The Science of Science

Dashun Wang and Albert-László Barabási

Introduction

This is the submitted version.
The printed version[1] is published by Cambridge University Press in 2021.

[1] Wang, Dashun, and Albert-László Barabási. 2021. *The Science of Science*. Cambridge University Press. doi:10.1017/9781108610834.

Scientific revolutions are often driven by the invention of new instruments—the microscope, the telescope, genome sequencing—each of which have radically changed our ability to sense, measure and reason about the world. The latest instrument at our disposal? A windfall of digital data that traces the entirety of the scientific enterprise, helping us capture its inner workings at a remarkable level of detail and scale. Indeed, scientists today produce millions of research articles, preprints, grant proposals, and patents each year, leaving detailed fingerprints of the work we admire and how they come about. Access to this data is catalyzing the emergence of a new multidisciplinary field, called *science of science*, which, by helping us to understand in a quantitative fashion the evolution of science, has the potential to unlock enormous scientific, technological, and educational value.

The increasing availability of all this data has created an unprecedented opportunity to explore scientific production and reward. Parallel developments in data science, network science, and artificial intelligence offer us powerful tools and techniques to make sense of these millions of data points. Together, they tell a complex yet insightful story about how scientific careers unfold, how collaborations contribute to discovery, and how scientific progress emerges through a combination of multiple interconnected factors. These opportunities—and challenges that come with them—have fueled the emergence of a new multidisciplinary community of scientists that are united by their goals of understanding science. These practitioners of science of science use the scientific methods to study themselves, examine projects that work as well as those that fail, quantify the patterns that characterize discovery and invention, and offer lessons to improve science as a whole. In this book, we aim to introduce this burgeoning field—its rich historical context, exciting recent developments, and promising future applications.

We had three core audiences in mind as we wrote this book. The primary audience includes any scientist or student curious about the mechanisms that govern our passion, science. One of the founding fathers of the science of science, Thomas Kuhn, a physicist turned a philosopher, triggered world-wide interest in the study of science back to 1962 with the publication of *The Structure of Scientific Revolutions*. Kuhn's notion of "paradigm shift" today is used in almost every creative activity, and continues to dominate the way we think about the emergence and acceptance of new ideas in science. In many ways, the science of science represents the next major milestone in this line of thinking, addressing a series of questions that are dear to the heart of every scientist but may well lay outside of the Kuhnian worldview: When do scientists do their best work? What is the lifecycle of scientific creativity? Are there signals for when a

scientific hit will occur in a career? Which kinds of collaboration triumph and which are destined to disasters? How can young researchers maximize their odds of success? For any working scientist, this book can be a tool, providing data-driven insight into the inner workings of science, and helping them navigate the institutional and scholarly landscape in order to better their career.

A broader impact of the science of science lies in its implications for policy. Hence, this book may be beneficial to academic administrators, who can use science of science to inform evidence-based decision making. From department chairs to deans to Vice Presidents of Research, university administrators face important personnel and investment decisions as they try to implement and direct strategic research. While they are often aware of a profusion of empirical evidence on this subject, they lack cohesive summaries that would allow them to extract signals from potential noise. As such, this book may offer the knowledge and the data to help them better take advantage of useful insights the science of science community has to offer. What does an *H* index of 25 tell us about a physics faculty member seeking tenure? What would the department most benefit from: a junior vs. a senior hire? When should we invest into hiring a superstar, and what can we expect her impact will be?

We also hope that program directors with NSF, NIH, and other public and private funding agencies will find the book useful for supporting high-performing individuals and teams to best address science's emerging challenges. Many civilian and military government agencies, nonprofits, and private foundations are already collecting data and developing tools rooted in science of science. The framework offered in the coming chapters will allow them to utilize this data in a way that best serves their own purposes, helping them set up more effective funding mechanisms, and ultimately benefitting both science and society.

The changing landscape of science also affects scholarly publishers, who often compete to publish articles that will impact the direction and the rate of future scientific progress. We hope journal editors will also find science of science useful for a range of practical purposes—from understanding the natural life-cycle of a discovery's impact to identifying hit ideas before they become hits—which may, in turn, augment the impact of what they publish.

Lastly, this book is intended for scientists who are currently involved in science of science research, or for those who wish to enter this exciting field. It is our aim to offer the first coherent overview of the key ideas that currently capture the discipline's practitioners. Such an overview is necessary, we believe, precisely because our community is highly interdisciplinary. Indeed, key advances in the science of science

have been generated by researchers in fields ranging from the information and library sciences to the social, physical, and biological sciences to engineering and design. As such, approaches and perspectives vary, and researchers often publish their results in venues with non-overlapping readership. Consequently, research on the science of science can be fragmented, often along disciplinary boundaries. Such boundaries encourage jargon, parochial terms, and local values. In the book we aimed to summarize and translate the insights from highly diverse disciplines, presenting them to students and researchers cohesively and comprehensively. We will not only emphasize the common intellectual heritage of the diverse set of ideas that coexist in the field, but also provide approaches to orient new research. Thus, we hope the book will be a resource for interested students and researchers just discovering the field.

The book is structured in four parts: *The science of career* focuses on the career path of individual scientists, asking when we do our best work and what distinguish us from one another. *The science of collaboration* explores the advantages and pitfalls of team work, from how to assemble a successful team to who gets the credit for the team's work. *The science of impact* explores the fundamental dynamics underlying scientific ideas and their impacts. Finally, the *Outlook* part summarizes some of the hottest frontiers, from the role of AI to bias and causality. Each part begins with its own introduction which illuminates the main theme using questions and anecdotes. These questions are then addressed in separate chapters that cover the science relevant to each.

By analyzing large-scale data on the prevailing production and reward systems in science, and identifying universal and domain-specific patterns, science of science not only offers novel insights into the nature of our discipline, it also has the potential to meaningfully improve our work. With a deeper understanding of the precursors of impactful science, it will be possible to develop systems and policies that more reliably improve the odds of success for each scientist and science investment, thus enhancing the prospects of science as a whole.