Algorithms under the Reign of Probability

In a 1930 issue of *Scientific Monthly*, US mathematician Warren Weaver proclaimed the twentieth century as the Reign of Probability. He argued that “it is nearly impossible to escape from this mathematical goddess of chance,” as probability tools were then being designed to quantify uncertainty across all areas of statistical research, even in slippery calculations of quantum particles. Weaver’s proclamation was a reflection of a growing embrace of new “probability tools” in statistical research and governance throughout interwar Europe, India, and the US. But it was also a future-looking declaration of an ascendant epistemological power over modern society. It was a promise that wide-ranging forms of knowledge would be subject to probability—its dominion stretching from the seemingly banal instruments of life insurance to the highest theoretical sciences. Probability would command all domains of data-driven inquiry. The interwar turn toward uncertainty management reverberated throughout science, industry, and government. It was a movement that produced powerful and long-lasting social infrastructures toward a world dictated by uncertainty. And we are currently experiencing this movement in revival under its big data iteration: algorithmic uncertainty.

**MAPPING UNCERTAINTY**

To understand the shifting powers of uncertainty management, we must map out its epistemological, technological, and political meanings. As a base definition, probability describes likelihoods of propositions and events, usually expressed as a percentage, where perfect certainty is 1 and uncertainty is < 1. Beyond this, probabilistic knowing is a commitment to greater analytic (laws, axioms, and definitions) and technological (computers and data systems) architectures needed to express limited information in terms of likelihoods. Under conditions of limited information, elaborate mathematical and social infrastructures are needed to sustain the probabilistic worldview.

My dissertation examines the significance of a dramatic inflection point in 1920s and 1930s-era uncertainty management, when new probability tools and infrastructures were designed to manage uncertainty in statistical research and governance. Studies of data uncertainty then persisted with drastically different applications throughout Cold War proxy wars. After 1980, computer scientists redesigned uncertainty models to work for big data, and since 2008, this research has ballooned, with over 1.3 million publications on algorithmic-driven uncertainty management. Algorithmic uncertainty constitutes a very different set of technologies and social politics from the interwar years, but there is nonetheless a clear lineage. Significantly, algorithmic uncertainty processes involve layered computer-implemented runs of probability tests and statistical estimations on computer-implemented sampling bodies.
A growing commitment to uncertainty management is resurfacing today, a newly reconfigured language of both truth and objectivity. It has overwritten earlier social doctrines—such as the mechanistic worldview—and stands at the exclusion of alternative possibilities. Today we test the validity of science through computer-simulated assessments of the statistical significance of its findings. The Intergovernmental Panel on Climate Change (IPCC) assesses the past, present, and future of climate knowledge in terms of uncertainty logics. And uncertainty-modeling practices are increasingly gaining power in finance innovation cultures. Through merging uncertainty practices with midcentury capital risk logics, “market makers” now promise to offload market risk accountability entirely to algorithms.

PREDIGITAL UNCERTAINTY

The 1920s and 1930s were a massive inflection point in uncertainty work that established a modernizing conception of data technologies and data-driven governance. What is significant here and what should be asked about algorithmic uncertainty is how uncertainty epistemology came to power and who benefits in a world purportedly dictated by probability. The popular adoption of interwar uncertainty tools was a response to an apparent loss of technocratic control over nineteenth-century modes of statistical governance.

The utility of statistical data was put to question as practitioners described both an overabundance and incompleteness to statistical information. Statistics was by then not a science of counting but of estimation, and error was an inherent component of this work. Anxious about a loss of public confidence in data-driven institutions, technocrats sought to command error in statistical estimation. In Weaver’s description: all knowledge was undergirded by an “insufficiently accurate and insufficiently extensive body of data” and needed probability tools to render that data into “probability data.” New probability tools were designed to delimit uncertainty in statistical research. These were based in translating common statistical research concepts—vagueness, error, and randomization—into the language of axiomatic probability theory.

The promissory power that probability tools could overcome and make useful tattered, porous, and nonexistent data was itself used to garner public confidence in much larger monetary and trade systems. By the start of World War II, statisticians widely proclaimed that probability tools had “improved” statistics, which were likewise invoked by administrators, trade commissioners, and bureaucrats. They were of interest and benefit to many as they promised control over preexisting and newly generated data bodies undergirding larger technical systems, and were operational across specialist domains of knowledge.

In efforts to further demystify uncertainty practices, I historicize two interwar-era probability tools: confidence intervals and the null hypothesis. These estimation architectures are now deeply embedded in algorithmic uncertainty processes and are invoked in ongoing tensions and negotiations between hypothetico-deductive and inferential methods used in data testing. Engaging the logical complexities undergirding how data is tested and interpreted clarifies the political processes by which uncertainty measures are translated into real-world decisions. This is a question of what kinds of conclusions our probability tools allow us to make.

ALGORITHMIC UNCERTAINTY AS CRITICAL REFLEXIVITY

After 1980, the field of uncertainty quantification galvanized in data-rich scientific and social scientific fields including physics, geography, and economics that were preoccupied with the impacts of informational error, vagueness, and ambiguity in their forecasting and predictions. Uncertainty could enter into the modeling process at any stage from defining model parameters to observational error. Researchers designed fuzzy mathematical methods, decision trees, and topologies to help them conceptualize and define uncertainties in their computer simulation processes (see, for example, P.A. Burrough’s work). Practitioners now embrace uncertainty as a defining characteristic of big data. Big data decision-making tools are designed to wrangle, tame, and reduce the fog of uncertainty.
Uncertainty processes are also a site for critical reflexivity, where interdisciplinary scholarship on the study of data and algorithms should meet. I invoke Mei-Po Kwan’s defining work on algorithmic-driven geographies. Kwan situates “uncertainty work” as a critical practice of paying “attention to the effects of algorithms on the content, reliability, and social implications [that] these algorithms help generate.” This sense of uncertainty work is part and parcel of the critical data studies (CDS) corpus, which has identified algorithmic bias, algorithmic oppression, and digital labor(s) in data processing. Historians and anthropologists in turn study how data processes stabilize larger processes of power such as modernism, capitalism and communism, and empire. Uncertainty is a site where historians can benefit from the theoretical work generated in the CDS corpus, just as historical perspective can enrich more recent studies.

**POLITICS OF CERTAINTY-MAKING**

Uncertainty politics constitute the translation of uncertainty to objectivity, truth, and certainty for political and economic interest. Creating certainty from uncertainty is not a new concept in the history and philosophy of science. Sociologist Susan Star shows how local contingencies in scientific and technological work are translated into a global certainty through justifying, diminishing, and erasing ad hoc actions in the name of established disciplines. In accord with this, the sociopolitical power of big data cannot be understood without its simultaneous reliance on uncertainty tools that also serve as a site of resistance.

Science, technology, and society scholars have shown how certainty is created through the strategic packaging and selective communication of uncertainty processes, tools, and calculations. For example, scientific certainty argumentation methods (SCAMs) are methods of hedging scientific uncertainty for benefit especially in legislative and judicial decision-making as in the cases of Big Tobacco. In these cases, confidence intervals and null-hypothesis tests were used to promote or diminish the significance of statistical information in political outcomes. On a global scale, philosopher of science Wendy Parker makes clear that uncertainty remains elusive and detrimental in current practices of climate change modeling. Parker maps out the epistemic and political dimensions of ensemble modeling and asks: Whose probabilities are at work in climate change policy? She intervenes with her own algorithmic countermeasures—ownership, justification, and robustness—to reinstate scientific and political accountability in uncertainty modeling processes.

Warren Weaver’s gendered depiction of probability as the “inescapable goddess” and, in his later book, as “Lady Luck” relinquishes accountability for those who wield power over probability tools and papers over the infrastructures that support it. By producing fuller maps of uncertainty models and practices before and after the advent of the digital computer and across various domains of knowledge, we can better understand the shifting political landscapes of certainty-making and their impacts.

**REFERENCES**


**ABOUT THE AUTHOR**

Theodora J. Dryer is a PhD candidate in the Department of History and Science Studies at the University of California, San Diego, who works on the history of predigital computing and transnational and cultural histories of economic and algorithmic modeling. Her dissertation focuses on the rise of twentieth-century algorithmic computing. Contact her at tdryer@ucsd.edu.