated by the different innovative analytical methods.

Our community can boast strong computational development groups that apply in-silico methods at all scales from simulations to data analytics of DNA/proteins, metabolomics, cells and cell-cell interactions, tissue, organ, and organism scale. Alignment of our research endeavours targets the big challenges of the 21st century (healthy living and ageing, the energy transition, climate change, urbanization, etc.) and is well positioned to significantly contribute to major scientific challenges (e.g. origins of life and sustainable planet). In that light, the Dutch bioinformatics and systems biology field has an internationally strong reputation and is recognized as ahead of the pack in FAIR-based data stewardship that is well connected to the major European data infrastructures.

This research community represents a very active research field as an established driver and enabler to advance the life sciences. A continuous investment in instrumentation development for the life science is pivotal to maintain the leading position of Dutch life science research.

3. Focus areas for the next ten years

The current level of independent technology platforms is very high. To be able to answer the pressing questions of the future across the breath of life sciences, necessitates integration of available technologies in combination with ‘scale-crossing’ and ‘discipline-crossing’ innovations. This will enable life science research over a variety of spatial and temporal scales in increasingly complex systems.

The next decade will show a focus on the systematic integration across the entire analytical workflow in the life sciences, including sample handling, methodology and technology, data acquisition and processing and finally data interpretation and visualization. This needs to be established in combination with the development of technology platforms that overcome temporal boundaries, facilitating measurements ranging from (sub)seconds to minutes, from minutes to hours and hours to days, weeks and even months. This will allow the investigation of contextual molecular dynamics ranging from direct interactions to obtaining dynamic profiles over extended periods of time. Moreover, these technologies need to be functional over a wide spatial range, which will allow investigations of single molecules or cells up to whole organisms and their environment. Integration of these different types of data, over multiple temporal and spatial scales is essential to understand living systems as a whole. Dynamics of a single molecule, a molecule in the context of a cell, a cell in its native tissue context and large scale environmental influences on this dynamics are going to be the focus areas of method development.

The technologies to acquire data and analyze it in the context of the interaction of all the scales in larger living ecosystems as well as the effect these ecosystems have on the future of a healthy, diverse and sustainable society is an additional strong focus in the next decade. Acquiring the data is however only the first step in a data analysis process. Data analysis tools are essential for extracting useful information from such complex, high-dimensional and “multi-layered” data. This requires the development of new computational methods based on the latest, state-of-the-art tools developed in computer science, data science and statistics, with artificial intelligence as a key focus.

Finally, training and accessibility are important parameters to successfully implement, use and maintain such state-of-the-art technology platforms in the future. Experienced researchers are required to develop, implement and maintain innovative technology, while users need expert training to understand the acquired data and extract all available information. The realization and straightforward access to High Performance - and High Capacity Computing and further development of advanced Artificial Intelligence (AI) infrastructure for the life sciences is required to utilize the full potential of the acquired data. Furthermore, affordable solutions for data storage and data sharing are needed to deal with the increasing size and complexity of the different data types and allow easy integration of results across all scales of life.

4. Impact on society and contribution to topsectors (and related missions), national science agenda-routes and SDG’s

With the shared aim of understanding the living world, our multidisciplinary community connects innovations from computer science, biology, physics, chemistry, math, and engineering with urgent questions in the life sciences. We continuously exploit novel world-wide technological developments, which demands a very dynamic way of working to remain competitive at the international level. A crucial reason for Dutch success in the adaptation of the latest technologies, the invention of novel techniques, and the coping with the increasing complexity of our advanced methods thus far, is the strong engagement of Dutch organizations in our research community on the study of living systems.

Impact in this field is generated when high-throughput, cutting-edge experimental technologies are paired with powerful data analysis methods to develop new strategies for dealing with practical problems (from food production to diagnostics and treatment development), gaining new insights into complex systems (e.g. disease biology) and developing new product designs. This has a broad impact across the topsectors Agri-Food, Horticulture and Starting Materials, LSH, Chemistry, HTSM. The protein transition in our food is largely based on our understanding of the molecular understanding of the cellular systems that make up food resources. New genetic crop varieties result from the technology development in our research community. The production of real meat for the world’s growing population based on cell culturing or vegetable proteins was pioneered in the Netherlands has the potential on reducing our ecological footprint.

Molecular analysis, imaging systems and data sciences combined can study healthy cells and tissue and optimize food production processes as well as disease treatments that make health care more efficient. The same innovative approach can be applied in the study of the surface of chemical catalysts or active nanotechologies in high tech systems and materials. The active embedding of multidisciplinary teams of research groups at universities creates an environment of fast knowledge transfer, sharing of specialized equipment, and open interdisciplinary discussion, stimulating new research lines. The interdisciplinary education, that combines life science topics with technologies, stimulates the fast developments necessary for Dutch innovative research. This approach has a huge impact on many of the government-defined missions for which this research community provides the key technologies. At the same time these developments directly contribute to the sustainable development goals (SDG’s). Technology that improves our understanding of living systems that contribute to CO2 reduction, biobased materials and plant based energy harvesting is paving the way for a better future.

The central mission of the topsector Life Science & Health, for example, defines an increase of life-expectancy of five years for all Dutch citizens in 2040 and a 30% reduction of health differences between population groups with the lowest and highest socio-economic status. This requires our community to develop cheaper, faster, better, precise and comprehensive diagnostic methods that allow medical professionals to make better-informed and precise decisions on personalized treatment plans. Additionally the advanced methods, data and analysis generated will provide innovative information that feeds into novel medication that can intervene in the dynamic complexity of life.

5. Infrastructure in national & international perspective

Advancing technologies and analytical methodologies – and making these available through shared research facilities – directly impact all fields of biology. Creating infrastructures around advanced experimental setups in biology requires a deep understanding of both technology and biological specimens. This has resulted in a typically rich landscape of specialized labs, each offering dedicated expertise and instrumentation to address the enormous richness of biological systems across scales and domains ranging from human health and food production to biodiversity and the environment. Meanwhile, life science labs deliver a growing plethora of data types and formats that need to be combined, integrated, modelled and shared in collaborative research in national and international teams.

The Netherlands already has some very strong data collections that express the scientific value of well-collected, well-structured, well-annotated data. Examples are our strong population-based cohort studies like The Rotterdam Study, The Maastricht Study, Dutch Tweelingenregister, LASA etc, as well as Fungal collections at Westerdijk Institute, Biodiversity collections of NL-BIF and the plant seed collection in Wageningen. These data collections and studies offer overwhelming potential to future research, e.g. to address societal challenges, in combining data from genetics, biology towards environment and ultimately health and disease, but also breakthrough applications of artificial intelligence and machine learning.

Advancing our infrastructure is of paramount importance to keep ‘ahead of the art’ in global life sciences and stand out in attracting talent. Rather than establishing islets of small stand-alone labs, we need to raise the impact of technology even further and create a strong Dutch national network of interconnected and well-accessible life science research facilities. We need a fully networked research infrastructure for the next generation of Dutch experimental and digital biology! Bringing communities together in technology platforms such as the X-Omics platform, NL-Bioimaging, NeurotechNL, NEMI, and NPEC, the Netherlands EcoPhenotyping Cluster, offer distributed infrastructure, with multiple specialized nodes focusing on particular methodologies and expertise. The Dutch research community benefits from these platforms by sharing cross-disciplinary expertise, state-of-the-art equipment, and by coordinating funding opportunities. Each shared research facility, large-scale or medium-sized, employs technology and methodology specialists, biologists as well as digital life science experts. In a coordinated fashion, they operate as networked national nodes in the framework of European (ESFRI) and other global infrastructure and life science networks, gaining access to further capacities and resources of international labs. Crucially, granting procedures must be set up to support all scientists to make use of these expert facility nodes, where they can be assisted and trained, leading to novel collaborations and reducing redundant investments across the field. Universities and funders define a collective code of conduct in offering access to scientists outside their institutes, also for industry.

This extensive network of large scale research infrastructure, operating across all areas of biology, also builds upon a next generation federated but connected national life science data and ICT (e-)infrastructure. This digital layer of the research infrastructure is essential to support knowledge discovery and innovation in the increasingly data-intensive fields of life science research. In the future, research teams are offered a digital workbench and have easy access to external datasets, to advanced bioinformatics tooling, to data science and modelling expertise and to computational infrastructure. This process of digital life sciences is heavily supported by implementing the FAIR data principles across the field: facilities automatically deliver data FAIR at the source and expose datasets using the global web of FAIR data and services. Well trained life science data stewards are a crucial factor to support digital biology across the national life science research infrastructure. The biology community in The Netherlands needs a large new generation of biologists that is fluent in data science and AI. This means that we have to invest in new faculty in biological data science that can provide new generations of students with the proper background. The digital infrastructure layer is equally well-embedded in the European digital research infrastructure landscape, operating as node in ELIXIR and European e-infrastructures.

The above networked infrastructure fits well in the geography and culture of the Dutch biology field, will give a boost to cross-disciplinary and public-private life science research, and it will immediately stand out globally. Realizing it requires an integrated vision on the establishment and operation of infrastructures in biology, agreements on facility access across organizations and on the sharing of costs among all stakeholders.

6. Involvement of Dutch organisations in your research community

Within and between the Dutch universities there is an excellent attitude towards the combination of each of the strengths. Examples of this are the strengthening interdisciplinary ties between technology and medicine in for example the joint strategy of both research and education of Leiden-Delft-Erasmus, the Maastricht-Eindhoven axis and the OnePlanet collaboration on health, food and agriculture between Nijmegen and Wageningen. Inter-university collaborations allow smaller scale access to advanced knowledge and techniques, supported by the researches themselves, their universities and the funding agencies both national and international. Universities of Applied Sciences (HBO's), can play important roles in more applied research, and translate scientific results to actual implementation in society.

The joined research and education programs between the life science and technology departments of universities, and the medical centres are significant in the development and application of advanced methods for medical and health related research questions. Both technical universities (TU Delft; TU Eindhoven; TU Twente) and the other universities have strong connections with medical centres (AMC; VUMC; LUMC; UMCG; MUMC; Radboud UMC; ErasmusMC; UMCU).

With their specialized knowledge and technologies, focused on specific domains, the engagement of the Dutch research institutes in our community is very important. Institutes like the NIN (Netherlands Institute for Neuroscience), NKI (Netherlands Cancer Institute), Hubrecht Institute, Westerdijk, NIOZ (Netherlands Institute for Sea Research), NIOO (Netherlands institute of ecology), AMOLF, Naturalis, The Maastricht MultiModal Molecular Imaging Institute (M4I), NeCEN (Netherlands Centre for Electron Nanoscopy), the Netherlands Proteomics Centre, and the DTL/BioSB/ELIXIR/Health-RI combination together cover a broad range of the life sciences. Many of these specialist nodes collaborate in National Large Scale Infrastructure coordinated platforms. These large-scale scientific infrastructures are imperative in enabling access to areas otherwise unavailable due to their significant cost or great complexity. National examples are the NEMI (Netherlands Electron Microscopy Infrastructure), X-OMICS, Proteins@Work and NPEC projects, but also international initiatives like the Eurobioimaging, eScience Center and SurfSara play important roles.

The presence of large international organizations, especially with R&D in the Netherlands (for example: DSM, GlaxoSmithKline, Johnson & Johnson, UCB, ASML, Amgen, Phillips), the significant number of smaller biotech companies, medical device industry, pharmaceutical industry and non-profit organizations, like the TNO healthy living initiative and the international VPH-institute (Virtual Physiological Human), make the important translation of discovery to applications for society. This transition is further stimulated by the organization of bioscience parks (only the Leiden Bioscience park hosts 103 biomedical companies and 11 research institutes), public-private institutes like the Chemelot Institute for Science & Technology (InSciTe), not-for-profit organizations like Lygature (driving discoveries to societal implemented solutions), the Hub initiative of Utrecht university, and international organizations like the Avicenna Alliance. Spin off companies from scientific research lines also contribute significantly to the international position of the Netherlands and the valorization of Dutch research.

Taken together, this engagement of the Dutch organizations in our research community creates a stimulating environment, promoting innovation, enabling fast adaptation to (technological) developments, ensuring to remain at the international top.

7. Specific challenges for the community

The community needs to support durable science: The advanced methods are reaching such complexity that projects at the standard short funding cycle might not be the most rewarding. This longer cycle and greater complexity also needs more strategic and long-term investments in support, maintenance and updates. This includes the development of big data handling and analysis approaches that go beyond proof of principle studies to their high throughput application required to achieve our 10-year goals. Existing facilities should be able to update instruments / software and to maintain high-level support staff such that our Dutch facilities stay competitive. Importantly, this needs different funding that supports innovative method development and application across the life sciences.