Patrick Bos

Curriculum Vitae

Personalia & contact information

Name	Evert Gerardus *Patrick* Bos	Birth date	27-02-1986
Address	Siriusstraat 28 9742KW Groningen	Marital status	Married
		Driver's license	yes (B, AM)
Telephone	06-10795874	Website	egpbos.nl
E-mail	egpbos@gmail.com		

Profile

Experienced scientist, research software engineer, research manager and entrepreneur looking for new challenges at the interface of research, technology and society. Background in physics, astrophysics, philosophy, theoretical and numerical mathematics, computational science, programming and (big) data science/analysis/visualization. Broad scientific appetite, having forayed into scientific fields as diverse as linguistics, (art) history, archeology, genetics and much more. Eager to keep combining my broad scientific interest with my passion for analytics and technology.

For my next step, I'm looking for ways to apply my helicopter view on digital technology for (academic) research in a role at a senior leadership level. I want to further develop the field and help others achieve maximum impact in it. My strong communicative skills and experience in outreach, teaching and (popular and scientific) writing will be vital assets in this effort.

As a practical idealist, I always strive to improve the world around me. I have been active for years in (university) policy making and steering democratic processes. I participate in strategy and policy making in the works council and working groups. As technology lead, I actively improve the organization's processes. On a political level, I contribute to the recognition of software as a first-class citizen in research, tying in to DORA and similar movements. In my new role, I want to continue adding value by taking on vital tasks within the organization and the broad landscape it operates in.

Work experience

2018 - **now** **Freelance Data Scientist, Research Software Engineer & Consultant**

- Started *PyData Groningen*, bringing together 200 participants from academia and industry in meetups with invited talks, sponsored for location, food and drinks.
- Ran engineering projects for e.g. for historians, archeologists and astronomers.
- Consulted on software quality, efficiency, usability and more, e.g. for Accelting.
- Led bid on M€+ SKA tender.
- Connected with core members of the Python data science ecosystem (Jupyter, Conda, etc.).

2022 - **now** **Technology Lead** *Software Quality* *Netherlands eScience Center*

- Formed "external projects team", advising on grant applications and processes
- Acquired the EU Horizon-Infra consortium project EVERSE (7.8M€ total, 570k€ for us)
- Management: hiring, coaching and evaluation of research engineers
- Strategy: e.g. wrote a technology forecast, gave input to policy and strategy
- Project management: keep overview of running projects, technologies used, share knowledge between projects, advise researchers and engineers.

- Knowledge development: define project programme for acquiring and developing key knowledge and technology in anticipation of their use in the near future.
- Knowledge dissemination: e.g. coordinated projects for engineering guidelines and Python & Julia templates to help beginning engineers

2014 - 2022 **Research Software Engineer** *Netherlands eScience Center*

- Cross-domain engineer with broad technological abilities, flexibly deployable.
- Successful lead engineer on high impact C++ particle physics projects and in diverse humanities projects with high deltas.
- Worked equally well solo, in duos or teams, switching effortlessly as demanded by circumstance. My presence and expertise is well appreciated by my collaborators.
- Provided IT support for the Center and led IT support team and IT strategy.

2007 - now **Teacher/mentor/supervisor** *Netherlands eScience Center, Kapteyn Institute, University of Groningen* and *University Education Center Groningen*

- NLeSC: developed and supervised internal training programme on software engineering and computer science (2022 - now)
- NLeSC: contributed to development of training programme for research software supporters (librarians, IT staff, etc.) (2023)
- NLeSC: supervisor/advisor on 3 AI projects *(2021 - 2024)*
- NLeSC: team mentor at *Workshop Machine Learning for Research (20-24 Jan 2020);* mentored **TNO Groningen** team on assimilation of seismographic data
- Kapteyn: *Programming and Numerical Methods* (2007, 2008, 2009).
- Kapteyn: *Astrophysical Hydrodynamics* (2011, 2012, 2013); rated **A+** by the students and education committee, the highest grade.
- UECG: Physics-teacher at *Exam Training* for high-school students (2009, 2010).

2007 - 2010 **Webmaster** *Kapteyn Institute, University of Groningen*

- Completely renewed website Kapteyn Institute (content, structure, lay-out).
- Maintained and technically supported this site for over three years.
- Webmaster and designer of our website for the Academic Year-Prize 2007. In doing so, provided our **communication** to the public.
- Webmaster and designer of the *national* astronomy education committee.

Education

2010 - 2016 **PhD** **Astronomy / Cosmology** *Kapteyn Institute, University of Groningen*

- Recovered initial conditions of the Universe based on observational data.
- Developed and programmed Bayesian inversion formalism.
- Processed large amounts of data in high-dimensional statistical model.
- Implemented efficient MCMC algorithm (C++) that cleverly makes use of the properties of the statistical and physical models we used.
- Successfully applied for X-ray space telescope (Suzaku) observation.

2008 - 2010 **MSc** **Astronomy** (cum laude) *Kapteyn Institute, University of Groningen*

Experience from thesis (e.g.):
- Ran N-body gravity simulations on large computer clusters.
- Made morphological analysis of large scale structures with Watershed Transforms.

2006 - 2009 **BA** **Philosophy of a Specific Discipline**
Faculty of Philosophy, University of Groningen

Thesis: classification theory, discussion analysis, logics, philosophy of science

2004 - 2008 **BSc** **Astronomy** *Kapteyn Institute, University of Groningen*

Thesis: big data (Astro-WISE, Virtual Observatory), classification

1998 - 2004 **VWO** **Nature & Technology** and **Nature & Health** (cum laude)
Esdal College, Emmen

Technology leader, expert and communicator

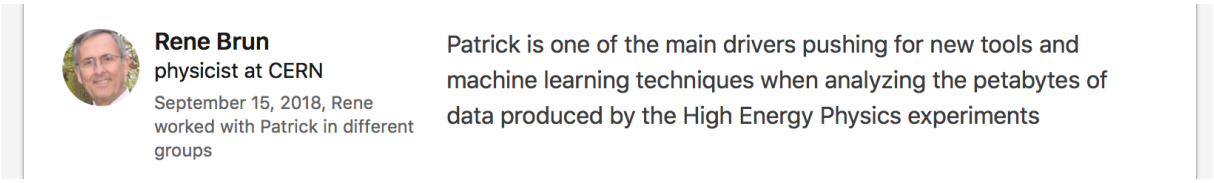
Recent invited talks

- 11 Apr. 2024, CERN, Geneva: RooFit parallelization developments
- 20 Feb. 2024, NRC newspaper, Amsterdam: Presented portfolio to data and science journalists
- 2 Nov. 2023, **TNO, Den Haag** (New Babylon): Research software engineering: why it matters

Gave many other talks at colloquia and conferences about project results, methods and other topics.

Recognition

- 2024 - now: Users' Committee member OTP project *SuperCode: SUstainability PER AI-driven CO-DEsign*
- 2024 - now: Expert member of EOSC *Opportunity Area 7 on Research Software*
- 2023 - now: Leader of EOSC *EVERSE* project work package on research software best practices.
- Testimonial (2018, LinkedIn) from creator of ROOT (particle physics data analysis package used by all large experiments over the past 2 decades) on my influence as a technological leader in the particle physics computing community:



Popular writing

I've written many well received technology-oriented articles between 2017 and now:

- Travis caching and incremental builds (DevOps/CI): 2700 views, 1100 reads
- Floating (point) butterfly effect (C++/numerics): 3900 views, 1300 reads
- 50 times faster data loading for Pandas (Python + C++): 111,000 views, 37,000 reads; second-most read blog post of the Center.
- C++ Compile-Time Exceptions (C++): 7100 views, 1500 reads
- Democracy (Python, election compass data, PCA): 1500 views, 340 reads
- Combining ZeroMQ & POSIX signals (C++): 4500 views, 909 reads
- Coalition polls for the people with Coalitiewijzer (Python/Jupyter, election data): 529 views, 172 reads
- Twitter Bots, for Science! (Python, COVID-19 vaccination tracker): 500 views, 134 reads
- Automate chores with GitHub Actions (DevOps/CI/CD): 498 views, 205 reads
- Ditching Docker Hub: serve research software with GHCR + Zenodo (DevOps): 394 views, 234 reads

I also wrote a conference report on the Information Universe conference of 2022 (151 views, 75 reads).

Total views: 133k. Total reads: 43k.
stats last updated: January 2025

International experience

2023 - now Work package leader of the *EVERSE* European consortium

2017, '19, '24 Work visits to CERN, Geneva

2018 Work visit to Paris, including visits to:
- the Orsay cosmology group of Nabila Aghanim
- QuantStack (**scientific computing company**; Jupyter & C++ devs)

2014 Work visit touring art historic institutes in Rome

2012 - 2014 Work visits (adding up to 2 months) to the AIP, Potsdam (Germany).

2010 *Astro-Computing* Summer School (3 weeks) at UC-HIPACC, Santa Cruz, CA, USA.

2008 - now **Presented work at (international) conferences**, among which:
- *The Cosmic Web in the Local Universe 2020 (Leiden)*
- *ACAT 2019* (on computational methods in physics, Saas Fee, 2019)
- *IEEE eScience 2018* (on digital technology for research, Amsterdam, 2018)
- *Digital Humanities* (Krakow, 2016)
- *Computational Astrostatistics workshop* (Lorentz Center, Leiden, 2014)
- *13th Marcel Grossman Meeting* (large physics conference, Stockholm, 2012)
- *Conferences and workshops in Haifa, Israel (2011), Warschau (2011), Potsdam, Germany (2012), Sarajevo (2018).*

Leadership, strategy and policy making

2022 - now As technology lead I performed numerous policy/strategic and management tasks:

Management
- Hiring of new research software engineers (RSEs)
- Performance evaluation of RSEs (e.g. for promotions or contract extensions)

Policy & Strategy
- Formed "external projects team" to *improve acquisition processes* and *advise colleagues on ongoing grant applications*
- Set out Generative AI policy (link)

Call programme
- Provide technological input for call texts and processes
- Join information events and consultation meetings with potential applicants to advise on technological aspects of proposals
- Review proposals: technical aspects (DMP, SMP, data & software health checks, etc.) as well as full proposals

2020 - now As **works council secretary** (until 2024), I was half of the "dagelijks bestuur" of the council.
- I managed communication and information flow in and around the council
- I organized meetings together with the council chair.
- We initiated professionalization of our communication and processes with the directors.
- We initiated regular open discussions to strengthen our connection with our constituency.
- We advised on policy for the 2021 restructuring, RSE role profiles, team work, the 2025 restructuring/downsizing and other issues of strategic importance.
- From 2024: regular works council member.
- From March 2025: **vice-chair**.

2017 - 2019 Led eScience Center IT support team

2011 - 2013 As **board member** of the Groningen Graduate School of Science I represented the interests of PhD students.

2011 - 2013 As **chair** of my faculty's PhD Council I organized activities and professionalized our tasks as a co-determination committee.

2009 - 2010 **Board member (secretary)** Bètastuf, overarching body for students in university councils.

2008 - 2010 Member *national* education committee astronomy (LOCNOC).

2007 - 2010 Member education committees astronomy (also as **vice-chair**) and philosophy.

Professional skills / trainings

2024 **Negotiation** by Bureau Zuidema

2023 **Strategy development, communication and persuasion** by YEARTH

2022 **Time management** by Ken & Ben

2022 **Conscious communication training** by Ken & Ben

2020-'25 yearly **Works council trainings** by Bureau Zuidema

2015 **Elevator pitch training** by Edo van Santen

2008 - 2014 **Public nights volunteer** at the Blaauw Observatory and other outreach activities.
2011 **Requested and obtained** observing time on Suzaku X-ray space telescope.
2011 **Presentation training** by Ludens Seminars & Private Coaching.
2007 - 2009 **Winning** Academic Year-Prize team (prize for **public science education**).
2006 Observed with the Isaac Newton Telescope on La Palma (Spain).

Miscellaneous skills

Language	Dutch (mother tongue), English (fluent), German (reasonable), Frisian (passive)
Social & organizational	Presenting, teaching, leadership, co-determination, entrepreneurship, networking, negotiation, coaching/mentoring, management, policy, strategy
Science	Data reduction and analysis, modeling, simulation, visualization and image processing, analytical and numerical mathematics, physics, astronomy, (Bayesian) statistics, (deep) neural networks, explainable AI, scientific publishing
Tech	**Python**, **C(++)**, Fortran, Java(script), Julia, HTML, SQL, LaTeX, shell scripting, macOS, Linux, AWS, Google Cloud Platform, Heroku, **ZeroMQ**, **Conda**, high performance computing, HPC clusters, GPU acceleration computing.
Research software	Jupyter, SciPy/PyData stacks, Matlab, Mathematica, many astronomical tools. *Created*: Barcode (cosmology), TICCLAT (OCR), DIANNA (XAI), Via Appia Visualization (archeology), xtensor-fftw (generic), kwandl (generic), ectopylasm (archeology). *Contributed to:* ROOT/RooFit (particle physics), GGIR (accelerometers), AMUSE (astronomy).

Selected scientific & scholarly publications

I. Bos, E.G.P., van de Weygaert, M.A.M., Dolag, K., Pettorino, V. (2012). "The darkness that shaped the void: dark energy and cosmic voids." *MNRAS 426* (1), 440-461. 200+ citations.
II. Shivashankar, N., Pranav. P., Natarajan, V., van de Weygaert, M.A.M., Bos, E.G.P., Rieder, S. (2015). "Felix: A topology based framework for visual exploration of cosmic filaments." *IEEE Transactions on Visualization and Computer Graphics 22* (6), 1745-1759. 100+ citations.
III. Bos, E.G.P., Wijfjes, H., Piscaer, M., Voerman, G. (2016). "Quantifying “Pillarization”: Extracting Political History from Large Databases of Digitized Media Collections." *The 3rd HistoInformatics Workshop* 1632, 57-66. 5 citations.
IV. Reynaert, M., van der Zwaan, J.M., Bos, E.G.P. (2019). "TICCLAT: a Dutch diachronical database of linked word-variants." In *Proceedings of DH Benelux Conference 2019*.
V. Bos, E.G.P., Kitaura, F.S., van de Weygaert, M.A.M. (2019). "Bayesian cosmic density field inference from redshift space dark matter maps." *MNRAS 488* (2), 2573-2604. 16 citations.
VI. Bos, E.G.P., Burgard, C.D., Croft, V.A., Hageboeck, S., Moneta, L., Pelupessy, I., Attema, J.J., Verkerke, W. (2020). "Faster RooFitting: Automated parallel calculation of collaborative statistical models." In *EPJ Web of Conferences 245*, 06027; CHEP 2019 conference proceedings.
VII. Wijfjes, H., Voerman, G., Bos, E.G.P. (2021). "Meten van verzuilde politiek in media: Een digitale benadering van katholieke en sociaaldemocratische dagbladen." *BMGN - Low Countries Historical Review 136* (3), 61-91.
VIII. Ranguelova, E., Meijer, C., Oostrum, L., Liu, Y., Bos, P., Crocioni, G., Laneuville, M., Cardenas Guevara, B., Bakhshi, R., Podareanu, D. (2022). "DIANNA: Deep Insight And Neural Network Analysis." Journal of Open Source Software, 7(80), 4493.
IX. Wolffs, Z., Bos, P., Brenner, L., Verkerke, W., & van Vulpen, I. (2024). "Efficient Parallelization of RooFit Computations for Accelerated Higgs Combination Fits." EPJ Web of Conferences, 295, 06007.
X. Kamiloğlu, R. G., Sun, R., Bos, P., Huber, F., Attema, J. J., & Sauter, D. A. (2024). "Tickling induces a unique type of spontaneous laughter." Biology Letters, 20(11), 20240543.
XI. Meijer, C., Bos, P. (2025). "Explainable embeddings." *in preparation.*

Past research projects

These are projects that I worked on myself as a researcher and engineer.

Data science

2014 - 2021 *PIDIMEHS - Measuring pillarization in digitized historical sources*
Question: can we measure political trends by combining data from digitized historical newspapers, parliamentary proceedings and other historical databases? In particular pillarization in 20th century Dutch society. How does the digital approach compare to traditional historiography?
Solutions: NoSQL data handling, data integration, information retrieval, analytics, visualization.
Answer: yes we can, but issues in large datasets like quality, accuracy and completeness is new to traditional "close reading" historians.

2014 - 2016 *HADRIANVS - A digital gateway to the Dutch presence in Rome through the ages*
Linked data, ontologies, triple stores, SQL databases, interoperability

Geometry & Visualization

2010 - 2013 *Topology as a tool for cosmology*
Involved in group efforts led by other group members, contributing with simulation data and software for data handling.

2013 - 2016 *Cosmic web component identification algorithm and statistics*
As a side-topic of my PhD thesis, I needed a simple way to identify separate components of the cosmic web. Building upon the group's Nexus+ method, I developed a simple algorithm to define "single" filaments and walls, delineated by voids and clusters. This in turn allowed for more detailed population studies of the cosmic web "objects".

2014 - 2016 *Via Appia Visualization*
3D pointcloud visualization of laser-scanned archeological site, web-based (JS)

2017-2018 *TICCLAT - Optical character recognition error correction visualization*

2019 *Sundial pointcloud interactive geometry tools*
Freelance: Built ectopylasm: a toolkit for geometrical analysis of ancient Greek sundials using laser-scanned pointcloud data.

Computational science

2011 - 2019 *Bayesian cosmological density field reconstruction*
The main topic of my PhD thesis.
Built a Bayesian inference code (C++) for sampling an ensemble of possible density fields given galaxy cluster input data, taking into account the variance from redshift distance estimates. The resulting reconstructions effectively cancel out anisotropies that similar works suffered from.

2016 - 2021 *Automatically parallelized likelihood fitting for particle physics*
Designed a custom asynchronous task execution framework within RooFit (C++). Frequentist statistics, Newton-Raphson fitting (DFP), POSIX signals, benchmarking, profiling, optimization, parallelization. Integration with CPU vectorization, GPU acceleration and automatic differentiation. Software design. Community management.

Machine learning

2015 - 2016 *Sherlock DL: car identification with deep learning models* (for NFI)

2019 - 2021 *Understanding visually grounded spoken language via multi-tasking*
Multi-modal modeling (images, audio, text), embedded spaces

Supervised research projects

These are projects in which I closely supervised (junior) colleagues, PhD students and advised research teams.

Machine Learning

2021 - 2022 *Epigenetic signatures for complex diseases*
Weekly supervisory/advisory meetings with Amsterdam UMC group of Peter

Henneman (genetics).
Question: can we identify pathology from (measures of) the epigenome?

2021 - 2024 *A different kind of laughter*
Weekly supervisory/advisory meetings with UvA group of Disa Sauter (social psychology).
Question: can we discriminate between audio signatures of different types of laughter using ML methods?

2023 - 2024 *Cerebrovascular Brain Age*
Mentored RSE Candace Makeda Moore (medical doctor) on software engineering best practices, code modernization, Python packaging.

Performance optimization

2022 *Review Argumentation at Scale*
Mentored RSE Laura Ootes (astronomer) on optimization, load balancing, benchmarking.

2021 - 2023 *Optimized parallel calculation of complex likelihood fits of LHC data*
Mentored particle physics PhD student Zef Wolffs on benchmarking, optimization, C++ programming, coding best practices and tools (git, code reviews, etc.).

Research Software Engineering

2024 Mentored RSE Olga Lyashevska (ecologist) on software engineering, community building, project management.

Managed research projects

As Technology Lead at NLeSC I also managed a portfolio of projects where my involvement was less intensive, reviewing and advising on matters like software design, project planning and evaluating overall progress, as well as keeping track of technologies used, and technology decisions made. I managed 40+ projects over the years. See my Research Software Directory profile page for a complete overview.

Ongoing research projects

I. ***Explainable AI (XAI)***
   A. Senior researcher and advisor on the **DIANNA** project (*Deep Insight and Neural Networks Analysis)*, exploring systematically the most useful XAI methods, selected and reproduced from the literature based on needs from the broad research landscape, made easily and uniformly accessible. Defined benchmarks to test these and new methods intuitively.
   B. **Explainable embeddings**: developed a new method specifically suited for so-called "embedded spaces": vector spaces that are used for representing many types of data in ML models. Made popular through early ML-based language models like Word2Vec where a vector would correspond to a word, they are now used for many other purposes, encoding meaning or other properties of a given data modality, or even multiple modalities at once (like images and corresponding text captions). There are no publicly available XAI methods for this key modern ML model component, so we developed one. The article is under final preparations, but the code is already available on GitHub.

II. ***Particle physics statistical analysis software***: As RooFit core developers team, we meet bi-weekly to discuss the latest challenges for the large collaborations at CERN LHC experiments like ATLAS, CMS, LHCb, and others, and the improvements necessary to accomodate these.

III. ***SuperCode: Sustainability PER AI-driven CO-DEsign***: I advise the group of Broekema (ASTRON) and Van Nieuwpoort (Leiden) on (astronomical and other) software to optimize using generative AI (Vision paper).

IV. ***Research software engineering***: quality, incentives, best practices, policy.
   A. **Acquired:** EVERSE (7.8M€ **Horizon-Infra grant**, **570k€** for NLeSC). A project to map the research software best practices used in the research communities across Europe, consolidate them and help researchers find them, improving the quality of research

software (and hence: research) across Europe. Vision paper on Zenodo: https://doi.org/10.5281/zenodo.10183077

B. Lead of NLeSC Guide project, which gathers a curated set of best practices for RSEs to get started in a project that requires programming technologies that we already have experience with. Latest release on Zenodo: https://doi.org/10.5281/zenodo.4020622
C. Lead of NLeSC Software Templates projects (Python, Julia) that essentially implement the best practices gathered in the Guide, helping RSEs to get going on new projects quickly.
    1. Latest release on Zenodo: https://doi.org/10.5281/zenodo.4734210
    2. Recent talk at RSECon: https://doi.org/10.5281/zenodo.14290439
    3. Recent workshop given: https://doi.org/10.5281/zenodo.14290341
D. Generative AI policy: I contributed to the definition of the "Policy concerning the use of Generative AI at the Netherlands eScience Center", available at Zenodo: https://doi.org/10.5281/zenodo.10363728. My contribution was mainly on software engineering aspects.

References

- prof. dr. Wouter Verkerke, ATLAS group leader at NIKHEF & professor of Data Analysis in particle and astroparticle physics at UvA. verkerke@nikhef.nl.
- prof. dr. Rob van Nieuwpoort, professor of Efficient Computing and eScience at UL. r.v.van.nieuwpoort@liacs.leidenuniv.nl
- prof. dr. Rien van de Weijgaert, professor of Cosmology and Large Scale Structure, RUG, weygaert@astro.rug.nl
- dr. Nicolas Renaud, Director of Technology, Netherlands eScience Center, n.renaud@esciencecenter.nl
- dr. Vincent van Hees, owner, software engineer and consultant, Accelting, v.vanhees@accelting.com