From physics to data science

Andrew Lowe

14 June 2017

Early history (1993–1997)

- 1993–1996: BSc in physics, Royal Holloway, University of London
- 1997: TV & VCR retuning engineer, Channel 5 TV
  - Was the 5th national terrestrial analogue TV channel in the UK
  - Problem: TV broadcast frequencies had been allocated to RF outputs from domestic video cassette recorders (VCRs)
  - To achieve national coverage, large numbers of VCRs had to be retuned at Channel 5's expense before broadcasts could start
  - Pre-launch, huge campaign to adapt public's VCRs to make them immune to RF interference

Pre-launch test transmission

Video playback with RF interference from Channel 5

Reflections on being a Channel 5 retuner

- My first customer-facing role
- Visited several hundred homes
- Had to gain the trust of the homeowner to grant me access to their home and their VCR
- Job provided me with many interesting glimpses of the inside of people's homes and how their owners live
- Many surprises behind doors only 5 minutes away from where I lived at the time!

Time metrologist (1998–2000)

- Assistant research scientist at the National Physical Laboratory (NPL)
- NPL: the UK's national measurement standards laboratory
- Provided technical and administrative support to a range of key activities relating to the maintenance and dissemination of the UK’s national time scale
- An interesting place to work during the countdown to the start of the new millennium
  - Much public concern about the “y2k bug” and the meltdown of civilisation that was forecast to result!

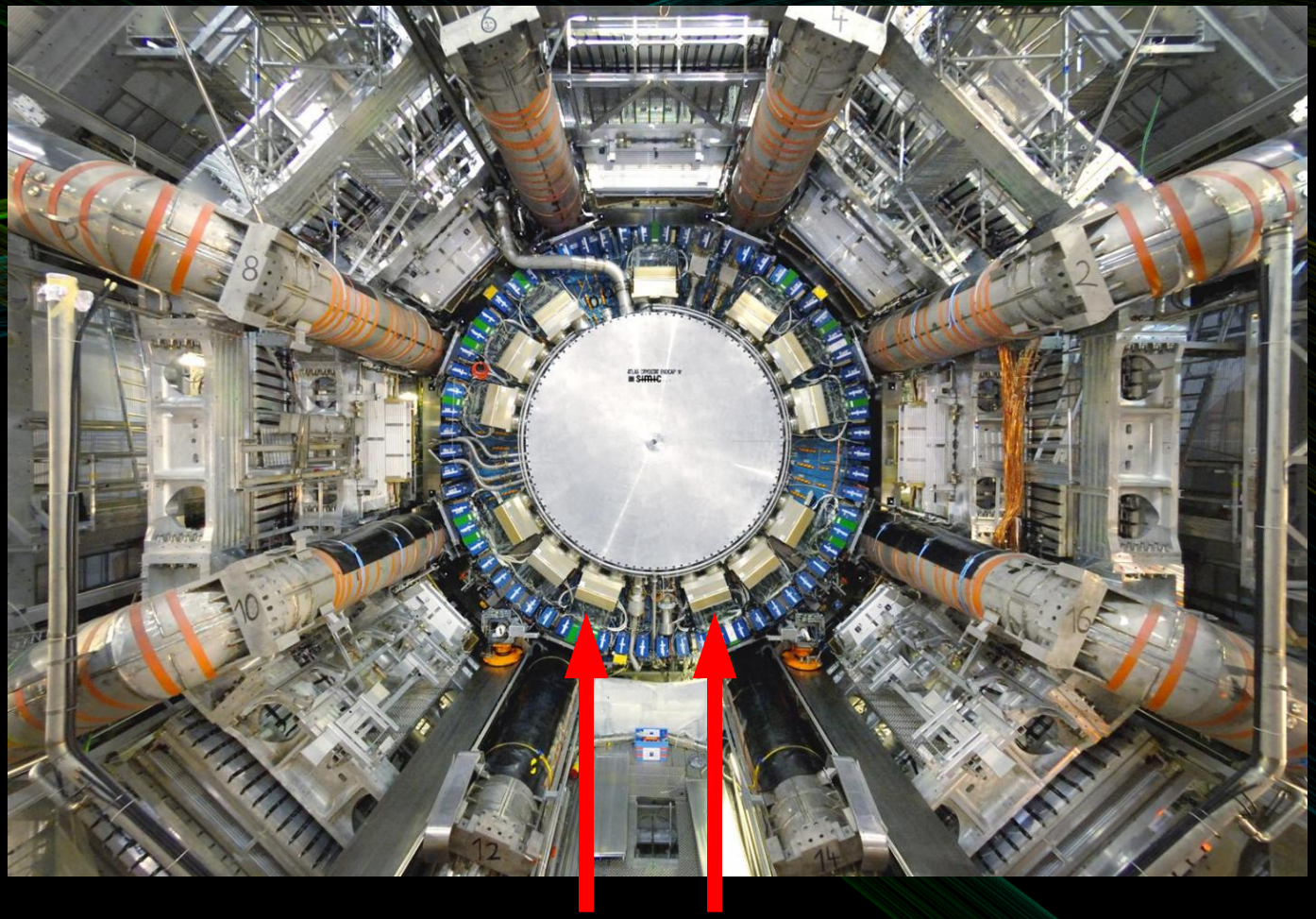
Things I did at NPL

- Maintained atomic clocks
- Programmed the start and end of daylight savings time
- Collected and analysed GPS timing data
- Operated a satellite earth station
- Published monthly bulletins
- Manned an enquiry service that received hundreds of phone calls, letters and emails from the public

Return to university (2000–2008)

- MSc and PhD in particle physics, Royal Holloway, University of London
- MSc thesis:
  - Data mining analysis of the search potential for the Higgs boson in a notoriously difficult decay channel
- PhD thesis:
  - Multiple key contributions to a real-time multi-stage cascade classifier that reduces data rate from 60 TB/s to a manageable 300 MB/s that can be written to permanent storage
  - Measured how performance of the data collection system is degraded when electronic components in the detector die

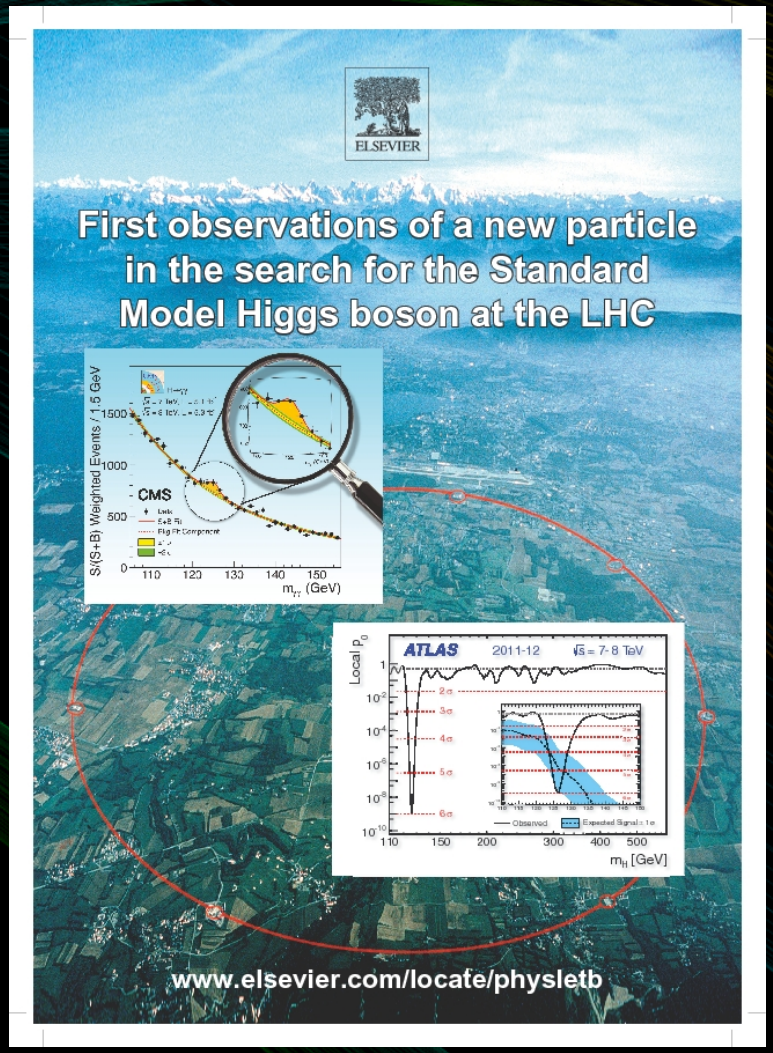
I studied how much data is lost when electronic components die



Electronic boards in crates mounted on-detector

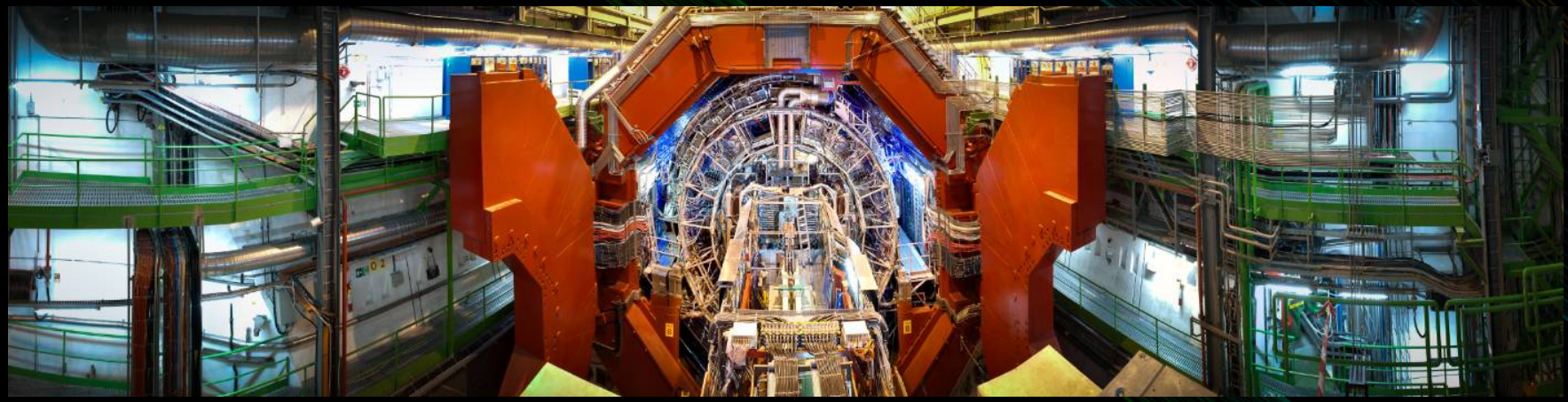
Life as an academic (2008–2017)

- Particle physicist (2008–2017)
  - Indiana University, USA
  - California State University, Fresno, USA
  - Wigner Research Centre for Physics, Hungary
- I 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, the observation of which led to the award of the 2013 Nobel Prize in Physics
- Moved to Hungary in 2013



Things I did as a particle physicist

- Developed algorithms for real-time particle identification in streaming data at an input rate of 1 GB/s (total data processed to date: 70 PB)
- Performed statistical data analysis for an experiment at CERN which recreates conditions that are believed to have existed a fraction of a second after the Big Bang
- Used machine learning to develop predictive classification algorithms for recognising subatomic particles based on their decay properties



Data scientist (2017—present)

- Now I'm at a major milestone again
- Recently started a new role as a data scientist for **EPAM**
- Opportunity to work with different types of data in different problem domains and have an impact on real-world problems
- As a physicist, I learned how to:
  - Work with data, including really big data
  - Understand and dissect complex problems and find creative solutions
  - Work in large cross-functional teams
  - Visualise and interpret results
- Easy to make the switch to data science, just needed to learn the language and tools