Dr. Marcel R. Haas - Astrophysicist turned data scientist
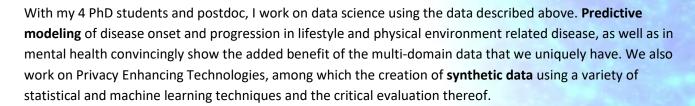
On this planet since 1983, now in Leiderdorp (NL). mail@marcelhaas.com

I am a passionate and investigative researcher in (applied) data science who loves a great, technical puzzle. Open Science and Recognition & Rewards are dear to my heart and I take joy in working with others to reach a large audience.

Professional Experience

Assistant Professor – Data Science and AI for population Health, Leiden University Medical Center (since 2023)

At the Health Campus The Hague of Leiden Univesity Medical Center, we maintain a data infrastructure that couples medical, socio-economic and socio-cultural data from many sources on the individual level. I **lead the technical team**, consisting of two data managers, a data steward, a data engineer and a data scientist in their work to maintain, extend and utilize this data. The outputs of this team are much more diverse than just scientific papers. We also assist the ecosystem of researchers with complying with **Open Science** standards, through data FAIRification and reproducibility.

With my 4 PhD students and postdoc, I work on data science using the data described above. **Predictive modeling** of disease onset and progression in lifestyle and physical environment related disease, as well as in mental health convincingly show the added benefit of the multi-domain data that we uniquely have. We also work on Privacy Enhancing Technologies, among which the creation of **synthetic data** using a variety of statistical and machine learning techniques and the critical evaluation thereof.

Astrophysics: (mostly) numerical simulations of galaxy formation (2006-2013)

My PhD thesis dealt with the sub-resolution physic in galaxy formation simulations. After graduating in Leiden in 2010, I did three years of postdoc in the USA. At Space Telescope Science Institute, we developed software for the first ever mock Hubble Space Telescope observations from the cosmological SPH simulations that I analyzed in my thesis. I led the team of 6 (including 2 graduate students) in our efforts to model astrophysical sources, the telescope and instrument and the methods observers would use on real data. At Rutgers University, I expanded that work to other galaxy formation models and made a start at simulating MeerKAT radio observations. During my second postdoc I decided to quit astronomy and turned to data science in industry.

Data Science in Industry (2013 – 2023)

At DSW Zorgverzekeraar (2013-2019), I established DataLab and was responsible for that team of four, using advanced analytics, machine learning, and data visualization for a broad variety of business problems (e.g. a multi-domain fraud detection system, customer journeys and satisfaction, churn, risk and IT performance). As senior data science consultant at ORTEC (2019-2020) I coached people at customers in technical roles (often data scientists) and I advised companies strategically in their digital and data driven transformations at all levels in the organization. I also built and deployed machine learning and AI solutions for a wide variety of applications in financial services, retail and energy industries. Based on sensor data and deep as well as reinforcement learning we develop a control model to minimize energy use and nitrogen emission from a Waste Water Treatment Plant that serves a million customers of Waternet. We built data-driven solutions for many different stakeholders, all in the Azure Cloud. A position as lead data scientist in the policy department of the University of Amsterdam opened the door for a re-entry into academia.

Teaching

At LUMC, I designed, coordinate and teach the Theme Week on AI and Medical Technology, a mandatory course for first year's medicine (BSc) students. I have also designed and will soon teach the Honour's College course Healthy AI, an interdisciplinary course focusing on the design and implementation of AI in Healthcare. I am the track coordinator for Data Science in the MSc program Population Health Management (PHM), where I also teach various data science modules throughout the first year and supervise students in their second year and thesis projects. I am also the competence coordinator and examiner of the Data Driven Thinking and Acting core competence of the PHM program. I have supervised bachelor and master thesis projects for PHM, biomedical data sciences, computer science, econometrics, data science and Health, Policy and Management and astrophysics.

At the University of Amsterdam, I taught guest lectures for Behavioral Data Science about data science and consulting. As a freelancer and for The Analytics Academy (joint venture between ORTEC and the University of Amsterdam) I developed and taught courses on analytics related topics, for both hard and soft skills, to a wide range of companies and at a range of skill levels (introductory to graduate level), in target audiences from management to analytics translators and data scientists, totaling about half a person-year FTE. I taught several guest lectures at The Johns Hopkins University ('Numerical Methods in Physics', 400 level) and was a TA for several (astro-) physics BSc and MSc courses in Leiden and Utrecht. From 2004 to 2006 I taught physics to HAVO 4&5 students at NOVA College Amstelveen.

Committees and Outreach

2025 – now: **UMCNL**: member of two AI expert groups: Education and Guidelines & Quality Management
2025 – now: **Future@Work**: Intervention design in a program for culture change in the workplace.
2025 – now: **AI and data science in LUMC education**: Developing guidelines and minimum learning goals for all educational programs affiliated to LUMC on AI and data science.
2025 – now: **(Gen) AI in education, research and business operations:** developing policy and preparing staff training for the use of generative AI in all core responsibilities of our department.
2024 – now: **Academia in Motion / Recognition & Rewards**: I take an active role in the Faculty AiM team (workstream R&R for (young) researchers), as well as in the departmental "PHEG in Motion" team.
2023 – now: **ELAN**: Member of the **Management Team** of the research and data infrastructure "Extramural LUMC Academic Network", as **technical lead** and coordinator of the Data Competence Center, defining its **strategy and technical roadmap**.
2024: **TV appearance** in FOCUS, a prime time popular science show, about AI in medicine.
2024: **Press coverage** in several national newspapers as well as other written and online media about our new BSc course on AI and Medical Technology.
2020 – 2022: Member of **Professional Field Committee**, Faculty of Mathematical Engineering, InHolland University of applied sciences
2019 – 2022: NumFOCUS Small Development Grants selection committee.
Before 2015: **National Academy of Sciences** (US), member of panel "On the State of the Postdoctoral Experience for Scientists and Engineers - Revisited". As postdoc representative in a committee at STScI, we designed and implemented a formal mentoring system. I co-organized the Dutch Astronomers Conference and was in a task force for the International Year of Astronomy. For 10 years, I was board member (including chair and treasurer positions) of amateur astronomy organizations in the Netherlands (JWG and KNVWS). I organized the first Dutch Astronomy Olympiad for high school students (worth the 2007 Kaiser Prize for astronomy popularization) and co-organized the EAAE summer school for high school science teachers.

Education

BKO (Dutch University teaching qualification): obtained at LUMC in 2025.
Leadership for Researchers: Leeuwendaal, 2024-2025
PhD in Astrophysics, Leiden Observatory (advisors: Joop Schaye and Marijn Franx), thesis: 'Nature and Nurture in Galaxy Formation Simulations', 2010
Master of Science Astrophysics ("cum Laude"), Utrecht University (supervisor: Henny Lamers), thesis: 'Star clusters in their host galaxies', 2006
Bachelor of Science Physics and Astronomy ("with honors"), Utrecht University, 2004
Gymnasium (grammar school, science major), Alberdingk Thijm College, Hilversum, 2001.
Other: Cloud computing on Azure, process mining, Deep Learning with SAS, Advanced Python Mastery (by David Beazley), Intercultural communication, Giving scientific and technical presentations.

Skills

Operating systems: Linux/UNIX, Windows, MacOS
Daily use: Python (including many data science and visualization packages), SQL, Version control (git), LaTeX, Microsoft 365
Experienced with SAS (and various related packages), R, NoSQL, Azure, AWS, Shell scripting (bash)
Basic knowledge: HTML/css, Fortran 77/90/95, MPI, IDL
Other languages: Dutch (mother tongue), English (fluent), German (basic)

Grant Support

1. PERFECT: Optimale dosering flucloxacilline voor effectieve behandeling cellulitis (ZonMw 10060012410020, PI: Martijn Sijbom)
2. ELAN - multidisciplinaire en transmurale versterking van hersen- en psychisch onderzoek (Nationaal Plan Hoofdzaken: kennis en data beter delen, ZonMw 10960302410003, PI: Erik Giltay, budget: 646k€)
3. GoHot: Gezond Omgaan met Hitte en pOllen in een veranderend klimaaT (task/deliverable lead, ZonMw 10241012410002, PI: Hein Daanen, VU, total budget: 800k€, LUMC budget: 125k€, Health Campus DH budget: 35 k€)
4. Phaeton: portable platform as a service for crowdsourced and privacy-respecting data analysis and modeling (WP-lead, ZonMw 10710062310030, PI: Jildau Bouman, TNO, total budget: 500k€, LUMC: 143k€)
5. LUMC Starter Grant: Avoiding Healthcare Avoidance (150k€)
6. Bid4Best, Innovative Training Network (industry partner, Marie Skłodowska-Curie grant agreement No 860744)
7. Observables in semi-analytic models of galaxy formation (NSF REU summer student grant, https://doi.org/10.13039/100000001, 30k$)

Private

I am the proud father of four lovely children. I enjoy puzzles and the boardgame Go (or: Baduk) and keep myself up to date on the current state of affairs in the astrophysics community. I am a certified beer sommelier and am active in the local and national craft beer scene.

Contact

mail@marcelhaas.com, +31 – 6 1151 5535