

Andrew John Lowe

Nagysándor József utca 8. I./4, 1195 Budapest, Hungary · Relocating to Berlin, Germany
✉ andrew.john.lowe@gmail.com · [in linkedin.com/in/andrewjohnlowe](https://www.linkedin.com/in/andrewjohnlowe) · [S andrewjohnlowe](https://twitter.com/andrewjohnlowe)

PROFESSIONAL PROFILE

I am a British data scientist with 4 years' professional experience consulting for a range of clients and over 15 years' experience working at the forefront of scientific research within large international collaborations. I have a PhD in particle physics and spent several years based at the European Organization for Nuclear Research (CERN) in Geneva and was a member of the team that discovered the Higgs boson. My core competencies include statistical data analysis, machine learning, quantitative problem solving in cross-functional teams, software development, mathematical modelling, data visualisation and interpretation of results. I am a co-author of 534 peer-reviewed scientific publications and have spoken in numerous international workshops and conferences.

EMPLOYMENT

Jun. 2017–
present **Data Scientist**
Enterprise Analytics Division, EPAM Systems Inc., Hungary

Acted as a consultant, working directly and collaboratively with clients on a range of projects:

London Stock Exchange Group (LSEG) (May 2018–present)
(Originally Thomson Reuters Finance & Risk, became Refinitiv in Oct. 2018, merged with LSEG in Feb. 2021.)

Devised, implemented and maintained software in R and Python to anonymise sensitive personal data to ensure compliance with EU GDPR data protection and privacy regulations, thereby derisking a 250 million US dollar per annum revenue stream for the client business. The software ingests the data and uses machine learning to generate a structurally and statistically identical synthetic dataset with strong privacy guarantees. The data itself is the World-Check database of Politically Exposed Persons (PEPs) and heightened risk individuals and organisations, used around the world to help financial institutions, the regulated non-banking sector and companies identify and manage financial, regulatory and reputational risk.

Bayer AG (Nov. 2017–Feb. 2018)

Built ETL processes using KNIME for automated Excel financial report generation. Reduced report generation time from over 24 hours to under 2 hours.

UBS Wealth Management Americas (Jul. 2017–Sep. 2017)

Helped develop a cloud-hosted chatbot, using NLP platforms Rasa and IBM Watson Conversation, to process customer service requests. Generated synthetic training data for chatbot.

Sept. 2013–
May 2017 **Scientific Research Fellow**
Wigner Research Centre for Physics, Hungarian Academy of Sciences, Hungary

Performed statistical data analysis for the ALICE experiment at CERN, which recreates conditions that are believed to have existed a fraction of a second after the Big Bang. Used state-of-the-art machine learning to develop predictive classification algorithms for recognising particles based on their decay properties. Co-founded the ALICE Statistics and Machine Learning Working Group. Co-organiser of the first CERN workshop dedicated to the use of machine learning in particle physics. Engaged with local data science community via public outreach talks and conference presentations.

Apr. 2010–
Oct. 2012 **Postdoctoral Fellow, Deputy Team Leader**
California State University, Fresno, USA (based 100% at CERN, Switzerland)

Systematically investigated the potential benefit of hundreds of different predictor variables for a range of data-mining analyses using Monte Carlo simulations written in C++. Found several new variables that provide significant improvements in sensitivity for a variety of Higgs boson and new particle searches, thereby vastly reducing the amount of expensive experiment running time required to collect sufficient data for offline analysis. Developed and productionised core software and algorithms for real-time data analysis.

Feb. 2008–
Aug. 2009 **Postdoctoral Fellow**
Indiana University, USA (based 100% at CERN, Switzerland)

Developed an algorithm in C++ and Python for real-time particle identification in streaming data at an input rate of 1 GB/s. Optimised algorithm parameters and achieved excellent performance. This algorithm underpins a large part of the ATLAS experiment's physics programme by providing the data used by about 3000 physicists for many analyses. It has been used in production for data-taking since 2010 and has processed tens of petabytes of data for subsequent offline analysis.

Mar. 1998–
Sept. 2000 **Assistant Research Scientist**
Centre for Time Metrology, National Physical Laboratory (NPL), UK

Provided technical and administrative support to a range of key activities relating to the maintenance and dissemination of the UK's national time scale.

EDUCATION

2001–2008 **PhD Particle Physics**
Royal Holloway, University of London, UK (including 17 months at CERN, Switzerland)
Thesis title: *Performance and robustness studies of the trigger for the ATLAS experiment*

Played a major role in the development of the core software and algorithms in C++ for a real-time multi-stage cascade classifier that filters and reduces the collision event data rate from 60 TB/s to a manageable 300 MB/s that can be written to permanent storage. Performed detailed time profiling of the core software and devised improvements that made it 8 times faster, thus meeting a critical requirement of the system. Wrote software that was used in the discovery of the Higgs boson.

2000–2001 **MSc Particle Physics**
Royal Holloway, University of London, UK
Thesis title: *Light Higgs ($H \rightarrow b\bar{b}$) at the LHC*

Investigated the search potential of the $H \rightarrow b\bar{b}$ decay channel for a light Higgs boson using the ATLAS detector at CERN. This was the first data-mining analysis of this type to be performed entirely in C++, and as such served as an example for other researchers to follow.

1993–1996 **BSc (Hons) Physics**
Royal Holloway, University of London, UK

TRANSFERABLE SKILLS

- **Communication:**
Can work closely with clients as a consultant and communicate effectively with other departments and business stakeholders to discuss complex data-driven findings and technical specifications. Can translate client requirements into highly specified project documentation. Invited speaker at numerous international conferences. Visual storytelling and data visualization best practices. Teamwork in a collaboration of about 3000 people.
- **Problem solving:**
Capacity to lead independent research, understand and dissect complex problems and find creative solutions.
- **Project management:**
Accustomed to handling parallel projects with strict deadlines, both on-site and remotely. Knowledge of Agile and Waterfall software development methodologies.

COMPUTING SKILLS

- **Programming languages:** R, Python (NumPy, pandas), C++, Octave
- **Software development:** Docker, Jira, UML, Git/GitHub/GitLab, RStudio, Valgrind, Visual Studio Code
- **Data mining software:** KNIME, Weka, XGBoost, H₂O, caret, mlr, ROOT
- **Documentation:** L^AT_EX, Markdown, R Markdown, Jupyter Notebooks
- **Operating systems:** Unix, Linux, Microsoft Windows
- **Data visualisation:** TIBCO Spotfire, Tableau, ggplot2, matplotlib
- **Other:** object-oriented analysis and design, CPU and time profiling, code optimisation, memory debugging.