Online Appendix to the discussion of Jones' Artificial Intelligence in Research and Development: The evolution of research in particle physics

Bronwyn H. Hall

27 September 2025

In order to provide a detailed example of the impact of automation on research, this section describes the evolution of fundamental research in an area with which I am somewhat familiar: "bump hunting" in high energy physics. This term describes the practice of looking for new particles or new phenomena by repeating the same interaction many times using some kind of particle accelerator and then examining the distribution of outcomes statistically. Beginning in the 1960s as an essentially manual procedure once the relevant experiment had been conducted, this type of research has evolved to use machine learning and AI to become (almost) fully automatic. Much of the improvement in efficiency has been due to the fall in computing cost and the shift away from humans to software for analysis.

In the 1960s and early 1970s, bump-hunting was a fundamentally human-centric activity. Physicists would meticulously plot experimental data by hand or with early computer programs, creating histograms of particle interaction properties such as invariant mass. They would then visually scan these plots for interesting deviation from what they expected to see (normally a rather smooth distribution) and identify these with new or known particles. The first step in this process was to define an interesting experiment that would be repeated thousands of times so that a distribution of the possible results could be collected. For example, a typical experiment at the Lawrence Radiation Laboratory (now the Lawrence Berkeley Laboratory) Bevatron involved firing a proton into the bubble chamber filled with liquid hydrogen and photographing its track as it interacted with the bubbles, possibly producing other particles that then decayed.

The experimental data itself was often collected by using the FSD (Flying Spot Digitizer) on each photograph of an interaction (of which there might be thousands). The operators responsible for this were skilled at the task, but were not trained physicists, nor were they required to even have a bachelors degree. The resulting data were then fitted by a software program to yield information on the particles emitted and their energies. Distributions of the results were plotted, usually manually, and examined for evidence of anomalies that would indicate the presence of a new particle or phenomenon. For some examples of this kind of research, see Bland et al. (1968) and Goldhaber et al. (1969).

Over time, several advances were made in this process: first, statistical techniques were developed to deal with the problem of false positives and make the finding and testing for anomalies more rigorous.[1] Second, the bubble chamber was replaced with other kinds of

[1] The problem should be familiar to economists – as the number of signal searches increased, the number of false signals also increased at any fixed significance level. The solution used in high-energy physics is to

detectors that could operate at much faster rates and provided electronic readout, obviating the need for FSD operators. For example, the Nobel prize-winning discovery of the charm quark *J/ψ* simultaneously at the Stanford Linear Accelerator (SLAC) and Brookhaven National Laboratory was done by the SLAC team using a series of spark chambers that recorded the results of the interactions of positive and negative electrons. (Augustin et al. 1974, 1975). The Brookhaven team used a proton-proton interaction and spectrometers to find the same result (Aubert et al. 1974).

Third and more important for the shift towards the use of machines and software, advances in computing allowed a number of tasks formerly done by humans to shift to machines and for new methods to become possible. For example, computational speed improvements allowed rapid scanning of vast amounts of data in search of rare "resonances," that is, particles not yet discovered. For example, the Collider Detector at Fermilab (CDF) was used to analyze approximately 10^{14} proton-antiproton interactions by searching for statistical anomalies in their outcomes relative to the outcomes predicted by the standard model of particle physics. This exercise produced a negative result as it detected only the number of anomalies predicted by simulation of the standard model (CDF Collaboration 2009).

The current frontier technique in this area augments the task with machine learning, implying that AI methods have already been introduced in this area. The successful search for the Higgs Boson at the CERN Large Hadron Collider made extensive use of machine learning to classify the outcome of the particle collisions and select those for further analysis (ATLAS 2012; CMS 2012). A glance at the more recent publications in this area gives an idea of the increase in complexity of the analysis since the 1960s (Collins et al. 2018, 2019).

How does this historical development of a research history relate to the model in the paper? For example, we might define the desired outcome Z as the discovery of new particles. However, as the perceived difficulty of finding yet another particle increased over time ($\varphi<0$), clearly the budget D was not held constant. Besides the enormous increase in the cost of the necessary particle accelerator over time,[2] the number of authors on the relevant papers cited increased from a handful to several hundred. It is probable that this dynamic (diminishing returns and increased budget) is characteristic of many individual lines of research, as suggested by the examples in Bloom et al. (2020).

A second observation is that even though increased computing speed and other technical improvements meant that much of the original effort devoted to recording the experimental data and analyzing it moved from humans to machines this was accompanied by the expansion of tasks to manage both the experiment and the analysis of the results.

---

simulate the background distribution from known physics many times and compute how often a random significant bump would arise.

[2] Initial costs in 2025 dollars for the Bevatron was $110 million, for Fermilab's Tevatron $375 million, and for CERN's Large Hadron Collider $8.5 billion. These are of course approximate figures and do not include important later enhancements.

That is, the impact of improvements in machine technology was mainly to increase the difficulty of the research undertaken and the cost of that research. in this respect the activity is not unlike other not-for-profit endeavors such as health care and performing arts (Bowen and Baumol 1965).

A last related observation is that although the impact of the increase in computing speed and detector design over the past 70 or so years has meant that many tasks have moved from humans to machines, many new tasks have arisen in making this transition, many of them in the form of software design done by humans. It remains to be seen whether AI can help with this task or will simply make the research more ambitious in both cost and desired outcomes.

References

ATLAS Collaboration (2012). Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters* B 716(1): 1-29. https://doi.org/10.1016/j.physletb.2012.08.020

Aubert, J. J., U. Becker, P. J. Biggs, J. Burger, M. Chen, G. Everhart, P. Goldhagen, J. Leong, T. McCorriston, T. G. Rhoades, M. Rohde, S. C. C. Ting, S. L. Wu and Y. Y. Lee (1974). Experimental Observation of a Heavy Particle *J*. *Physical Review Letters* 33(23): 1404-1406. https://doi.org/ 10.1103/PhysRevLett.33.1404

Augustin, J. E., A. M. Boyarski, M. Breidenbach, F. Bulos, J. T. Dakin, G. J. Feldman, G. E. Fischer, D. Fryberger, G. Hanson, B. Jean-Marie, R. R. Larsen, V. Lüth, H. L. Lynch, D. Lyon, C. C. Morehouse, J. M. Paterson, M. L. Perl, B. Richter, P. Rapidis, R. F. Schwitters, W. M. Tanenbaum, F. Vannucci, G. S. Abrams, D. Briggs, W. Chinowsky, C. E. Friedberg, G. Goldhaber, R. J. Hollebeek, J. A. Kadyk, B. Lulu, F. Pierre, G. H. Trilling, J. S. Whitaker, J. Wiss and J. E. Zipse (1974). Discovery of a Narrow Resonance in *e+ e-* Annihilation. *Physical Review Letters* 33(23): 1406-1408. https://doi.org/10.1103/PhysRevLett.33.1406

Augustin, J. E., A. M. Boyarski, M. Breidenbach, F. Bulos, J. T. Dakin, G. J. Feldman, G. E. Fischer, D. Fryberger, G. Hanson, B. Jean-Marie, R. R. Larsen, V. Lüth, H. L. Lynch, D. Lyon, C. C. Morehouse, J. M. Paterson, M. L. Perl, B. Richter, R. F. Schwitters, F. Vannucci, G. S. Abrams, D. Briggs, W. Chinowsky, C. E. Friedberg, G. Goldhaber, R. J. Hollebeek, J. A. Kadyk, G. H. Trilling, J. S. Whitaker and J. E. Zipse (1975). Measurement of $e^+ e^- \rightarrow e^+ e^-$ and $e^+ e^- \rightarrow \mu^+ \mu^-$. *Physical Review Letters* 34(4): 233-236. https://doi.org/10.1103/PhysRevLett.34.233

Baumol, W. J. and W. G. Bowen (1965). On the Performing Arts: The Anatomy of Their Economic Problems. *American Economic Review* 55(1/2): 495-502.

Bland, R. W., G. Goldhaber, B. H. Hall, and G. H. Trilling (1968). Search for A K* Below Kπ Threshold. *Physical Review Letters* 21(3): 173-174. https://doi.org/10.1103/PhysRevLett.21.173

CDF Collaboration, (2009). Global search for new physics with *2.0 fb*$^{-1}$ at CDF. *Physical Review* D 79(1): 011101. https://doi.org/10.1103/PhysRevD.79.011101

CMS Collaboration(2012). Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters* B 716(1): 30-61. https://doi.org/10.1016/j.physletb.2012.08.021

Collins, J. H., K. Howe and B. Nachman (2019). Extending the search for new resonances with machine learning. *Physical Review D* 99(1): 014038. https://doi.org/10.1103/PhysRevD.99.014038

Collins, J., K. Howe and B. Nachman (2018). Anomaly Detection for Resonant New Physics with Machine Learning. *Physical Review Letters* 121(24): 241803. https://doi.org/10.1103/PhysRevLett.121.241803

Goldhaber, G., W. R. Butler, D. G. Coyne, B. H. Hall, J. N. MacNaughton and G. H. Trilling (1969). $\omega\rho^0$ Interference Effect in the $\pi^+p \rightarrow \pi^+ \pi^- \Delta^{++}$ Interaction. *Physical Review Letters* 23 (23): 1351-1354. https://doi.org/10.1103/PhysRevLett.23.1351