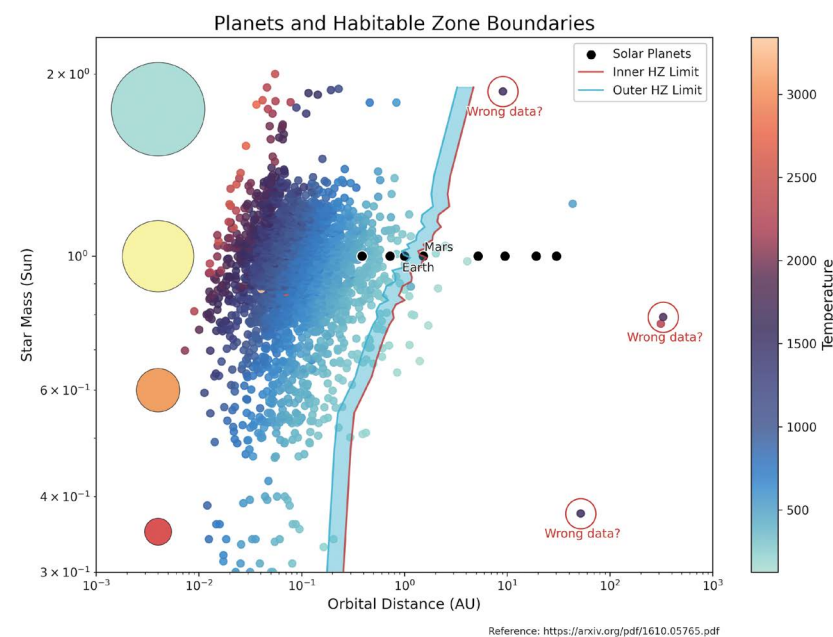# Training the next generation of data scientists

Research in life science is reliant on vast amounts of data, computational tools and statistics. Thus, there is a critical need for data science and bioinformatics expertise in the field. Researchers must learn skills in computational data analysis early on and develop an understanding of concepts such as machine learning and artificial intelligence (AI).

"Progress toward educating and training life scientists in the fundamentals of bioinformatics has been slow," explains Dr Venkata Satagopam, researcher in the Bioinformatics Core at the LCSB and co-chair of the Education committee at the International Society for Computational Biology (ISCB). With a group of international collaborators, he regularly reflects on what is needed to improve education in the field. In 2023, they published an article in *Nature Biotechnology* identifying the existing challenges, from defining the core competencies that students should master to supporting lifelong learning for researchers and keeping pace with computing and technology advances. "We want to provide a framework to help develop curricula that match the needs of different stakeholders and harmonise training worldwide."

At the LCSB, data scientists and bioinformaticians are highly sought. "We are having difficulties recruiting analysts," says Dr Roland Krause, researcher in the Bioinformatics Core. "So it was obvious that we should contribute to educate and train the future generations ourselves." When the University of Luxembourg launched a Master in Data Science directed by Prof. Yannick Baraud from the Faculty of Science, Technology and Medicine, several LCSB members were involved from the start, teaching courses ranging from data visualisation and AI to integrative bioinformatics and an introduction to biology for non-biologists. Over the course of the two-year programme, Dr Krause and colleagues Dr Andreas Husch, head of the Imaging AI group, Prof. Emma Schymanski, head of the Environmental Cheminformatics group, Prof. Alexander Skupin, head of the Integrative Cell Signalling group, and Prof. Anne Grünewald, head of the Molecular & Functional Neurobiology group, teach around 250 hours of lectures and practical workshops.

Data science has blossomed in the past decade as the amount of data collected in all aspects of life increased massively through digitalisation. "It is an interdisciplinary hybrid field that relies on computer science, statistics and analytical science," details Dr Krause. "It has strong ties to mathematics, but it is very connected to real-world questions and by now it is at the core of the scientific research process." By teaching students the basics of data science – collecting, managing and analysing large amounts of data while ensuring data privacy and security – they can then apply it to any field they want.

Plot created by a student of the master during the data visualisation course.

From left to right: Prof. Emma Schymanski, Prof. Anne Grünewald, Dr Roland Krause, Dr Andreas Husch and Prof. Alexander Skupin, lecturers for the Master in Data Science.

> Data management and analysis are now at the core of the research process.

Dr Krause continues: "These students will strengthen Luxembourg's work force in general and some will contribute to biomedical research. Several of them stayed in academia for their master's thesis and trained by working on research projects, so they are already helping with our needs in term of data scientists." Along the same lines, Dr Andreas Husch, who introduces the students to imaging AI and applying it to medical problems during the last semester, hopes to train scientists who could become members of his research group later on.

In December 2023, the Master in Data Science awarded its first diplomas to 14 graduates who are already forging their careers as analysts, data scientists and research assistants across various sectors, both in Luxembourg and internationally. Their teachers are now training around 30 new students from all over the world and from diverse academic backgrounds who will become the next generation of experts in this expanding field. ◼