

Faculty of Science and Engineering

Profile report: Biological Data Science and Biostatistics, Biologische Datawetenschap en Biostatistiek

- Discipline: data science, statistics, computational biology
- Level: Tenure-track assistant professor with education profile
- Fte: 1,0 fte

1. Scientific discipline

Biological data science is a discipline that develops methods and tools for the analysis of complex biological data, based on theory drawn from mathematics, statistics, computer science and information science. It also applies such methods to structured and unstructured biological datasets, in a way that is informed by biological knowledge, in order to extract new knowledge and test biological hypotheses. Examples of data science applications in biology include: the reconstruction of microbial community dynamics from metagenomic data; the inference of animal dispersal strategies from automated tracking data; or the application of machine learning to ecological niche modelling.

This position will predominantly focus on further strengthening and implementing new biostatistics and data science approaches in various courses in the Biology Bachelor and Master programs taught in GELIFES as well as contribute to ongoing research with biostatistics. The appointed candidate's research should preferably contribute to current research lines in the Groningen Institute for Evolutionary Life Sciences on biological Data Science and Biostatistics.

2. Vacancy

This position is opened by the Board of the Faculty (PT/gl/21/00266) and will be embedded in the Groningen Institute for Evolutionary Life Sciences (GELIFES). The criteria and conditions pertaining to the position are described in the document 'Assistant professor with an education profile'.

3. Selection committee (BAC)

- Prof.dr Rampal Etienne (Director GELIFES)
- Prof.dr Theo Elzenga (GELIFES education director; Chair)
- Prof.dr Sander van Doorn (Program Director BSc Biology)
- Prof.dr Bregje Wertheim (Program Director MSc Ecology & Evolution)
- Prof.dr Ido Pen (Professor of Theoretical Biology)
- Prof.dr t.b.d. (external member)
- student t.b.d.

Advisors: Linda Bosveld Verburg (HR), Prof.dr Simon Verhulst

4. Area of expertise

Biology is rapidly transforming into a scientific discipline where progress is no longer limited by the quantity of available data, but by the ability to analyse and interpret complex datasets in meaningful ways. Biological data science and bio-statistics confront these challenges by unlocking theories and techniques from statistics, mathematics and computer science for applications to biological data, in ways that are informed by biological knowledge. Research in this area is inherently interdisciplinary and

collaborative, and requires familiarity with state-of-the-art quantitative methods, as well as a solid understanding of biological principles, current developments in biological research, and the nature of biological data.

Multiple groups within GELIFES employ data science approaches in their research. Examples include the use of data reduction methods for the analysis of metagenomic sequencing data by microbial community ecologists; pattern detection techniques for the analysis of tracking data in animal ecology; or machine learning approaches to capture the dynamics of knowledge accumulation in models of cultural evolution in theoretical biology. However, none of the staff members currently involved in these projects qualify as a data science expert, so that available knowledge in this area is fragmented and insufficient to support the research ambitions of the institute. A data scientist would thus be able to fill a vacant niche, with many opportunities for research collaboration with established groups in different fields of biology.

Data science also deserves a strong position in any modern biology curriculum. In order to prepare our students for the sheer quantity and complexity of biological data that is collected in modern biological research, they need to be equipped with a sufficient basis in statistics, mathematics and computer science. After that, they need to become skilled in data science applications within their own field of specialization. The involvement of a data science expert in the teaching programs is direly needed to realize these ambitions, and to create coherence between data science elements in introductory courses and field-specific applications across different biological specializations. In addition, expertise is needed to provide advanced-level education in biological data science, for students who decide to specialize in this discipline. All of these initiatives require a high commitment to education, educational innovation and didactic skill development, because teaching data science and statistics to biology students depends on a deliberate strategy for promoting cross-disciplinary learning, starting from the development of elementary quantitative skills.

5. Embedding: institute (and expertise group)

The Groningen Institute for Evolutionary Life Sciences (GELIFES) aims to enhance the understanding of adaptive processes and the maladaptive consequences of their limitations, across all levels of biological organization (from molecules and genes to individuals and ecosystems), to inform the society and contribute solutions to societal problems. The institute has tight connections with the Faculty of Medical Sciences (FMS) and University Medical Centre Groningen (UMCG). It coordinates master programs in evolution and ecology as well as in medical and behavioural neurobiology.

GELIFES is organized in a non-hierarchical manner, and staff associate with one (or more) informal expertise groups. The tenure-track assistant professor is free to choose their expertise group. GELIFES currently has six expertise groups, each consisting of several professors and assistant professors with their groups: *Genomics Research in Ecology & Evolution in Nature* (GREEN), *Theoretical Research in Evolutionary Life Sciences* (TRES), *Evolutionary Genetics, Development and Behaviour* (EGDB), *Behavioural and Physiological Ecology* (BPE), *Conservation Ecology* (CONSECO) and *Neurobiology*.

The candidate will have access to GELIFES' excellent facilities, including IT facilities for large data processing.

6. Local and (inter)national position

Local:

Within FSE we teach the basics of Biostatistics in the Biology and Life Science & Technology bachelor programs. More advanced statistics courses are offered as electives for the various MSc programs in the general field of Biology. Our institute has strong links with the medical sciences at the UMCG and intends to participate in their initiative to offer a MSc track focused on data science in biomedical research. Data Science and Systems Complexity is a priority research theme of the Faculty of Science and Engineering, providing fertile ground for collaboration with sister research institutes in the faculty, primarily the Bernoulli institute for Mathematics, Computer science and Artificial intelligence.

National:

GELIFES has a strong reputation in research and education in ecology, evolution, behaviour and neurobiology. GELIFES is the only life science institute in the Netherlands that specifically aims at integrating the study of physiological mechanisms with those of ecology and evolution. Many collaborations exist with other universities and research institutes in The Netherlands on a wide variety of topics, including the universities of Wageningen, Utrecht and Amsterdam and the Royal Dutch Academy Institutes Netherlands Institute for Ecological Research and the Netherlands Institute for Sea Research as well as the Naturalis Biodiversity Center.

International:

There is no other institute in the Netherlands, and only very few in the world, that specifically aim at the integration of ecological and evolutionary approaches with neurobiology and physiology in the Life Sciences. Our international collaborations are too many to list but our research topics that are internationally very well recognized and relevant for the new staff member are among others the behavioural ecology of animal dispersal and migration, adaptation to climate change, microbial genetics and community ecology, evolutionary genomics, phylogenetics, self-organisation and biological complexity, and theoretical biology.

7. Expected contributions to teaching

The candidate is expected to contribute to the bachelor programs in Biology and the master programs in Biology and Ecology and Evolution. Existing education in which the assistant professor will be involved, include the courses Biostatistics 1 and 2 (BSc Biology), Big Data Management in Ecology and Evolution (BSc Biology), the MSc track Modelling in the Life Sciences (MSc Biology), Advanced Statistics (MSc). In addition, the candidate will develop a coherent program of data science education across the curriculum, ranging from introductory elements in existing BSc courses, to field-specific applications in BSc and MSc level courses, and culminating in an advanced-level data science course that is to be newly developed. The candidate is also expected to develop innovations in teaching in general. This contribution to the curricular development will be supported by successful applications for grants for educational innovation. The candidate will also supervise final research projects of bachelor and master students.

8. Expected contributions to research

The assistant professor is expected to establish a research line in biological data science and biostatistics, applied to a relevant field in biology (depending on the candidate's biological background). Second, the candidate is expected to contribute to efforts in the field of data science and biostatistics within the institute GELIFES by establishing research collaborations with existing groups. Finally, the candidate is expected to contribute to the

supervision of PhD students, supported by personal research funding and/or through collaborative/consortium grants.

9. Expected contributions to the organization

The candidate is expected to have an active interest and to provide a positive contribution to the management and organizational tasks of the institute. At the level of FSE, the candidate will contribute to the organization of the faculty, for example by participating in working groups and committees in the area of education. The candidate will participate in relevant national and international organizations.