New Scientific Director: Kai Polsterer

As of January 2025, **Kai Polsterer** has been the new HITS Scientific Director. He took over from Tilmann Gneiting, who had been Scientific Director for the last two years. The new Deputy Scientific Director is Rebecca Wade.

Kai Polsterer has served as head of the "Astroinformatics" group at HITS since 2013, the first European research group in that field. Polsterer has developed new methods and tools to deal with the exponentially increasing amount of data in that field, with a focus on developing tailored machine learning solutions. He works in various interdisciplinary projects with statisticians and medical scientists, to optimize weather forecasts as well as cardiological analyses.

Rebecca Wade leads the Molecular and Cellular Modeling Group at HITS. She was the first Scientific Director of HITS from 2015 to 2016. Rebecca Wade's research focuses on the development and application of computer-aided methods to model and simulate biomolecular interactions. Her research group has developed novel protein structure-based methods for drug discovery and protein engineering.

The Scientific Director is selected among the HITS group leaders and appointed by the shareholders. As the position of Scientific Director rotates through the group leaders, this is a planned change in the HITS management. The Scientific Director represents the Institute in all scientific matters vis-à-vis cooperation partners and the public.

Music from the Stars: Girls' Day 2025 at HITS
On 03 April 2025 from 9AM-1PM at Studio Villa Bosch.
Registration is now open!

HITS

Humboldt Research Prize winner at HITS

Victoria Stodden, Statistician and Associate professor at the Daniel J. Epstein Department of Industrial and Systems Engineering at the University of Southern California, USA, has joined the CST group at HITS in February 2025 as a recipient of the Humboldt Research Award.

The award, granted by the Alexander von Humboldt Foundation, honors internationally leading scholars for outstanding academic achievements. A key aim of the award is to promote international scientific collaboration. Award recipients spend up to one year at research institutions in Germany, collaborating with colleagues in their field.

Stodden's research is considered globally leading in the area of reproducibility in computer and data sciences. She investigates how to ensure the reliability and replicability of increasingly complex analytical methods. Since 2022, she has been Deputy Chair of the Scientific Advisory Board at HITS. During her research stay in Germany, her membership in the advisory board will be put on hold.

New employees and visiting scientists

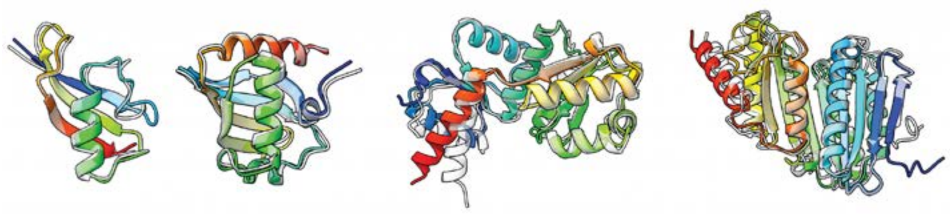
Postdoc:	Jordan Van Beeck (TOS)
PhD students:	Kristof Kraus (CST), Max Marius Heller (PSO)
Research Associate:	Beatrice Radice (SDBV)
Visiting scientists:	Victoria Stodden (CST, University of Southern California, USA, Humboldt Research Award), Lu Yang (CST, University of Minnesota), Federico Rizzuti (PSO, Heidelberg University), Nicolas Wolf (MLI, Heidelberg University), Marco Canossa Gosteinski (AIN, University of Groningen, The Netherlands)

HITS groups (03/2025): *Astroinformatics (AIN), Computational Molecular Evolution (CME), Computational Statistics (CST), Data Mining and Uncertainty Quantification (DMQ), Machine Learning and Artificial Intelligence (MLI), Molecular Biomechanics (MBM), Molecular and Cellular Modeling (MCM), Natural Language Processing (NLP), Physics of Stellar Objects (PSO), Scientific Databases and Visualization (SDBV), Stellar Evolution Theory (SET), Theory and Observations of Stars (TOS).*

HITSters

Machine learning: Advancing Protein Design

Proteins are fundamental to the existence and functioning of all living organisms and are responsible for numerous functions in biological systems. Artificially created, custom-designed proteins open up new possibilities for research, enabling their use for e.g., drugs, vaccines, or sensors. Machine learning has only recently made it possible to design entirely new proteins. However, one of the major challenges in protein design remains the creation of structurally rich and physically feasible "backbones" – chains of atoms that form the structure of macromolecules.
HITS researchers from the Machine Learning and Artificial Intelligence (MLI) and the Molecular Biomechanics (MBM) groups have now introduced a generative model for protein backbone design, utilizing geometric products and higher order message passing called GAFL (Geometric Algebra Flow Matching). This model is based on an extension of AlphaFold2, a Nobel Prize-winning machine learning software.



Efficient and realistic protein design
"In our model, the backbone fragments are represented as elements of projective geometric algebra – a mathematical construct for describing the geometry of points, lines and layers. We combine this new architecture with Flow Matching, a modern method for generative machine learning, to train a model that can create protein structures. One challenge here is to achieve a high designability without limiting the diversity of the structures," says **Leif Seute**, PhD student in the MLI group and one of the paper's lead authors.
GAFL was trained on approx. 25,000 protein structures with up to 512 backbone fragments. It is three times faster than comparable models and not only achieves high designability, but also a realistic distribution of secondary structures in the generated proteins. This is a property that many current methods only achieve inadequately. However, there is still room for improvement: the diversity of the generated structures decreases for particularly large proteins.

Paving the Way for Future Models
Driven by the positive findings, the groups are currently developing two new models that are based on GAFL but differ fundamentally in their mode of operation. The first approach involves generating so-called "conformational ensembles" of de novo proteins. "These ensembles show how the proteins move and change – an important property for the design of enzymes," says Leif Seute. The second model will investigate how GAFL can be adapted in order to determine the flexibility of the generated proteins. In the future, the researchers want to use GAFL to investigate further advantages for various protein-related tasks.

Seute, L., Wagner, S., Viliuga, V., Wolf, N., Gräter, F., and Stühmer, J.: Generating Highly Designable Proteins with Geometric Algebra Flow Matching, in: arXiv, 7 November 2024. https://arxiv.org/abs/2411.05238

Research

Guillermo Cabrera-Vives, Klaus Tschira Guest Professor

"I admire the way how research is done here," **Guillermo Cabrera-Vives** wraps up his first months as a Klaus Tschira Guest Professor at HITS. The astronomer and computer scientist from the University of Concepción, Chile, has been at the institute since September 2024.In this time period, he already participated in the Heidelberg Laureate Forum, held a HITS colloquium talk in October on machine learning and astronomy, and used the opportunity to delve into deep discussions with HITS researchers.
Since 2022, the Klaus Tschira Guest Professorship Program aims to enhance international exchange and scientific collaboration at HITS. Internationally renowned scientists are invited for sabbaticals or extended research stays. They collaborate with scientists at HITS, develop joint research projects, and engage with the wider scientific community at the institute and in the Heidelberg region.

Cabrera-Vives' research focus is on Machine learning in different application fields, including astronomy, data genomics, biomedical imaging, and satellite images. "HITS seems to me as a unique place where all this research happens at the same time." His main contact has been **Kai Polsterer** (Astroinformatics). They've known each other for several years. "Actually, I have been to this place before, when I participated in the Astroinformatics Conference in 2018," he says, an event that Polsterer had organized at the Studio Villa Bosch. Guillermo Cabrera-Vives also met **Tilmann Gneiting** (Computational Statistics) and **Jan Stühmer** (Machine Learning and Artificial Intelligence), and their group members to discuss applications, theoretical aspects and problems in Machine Learning. "I learned a lot of techniques how I can calibrate my models and measure their uncertainty," he states. "I would just have missed that not being here."

As an open-minded, interdisciplinary-oriented scientist, Cabrera-Vives is also interested in other fields: "I have talked with **Alexander Zeilmann** from the DMQ group about how Machine Learning can help in cancer treatment." Cabrera is pursuing a project where he uses AI algorithms to segment tumors to help radiologists monitor cancer treatment effectively. "We discussed about methods Alex uses to segment fossils, which are similar to the methods we use for tumor segmentation."

Guillermo Cabrera-Vives has also enjoyed the atmosphere at the institute: "I like the collaborative work and the spirit, you can feel it." He is also fond of the surroundings. "I live on the other side of the river, and I love to walk up the hill, it's a good exercise." Coming back to his first statement on how science is done here, he resumes: "Scientists usually think about the next paper. Here, people talk about the field, not about short-term, but long-term projects. They have the bigger picture in mind."

Imprint | Dr. Peter Saueressig (V.i.S.d.P.), saueressig@h-its.org, Tel. +49 6221 533 245
Pictures: HITS, ChatGPT (Girls' Day) | www.h-its.org

Beyond the limits

HITS

The Charts

Ausgabe 1 /2025