Forecasting innovations in science, technology, and education

Katy Börner[a,b,1], William B. Rouse[c], Paul Trunfio[d,e], and H. Eugene Stanley[d,e]

Human survival depends on our ability to predict future outcomes so that we can make informed decisions. Human cognition and perception are optimized for local, short-term decision-making, such as deciding when to fight or flight, whom to mate, or what to eat. For more elaborate decisions (e.g., when to harvest, when to go to war or not, and whom to marry), people used to consult oracles—prophetic predictions of the future inspired by the gods. Over time, oracles were replaced by models of the structure and dynamics of natural, technological, and social systems. In the 21st century, computational models and visualizations of model results inform much of our decision-making: near real-time weather forecasts help us decide when to take an umbrella, plant, or harvest; where to ground airplanes; or when to evacuate inhabitants in the path of a hurricane, tornado, or flood (1). Long-term weather and climate forecasts predict a future with increasing torrential rains, stronger winds, and more frequent drought, landslides, and forest fires as well as rising sea levels, enabling decision makers to prepare for these changes by building dikes, moving cities and roads, and building larger water reservoirs and better storm sewers (2).

Power of Forecasts

Computational models are particularly useful if they are combined with high-quality data and if they are widely used and understood. As early as 1960, Buckminster Fuller proposed the "World Game" to address the world's problems through a holistic and anticipatory systems approach (3, 4). The game used Fuller's Dymaxion Map to visualize resources, trends, and scenarios. It was meant to be accessible to everyone (not just experts); therefore, decisions could be made collectively, and results could be used by anyone. In the 1970s, "The limits to growth: A report to the club of Rome" (5) used simulations to forecast future states of the world; a 30-y update was published in 2004 (6). Today, computers and the internet provide the technological infrastructure for universal access to data and compute power. Advances in mathematics, computer science, engineering, and other disciplines have made it possible to implement scalable, empirically validated computational models that render data into actionable insights. Because datasets are huge and multidimensional and models are complex, both tend to exceed human comprehension. Data visualizations and novel interfaces are being developed to help communicate the inner workings of computational models as well as model results to diverse stakeholder groups, including public audiences. For example, the *Places & Spaces: Mapping Science* exhibit (scimaps.org) features more than 100 large-scale maps of science and 18 interactive data visualizations designed by more than 230 authors. Elsewhere, the Data & Network Science in K-20 Education initiative (www.bu.edu/networks) gives students an entry point to understand and make meaning of science and technology (S&T) network models, data visualizations, and the role of forecasting.

Diverse Forecasts for Different Stakeholders

Most decision makers prefer orderly, predictable conditions; little disruption; and sufficient resources (e.g., money, talent, compassion) to pursue desirable futures. Many realize that, in the knowledge age, scientific progress, technological innovation, and affordable high-quality education are of central importance to the success of individuals, regions, and nations. Hence, decision makers have a deep interest in—and are willing to pay for—easy to use, near real-time access to data and models that help them make sense of, communicate with, and proactively manage science, technology, and education. Global operation rooms that provide visualizations of current data and predictions of possible futures are already commonplace in meteorology, finance, epidemiology, and defense. Science, technology, and education "observatories" for experts and novice users are actively being researched and developed.

[a]School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN 47408; [b]Educational Technology/Media Centre, Dresden University of Technology, 01062 Dresden, Germany; [c]Center for Complex Systems and Enterprises, Stevens Institute of Technology, Hoboken, NJ 07030; [d]Center for Polymer Studies, Boston University, Boston, MA 02215; and [e]Physics Department, Boston University, Boston, MA 02215

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "Modeling and Visualizing Science and Technology Developments," held December 4–5, 2017, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and video recordings of most presentations are available on the NAS website at www.nasonline.org/modeling_and_visualizing.

Author contributions: K.B., W.B.R., P.T., and H.E.S. designed research, performed research, and wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. Email: katy@indiana.edu.

Published online December 10, 2018.

The Sackler Colloquium "Modeling and Visualization of Science and Technology Developments" held December 4 and 5, 2017 brought together power users, designers, and early adopters of this new type of observatory. The papers in this special issue review existing research and introduce recent developments regarding computational models and visualizations for use in academia, government, and industry, with a special focus on innovation in science, technology, and education.

Academia: Understanding Science. Recent reviews of studying science using the scientific approach, also called "science of science" (7, 8), and policy forum vignettes (9) provide an overview of existing methods, tools, and insights. Prevailing metrics and models use large-scale datasets (e.g., publications, patents, funding, clinical trials, stock market, social media data) to simulate the structure and evolution of the S&T landscape. They aim to quantify and predict scientific research, impact, and outcomes; help support the selection of candidate faculty members by universities; identify the best reviewers; prioritize the development of research fields in which a country should invest; or evaluate the scientific impact of scholars, journals, academic institutions, or nations. There exist models of the structure and dynamics of the whole S&T system at diverse geographic and temporal scales (8, 10, 11) as well as many different ways to communicate the structure and dynamics of S&T (12, 13). As new datasets become available, such as the Institute for Research on Innovation and Science (IRIS) database (https://iris.isr.umich.edu) that tracks university-sponsored project expenditures at the transaction level for 26 research universities over the period from 2001 to the present (14), new models can be designed and validated [e.g., models on scientific productivity (15), economic impact (16), or workforce development (17)].

Government: Data-Driven Policy Making. "Report of the Commission on Evidence-Based Policymaking" from 2017 (18) opens with this statement: "The American people want a government that functions efficiently and responsibly addresses the problems that face this country. Policymakers must have good information on which to base their decisions about improving the viability and effectiveness of government programs and policies. Today, too little evidence is produced to meet this need." The report argues for a future in which "rigorous evidence is created efficiently, as a routine part of government operations, and used to construct effective public policy."

Improved access to data will make it possible to improve the quantity and the quality of evidence that informs important program and policy decisions—without decreasing data security or significantly increasing privacy risks. At the colloquium, Azer Bestavros presented "Sharing knowledge without sharing data" (19), showing how cryptographic approaches can be used to facilitate secure multiparty computation in accessible and scalable ways. The privacy-preserving approach has been used to assess and address economic inequalities (20) and to perform analytics in health care (21).

Existing models of S&T have been used to simulate the impact of population explosion and aging (22), alternative funding schemas (23), or the probable outcomes of different policy decisions (24) while being fully aware of the limits of predictive models (25).

In 2011, Helbing and Balietti (26) led a 10 billion Euros proposal effort to create an "Innovation Accelerator" meant to "identify new ways of publishing, evaluating, and reporting scientific progress; promote ICT (information and communications technology) solutions to increase the awareness of new emerging trends; invent tools to enhance Europe's innovation potential; develop new strategies to support a sustainable technological development; and lay the foundations for new ways to reach societal benefits and respond to industrial needs using ICT" (26, 27). While the proposal was not funded, it presented a possible blueprint and showcased the value of an infrastructure that supports evidence-based decision-making in government.

Industry: Predicting Innovation. Invention is the creation of a new process or device, and innovation is the creation of change in the marketplace. Innovation may or may not rely on one or more inventions, and technology has been adopted when it has transitioned from invention to innovation. Patented inventions and high-impact scientific works frequently build on unconventional combinations of existing knowledge (28, 29). Decision makers in industry must determine how to utilize limited resources to increase innovation, labor productivity, inventory turnover, and asset utilization (30). Research collaboration and workforce development decisions require knowing where the most productive research is being done and the best experts are trained as well as how that production has changed over time and across individuals and institutions (31). Such knowledge will foster and allow for better strategic planning, hiring, and resource allocation.

All Together Now: Population Health. Government, academia, industry, and the public are involved in efforts that aim to improve the health of an entire human population. Health challenges are due to urbanization, epidemiologic shifts, aging, and climate change. With more than one-half of the population living in urban environments, there is an increase in density, diversity, complexity, and inequality. Today, people's ZIP codes are a better predictor of health than their genetic code (32) (Fig. 1).

Global maps of health monitor, predict, visualize, and communicate health. An example is Predict: HealthMap (https://www.healthmap.org/predict) created by USAID's Emerging Pandemic Threats program (33), which aims to increase capacity in the developing world for early detection of viruses from wildlife with pandemic potential.

Health observatories could provide access to data, research results and funding, and new cures. Analogous to weather forecasts, they would analyze and visualize data to identify outliers and trends and broadcast "Health News" to communicate key developments to expert and lay audiences. They would make it possible to quantify and make visible the relationships between scientific discoveries and health advances, increasing public understanding and funding support (34).

Envisioning and Implementing Desirable Futures

Analytical and predictive computational data models can help us understand the past and present and predict and implement desirable futures. While many of us use data models on a daily basis (e.g., online games or shopping recommendations), few truly understand how these models work, and even fewer are able to use models to answer new questions.

The Sackler Colloquium brought together academic, industry, and government experts from more than 30 different disciplines. Resulting papers are grouped here in four major development trusts, and key insights are discussed to show their contribution to an evolving understanding of innovations in science, technology, and education that is larger than the sum of parts.

Integrated Systems Models. In the industrial age, task specialization, respect for authority, and predictability were important,

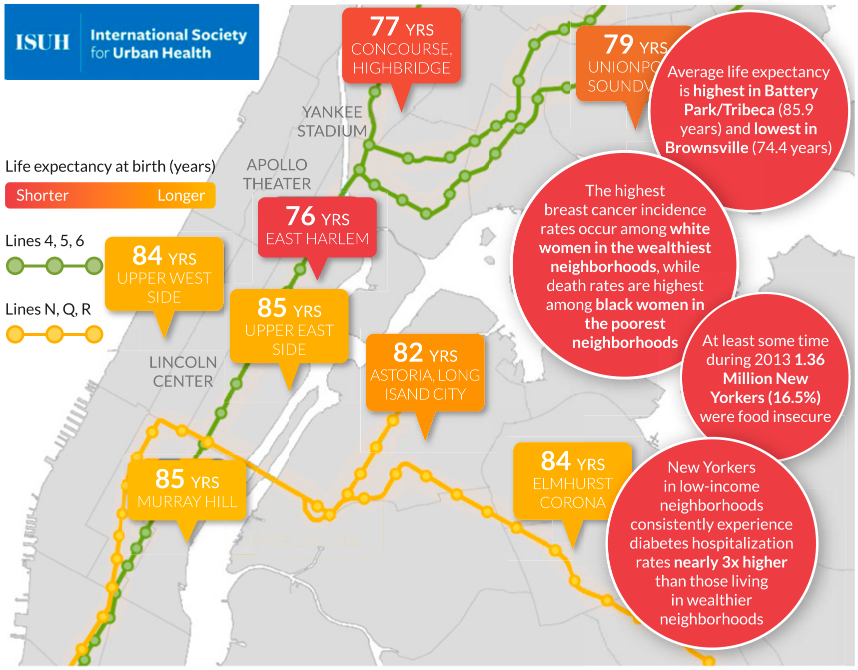

Fig. 1. Short distances to large gaps in health in New York City. Image courtesy of the International Society for Health at the New York Academy of Medicine.

and divide-and-conquer strategies widely applied. In the innovation age, agility, team work, and thriving in uncertainty are more valuable, and a holistic, system science approach to problem solving is beneficial. However, most ecosystems are fragmented (i.e., the organization of production and service delivery across different stages of production and service are provided, managed, and governed by different, independent, and often geographically dispersed organizational entities) (Fig. 2, *Left*). This fragmentation creates barriers to developing holistic solutions and even greater barriers to implementing them. Examples are health care, where critical processes are managed by providers but also by payers and regulators at local, state, and federal levels; education, with responsibilities distributed over local control motivated and constrained by state and federal funding; or government, with a separation of powers at local, state, and federal levels. In all three organizational systems, "workflows become a complex series of handoffs between functions, jobs, and information systems. Each handoff represents an opportunity to introduce error, delay, and added cost. When organizations become fragmented, it requires more work to deliver value to the customer and the ability of the organization to adapt to environmental changes is diminished. In extreme cases, the loss of value is deadly and (organizations) go extinct" (35).

Fig. 2, *Right* introduces an integrated system that supports innovations across different levels involving services, processes, capacities, consumables, and information and that facilitates modular changes (i.e., "plug and play") without major disruption. The three

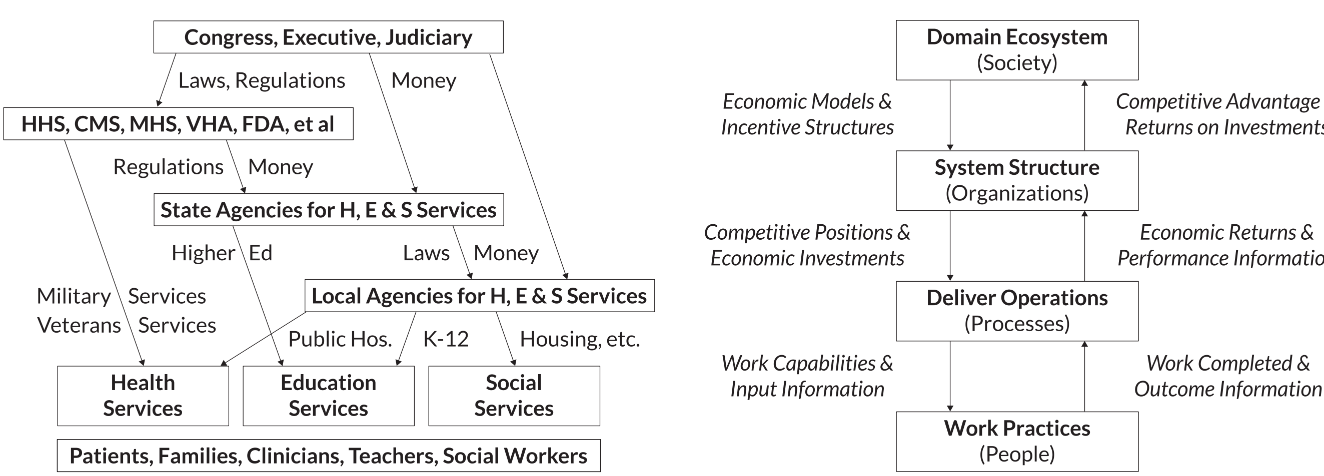

Fig. 2. Fragmented system (*Left*) and integrated organizational system (*Right*). CMS, Centers for Medicare and Medicaid Services; E, Education Services; FDA, Food and Drug Administration; H, Health Services; HHS, Department of Health and Human Services; MHS, Military Health System; S, Social Services; VHA, Veterans Health Administration.

system models featured in this special issue each present an integrated approach in support of evidence-based policy (36).

"Modeling research universities: Predicting probable futures of public vs. private and large vs. small research universities" by Rouse et al. (37) extends the computational model of research universities presented in ref. 38 using three strategic scenarios: (*i*) status quo, (*ii*) steady decline in foreign graduate student enrollments, and (*iii*) downward tuition pressures from high-quality online professional master's programs. Robust data are used to project four types of research universities (large public and private and small public and private) into the future. Computations show that, while research requires high subsidies, it serves to create reputation; research investments create "brand value" that can be converted into tuition income. The model was validated by applying it to different types of institutions and inviting feedback from over 20 domain experts. Model results predict the rise and decline of institutions and suggest possible revisions in business strategies (e.g., restricting research activities to avoid the inherent subsidies that these activities require) to address competitive forces.

"Twin-Win Model: A human-centered approach to research success" by Shneiderman (39) expands on ref. 40 and presents a model that encourages teams of researchers, academic leaders, business managers, and government funding policy makers to embrace a problem-oriented approach to research. It argues that teams should aim to pursue "breakthrough theories in published papers and validated solutions that are ready for widespread dissemination" simultaneously to increase the number of foundational discoveries and to speed up the translation of innovations into practice. The work shows that working on real-world problems with partners deeply invested in the solutions accelerates (and helps fund) both applied and basic research. Evidence is provided by means of citation analysis that compares six US public universities with six US private universities for the years 2012–2016, all of which show increased citation impact for papers that list authors from off-campus partners from business, government, and nongovernmental organizations. Researcher interviews provide deeper insights as to why such collaborations can advance basic and applied research.

"Vision for a systems architecture to integrate and transform population health" by Madhavan et al. (41) presents a perspective combining expertise from the National Academies of Sciences, Engineering, and Medicine with that of academia and industry to propose a visionary model for harmonizing programs, policies, regulations, legal arrangements, and practices to understand and improve the state of population health. As the paper notes, everyone is involved in population health, but no one is in charge of it. Hence, the specific focus is on a system architecture for real-time "situation awareness" that uses rich global data, a suite of computational models, and visualizations (providing different views, alerts, and scenarios) to improve proactive planning, monitoring, exploration, and decision support. The paper also reflects on the necessary changes in education, research, and joint action in support of greater coordination and better synergies of population health efforts.

Models of Academic Networks, Impact, and Awards. Two papers aim to quantify the impact of scientific apprenticeship, mentorship, and coauthorship on scientific excellence (*Academia: Understanding Science*).

"The chaperone effect in scientific publishing" by Sekara et al. (42) shows the critical importance of acquiring the expertise needed to publish in prestigious journals. It studies the impact of experienced authors on teaching young scientists to ask "the right question" and to acquire high-level scientific communication skills. The paper defines and quantifies this "chaperone effect" by computing how scientists transition into senior status given multiple publications within the same journal; shows that the effect is stronger in medical and biological sciences and weaker in natural sciences, with effect sizes growing over the last decade; and discusses implications on long-term citation patterns of papers, with chaperoned authors tending to have higher long-term impact than nonchaperoned authors. Their findings shed light on the role played by experience and skills required to publish in prestigious venues.

"Scientific prize network predicts who pushes the boundaries of science" by Ma and Uzzi (43) shows the critical link between the worldwide and transdisciplinary scientific prize network and the dynamics of reward stratification and prizewinning in science. The study uses original data collected on 3,000 different scientific prizes in diverse disciplines and the career histories of 10,455 worldwide prizewinners covering more than 100 years of science. Their work uncovers (*i*) the relatively small and densely clustered number of ideas and scholars who lead scientific thinking (e.g., 64.1% of prizewinners have won two prizes, and 13.7% have won five or more prizes); (*ii*) the interlocks among different prizes within and between disciplines, which are formed by multiple prizes being won by the same scientist whose ideas then gain credit and spread through the prize network; and (*iii*) the genealogical and coauthorship networks that predict who wins multiple prizes. Whereas scientific prizes were once mainly thought to be measures only of personal acclaim, they can now be recognized as performing multiple functions in science regarding the legitimation (44), spread, and stratification of ideas and having a network structure that reveals the "high level of interconnectedness among acclaimed scientists and their path breaking ideas."

Models of Job Market Needs and Educational Offerings → Training the Workforce of Tomorrow. There is no progress in science and no research and development (R&D) innovation without education. Most of today's jobs require at least a high school diploma and in many cases, a college degree or higher education. Workers change jobs frequently and need to upskill continuously—particularly in science, technology, engineering, and math (STEM) fields. Training the workforce of the future requires a deep understanding of how people learn (45).

In the United States, higher education is a major investment that a decreasing number of students can afford. While much data exist for other types of major long-term expenses and commitments (e.g., purchasing a home or a car), there is a major debate about the true value of a college education. Some equate value with the earning power that comes with a particular college degree; others value job satisfaction and/or the social skills and networks or intellectual rewards gained from the college experience. Many stakeholders are interested in ensuring that universities remain key creators of intellectual capital and economic growth (46) while competition among institutions grows.

Industry is concerned about the continuous high-quality training needed to keep engineers and others up to date when "products and processes are constantly changing due to technology, innovation, economic factors, and the encompassing influences of society and culture" (47). Learning scientists from Microsoft Corporation and The Boeing Company presented at the Sackler Colloquium, showcasing the urgent need for well-trained employees and efficient workforce development. STEM industries in general are fiercely competing for the best and brightest—offering high salaries, flexible work time, and much freedom. The aerospace industry and NASA have a disproportionately large percentage of workers aged 50 years old and older compared with the national average (48), and up to one-half of the current workforce will be eligible for retirement within the coming 5 years. Hence, many corporations have developed extensive in-house training programs or are partnering with institutions of higher education. Personalized, contextualized (just-in-time) learning tailored to a student's individual needs is the ultimate goal (49). For example, The Boeing Company supports the National Academy of Engineering Grand Challenge of Advancing Personalized Learning through competency-based learning, leveraging online learner analytics to develop metrics and approaches that support efficient experiential learning by doing that is social (i.e., facilitated by peers, mentors, coaches) and has a measurable return on investment (50).

The rise of artificial intelligence will lead to the displacement of millions of blue collar as well as white collar jobs in the coming decade (51–54). Jobs that require capabilities that are easy to automate (e.g., recognizing and analyzing pictures, voice, and video; parsing, translating, and generating natural language; driving vehicles within well-defined environments; or retrieving and summarizing all data on a topic) will soon be fully automated, while essentially human capabilities (e.g., taking on responsibility for unforeseen tasks, having consciousness and being capable of reflection, having feelings) might not be realized in the foreseeable future (55).

Research on "convergent technologies" argues that the synergistic combination of nanoscience and nanotechnology, biotechnology and biomedicine, information technology, and cognitive and neuroscience will make it possible to substantially augment human mental, physical, and social abilities (56). Exemplarily, the bottom level of Fig. 3 depicts the technologies on which we have come to depend. These are rapidly evolving, providing new, more powerful capabilities. The middle level depicts how these technologies are exploited in the context of research, design, manufacturing, production, and supply chains. The top level depicts the range of users, including people who create and communicate new knowledge and technologies; people who use knowledge and technologies to design new products and services; and people who operate, maintain, and manage the resulting capabilities. All of these levels and users are becoming increasingly integrated and interdependent.

Those humans who remain in the workforce will soon have AI-supported robotic coworkers (Fig. 3). They will become part of the internet-of-things via wearable outfits, virtual reality glasses, robotic companions, and smart environments. Given this outlook, how do we best train today's students for future jobs in academia, industry, or government?

"Changing demographics of scientific careers: The rise of the temporary workforce" by Milojević et al. (57) models academic success, defined as the ability to maintain a long, active career in science. Using bibliographic data for astronomy, ecology, and robotics extracted from the full Clarivate Analytics Web of Science database spanning the years 1900–2015, the authors find a dramatic shortening of careers of scientists across all disciplines studied. The half-life of a cohort went from 35 years for scientists who started their careers in the 1960s to only 5 years in the 2010s, reflecting the rise of the phenomenon the authors call the "expendable scientist." In addition, they show a rapid rise (from 25 to 60% since 1960s) of scientists who spend their entire career as supporting authors rather than lead authors. Cohort attrition is successfully modeled by a hazard probability function; however, neither an author's early productivity or citation impact nor the level of initial collaboration reliably predict ultimate survivability.

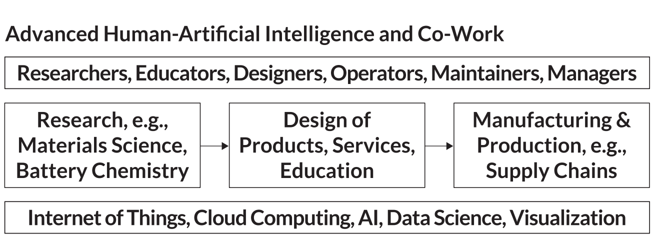

Fig. 3. Advanced human–artificial intelligence and cowork.

"How science and technology developments impact employment and education" by Martinez (58) details how the Bureau of Labor Statistics uses data sources from the US federal government to publish 10-y employment projections that are widely used by students, jobseekers, and policy makers. Since 1960, the projections have been computed every 2 years—the October 2017 projection covers over 300 different industries and 800 occupations for the period 2016–2026. The perspectives paper "argues for a better understanding of how developments in science and technology influence the creation of new occupations and how subsequent changes in educational programs can help decision makers at all levels of our society…it discusses several data sources that might help us to explore the relationship between advancements in industry, emerging occupations, and educational changes over time" (58).

"Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy" by Börner et al. (59) explores the gap between S&T developments, educational offerings, and job market needs. Using millions of publications, course syllabi, and job advertisements published between 2010 and 2016, they present visualizations and causal models that quantify the dynamic skill (mis-)alignment between academic push, industry pull, and educational offerings; reveal the increasing importance and demand for uniquely human skills, such as communication, negotiation, and persuasion, in a data-driven economy; and present results from a survey that asked 20 labor market and academic experts to examine the readability of the visualizations. Results show a substantial gap between the centrality of uniquely human "soft" skills for technical jobs and their relative peripherality to technical coursework and research publications.

Importance of Intangibles → Ecosystems Go Beyond Institutions and Laboratories. Haskel and Westlake (60), in their book *Capitalism Without Capital: The Rise of the Intangible Economy*, argue for the importance of intangible assets (e.g., skills, design, branding, research, or software) for understanding the rise and fall of corporations and major economic change. While most economists still count and report tangible assets (e.g., land, buildings, machinery, and computers), intangible assets are what drives economic progress today.

Understanding the world in which we live as a network of networks is beneficial. Batty (61) was among the first to argue that, to understand cities, we must model them not just as places in space but as systems of networks and flows. He and others combined theory and methods from network science, social physics, transportation theory, urban geography, and urban economics to reveal how cities function and inform decision-making.

"Opportunities to observe and measure intangible inputs to innovation: Definitions, operationalization, and examples" by Keller et al. (62) aims to quantify the value of intangibles using administrative data and repositories captured on web pages. They strongly argue for the development of accurate and repeatable measures to estimate the value of a company's ownership of databases and software, the tacit knowledge of their workers, and the investments in R&D and design. By means of two case studies, they exemplify processes to discover, acquire, profile, clean, link, and explore the fitness for use and methods to statistically analyze and visualize data about intangible assets, demonstrating the feasibility of different approaches.

"The role of industry-specific, occupation-specific, and location-specific knowledge in the growth and survival of new firms" by Jara-Figueroa et al. (63) introduces new measures of knowledge based on the work history of individuals to understand the impact of different forms of tacit knowledge on industrial diversification and growth. The authors look at the growth and survival rates of pioneer firms, which are the first firms to operate in an industry that was not present in a region. They find that the growth and survival of pioneer firms increase significantly when their first hires are workers with experience in a related industry (but not with experience in a related occupation) in the same geolocation. To address endogeneity concerns, the authors use Bartik instruments, which leverage national fluctuations in the demand for an activity, as shocks for local labor supply. This instrumental variable supports the finding that industry-related knowledge is a predictor of the survival and growth of pioneer firms. Results can help decision makers understand the micromechanisms that drive regional economic diversification and entrepreneurial success.

"Macroscopic dynamics and the collapse of urban traffic" by Olmos et al. (64) aims to inform planning and infrastructure interventions. Many workers—scholars included—endure long commute times (and long-distance relationships). Traffic routes affect commute times, which affect location and employment decisions. Using traffic data from multiple cities, this work studies changes in the travel time of individual drivers under various conditions of demand. Keeping road capacity and travel origins and destinations unchanged, the model is used to study the effects of increasing the volume of cars in the traffic network. Three general states are identified that are separated by two phase transitions: the appearance of bottlenecks and collapse of the system.

Actionable Models

The economic and political (in some cases, violent) turmoil of recent years adds an urgency to the development of models and visualizations that will help humans make informed decisions. In *2052: A Global Forecast for the Next Forty Years*, Randers (2) argues that human response is too slow to adapt humanity to the limits of planet Earth and "catastrophic overshoot and collapse" or "well-managed peak and decline" are more likely. Part of the problem is in the fact that, even with easy access to high-quality modeling results, making the right decision and implementing the right action are rather challenging for human beings as follows.

Humans often act based on emotions, not evidence (e.g., individuals will vote against their own economic interests if there is a strong emotional appeal made to a contentious cultural issue).

Most of our organizational systems favor short-term return on investment. The nature of the voter—expecting results within 4 y, not decades or hundreds of years—makes long-term investments nearly impossible. There is no effective spokesperson for those who are unborn.

Most people and nations react (i.e., action is often taken only after disaster has struck and lives have been lost) instead of acting in a systematic, proactive manner based on rich global data.

However, humans now do have a global impact on Earth, and connectivity has replaced the old divide-and-conquer approach as the new paradigm of global organization (65). The global environment, transportation, health, and other systems are best understood using a network science approach; networks and activity patterns are used to redraw existing maps. For example, data on commuter trajectories help redraw state boundaries, suggesting that new kinds of geographic categories are necessary if we wish to accurately describe the functional network of flows and relationships that shape our lives in the modern world (66) (Fig. 4).

Highways, railways, and air traffic routes transport people and other tangibles, and pipes supply water, energy, air, and remove waste; radio waves and internet cables diffuse intangibles, such as ideas, news, and innovations. Geopolitical competition is changing from war over territory to competition over consecutiveness (e.g., global supply chains; energy markets; hubs for traffic, trade, finance, innovation, talent) (65). Trade, manufacturing, education, and politics are all global. We need to understand this system of systems, learn to manage and use feedback cycles, and identify and implement pathways to collective resilience.

Computational models and associated visualizations make it possible to externalize, run, and compare different mental models of how biological, technological, and/or social systems might function. Many of us hold bits and pieces of the overall puzzle in our heads and hands, and these were developed using different questions, datasets, and methods and are relevant for different levels of abstraction. Combining these different views into a more holistic world view—one that is greater than the sum of its parts—is challenging.

"The limits to growth: A report to the club of Rome" (5) and the Google Flu Trends predictive model that went bad (67) tell important lessons: caution is required in an age when computational models are being used in disgraceful ways, producing bias and harm as described by Cathy O'Neil (68) in *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*.

For computational experiments and models to be useful, they must be properly designed and well documented (69). The ability to reproduce experiments, rerun computation models, or reexamine model results to get the same insight is one of the hallmarks of science. However, the different cultures and approaches used by those that develop models—mathematicians, physicists, computer scientists, economists, and social scientists to name just a few—make it hard to agree on a shared language to represent models or a uniform process to share those models. Harris (70) in *Rigor Mortis: How Sloppy Science Creates Worthless Cures, Crushes Hope, and Wastes Billions* elaborates the sources and implications of the reproducibility crisis in research.

Colloquium speakers and authors of this special issue were strongly encouraged to develop and present models that answer well-defined research questions (e.g., make sure that optimality criteria are stated clearly) or real-world questions that matter; collaborate closely with industry, government agencies, or other academics to gain access to relevant datasets; consult and conform to the manifesto for reproducible science (71); and obey

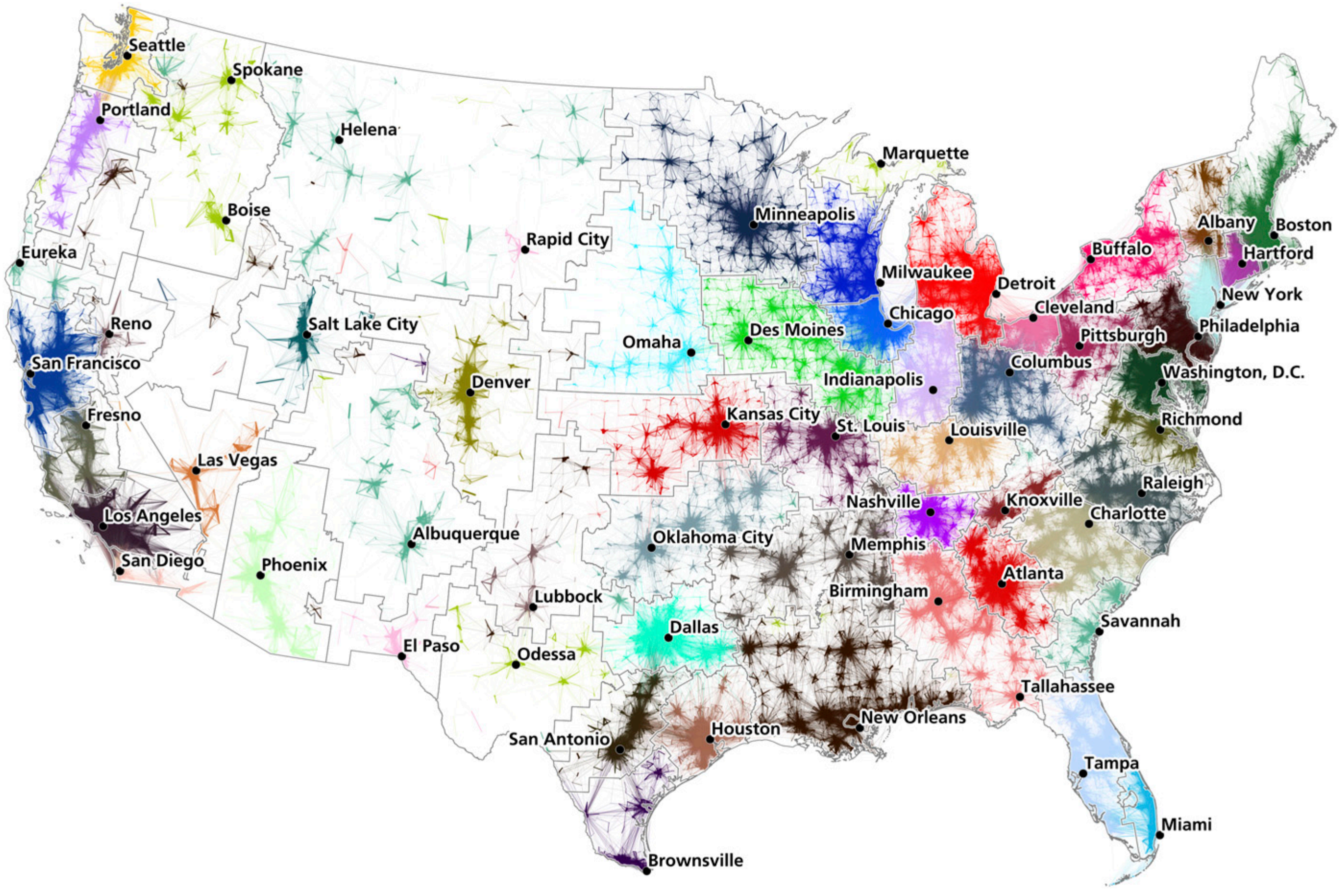

Fig. 4. Megaregions of the United States with state boundaries redrawn based on commuter patterns.

the 10 rules for credible practice of computational modeling detailed in ref. 72. It is our hope that future modeling and visualization efforts will aim to follow and advance those guidelines and rules.

Open code and data, visualizations of models that can be understood and explored by many, and stories that highlight key insights of different scenarios all provide an exciting opportunity to combine existing models from different areas of research (e.g., physics, economics, social sciences, information sciences). They have the power to inspire novel collaborations that overcome scientific boundaries and synergistically combine expertise and results from different domains.

Ultimately, it will be important that all humans be able to read and develop their own computational models and associated visualizations. In fact, in a world filled with data, data literacy is becoming as important as textual and math literacy. The larger the number of people who can make data-driven global decisions, the better futures we can implement together.

Acknowledgments

We thank the Sackler Colloquium speakers for their unique contributions and Albert-László Barabási, César A. Hidalgo, Sallie Keller, Guru Madhavan, Staša Milojević, Ben Shneiderman, and Brian Uzzi for their expert comments on an earlier version of this paper. We also thank Alasdair Rae and Garrett D. Nelson for a custom rendering of Fig. 4. We appreciate support of figure design by Leonard E. Cross. The Boston University work is supported by NSF Grants PHY-1505000 and CHE-1213217 and Defense Threat Reduction Agency Grant HDTRA1-14-1-0017. The work of K.B. was partially supported by NIH Awards U01CA198934 and OT2OD026671 and NSF Awards 1546824, 1713567, 1735095, and 1839167. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

1 Monmonier M (1999) *Air Apparent: How Meteorologists Learned to Map, Predict, and Dramatize Weather* (Univ of Chicago Press, Chicago).
2 Randers J (2012) *2052: A Global Forecast for the Next Forty Years* (Chelsea Green, Hartfort, VT).
3 Fuller B (1969) *Operating Manual for Spaceship Earth* (Southern Illinois Univ Press, Carbondale, IL).
4 Fuller B (1973) *Earth, Inc* (Anchor, Garden City, NY).
5 Meadows DH, Meadows DL, Randers J, Behrens WW, III (1972) The limits to growth: A report to the club of Rome. Available at www.ask-force.org/web/Global-Warming/Meadows-Limits-to-Growth-Short-1972.pdf. Accessed October 29, 2018.
6 Meadows D, Randers J (2004) *The Limits to Growth: The 30-Year Update* (Routledge, London).
7 Zeng A, et al. (2017) The science of science: From the perspective of complex systems. *Phys Rep* 714–715:1–73.
8 Fortunato S, et al. (2018) Science of science. *Science* 359:eaao0185.
9 Azoulay P, et al. (2018) Toward a more scientific science. *Science* 361:1194–1197.
10 Scharnhorst A, Börner K, van den Besselaar P (2012) *Models of Science Dynamics: Encounters Between Complexity Theory and Information Sciences* (Springer Science & Business Media, Berlin).

11 Milojević S, Börner K (2016) Modeling science, technology and innovation: NSF conference report. Available at https://modsti.cns.iu.edu/wp-content/uploads/2016/09/ModSTI-Conference-Report.pdf. Accessed October 29, 2018.
12 Börner K (2010) *Atlas of Science* (MIT Press, Cambridge, MA).
13 Börner K (2015) *Atlas of Knowledge: Anyone Can Map* (MIT Press, Cambridge, MA).
14 The Institute for Research on Innovation & Science (IRIS) (2017) *Full Documentation for UMETRICS 2016Q3a Dataset* (IRIS, Ann Arbor, MI).
15 Zolas N, et al. (2015) Wrapping it up in a person: Examining employment and earnings outcomes for Ph.D. recipients. *Science* 350:1367–1371.
16 Lane J, Owen-Smith J, Rosen R, Weinberg B (2015) New linked data on research investments: Scientific workforce, productivity, and public value. *Res Policy* 44:1659–1671.
17 Weinberg BA, et al. (2014) Research funding. Science funding and short-term economic activity. *Science* 344:41–43.
18 Commission on Evidence-Based Policymaking (2017) Report of the Commission on Evidence-Based Policymaking. Available at https://www.cep.gov/content/dam/cep/report/cep-final-report.pdf. Accessed October 29, 2018.
19 Bestavros A (2017) Sharing knowledge without sharing data. Arthur M. Sackler Colloquia [video recording]. Available at https://www.youtube.com/watch?v=P2MmO458xu4. Accessed October 29, 2018.
20 Lapets A, et al. (2018) Accessible privacy-preserving web-based data analysis for assessing and addressing economic inequalities. *Proceedings of the First ACM SIGCAS Conference on Computing and Sustainable Societies* (ACM, New York), pp 1–5.
21 Bestavros A, Lapets A, Varia M (2017) User-centric distributed solutions for privacy-preserving analytics. *Commun ACM* 60:37–39.
22 Blau DM, Weinberg BA (2017) Why the US science and engineering workforce is aging rapidly. *Proc Natl Acad Sci USA* 114:3879–3884.
23 Bollen J, Crandall D, Junk D, Ding Y, Börner K (2017) An efficient system to fund science: From proposal review to peer-to-peer distributions. *Scientometrics* 110:521–528.
24 Ahrweiler P, Schilperoord M, Pyka A, Gilbert N (2015) Modelling research policy: Ex-ante evaluation of complex policy instruments. *J Artif Soc Simul* 18:5.
25 Jasny BR, Stone R (2017) Prediction and its limits. *Science* 355:468–469.
26 Helbing D, Balietti S (2011) How to create an innovation accelerator. *Eur Phys J Spec Top* 195:101–136.
27 van Harmelen F, et al. (2012) Theoretical and technological building blocks for an innovation accelerator. *Eur Phys J Spec Top* 214:183–214.
28 Ahmadpoor M, Jones BF (2017) The dual frontier: Patented inventions and prior scientific advance. *Science* 357:583–587.
29 Uzzi B, Mukherjee S, Stringer M, Jones B (2013) Atypical combinations and scientific impact. *Science* 342:468–472.
30 McAfee A, Brynjolfsson E (2017) *Machine, Platform, Crowd: Harnessing Our Digital Future* (W.W. Norton & Company, Hartford, VT).
31 Mukherjee S, Romero DM, Jones B, Uzzi B (2017) The nearly universal link between the age of past knowledge and tomorrow's breakthroughs in science and technology: The hotspot. *Sci Adv* 3:e1601315.
32 Virginia Commonwealth University, Center on Society and Health, Foundation RWJ (2015) Short distances to large gaps in health—New York City. Available at wp.vcu.edu/vcusochealth/wp-content/uploads/sites/3740/2015/04/LE-Map-NewYork-Black.jpg. Accessed October 29, 2018.
33 Brownstein J, et al. (2014) PREDICT: HealthMap. Courtesy of USAID, UC Davis School of Veterinary Medicine, and Children's Hospital Boston. In "10th Iteration (2014): The Future of Science Mapping," *Places & Spaces: Mapping Science*, edited by Katy Börner and Samuel Mills. Available at scimaps.org. Accessed October 29, 2018.
34 Williams RS, Lotia S, Holloway AK, Pico AR (2015) From scientific discovery to cures: Bright stars within a galaxy. *Cell* 163:21–23.
35 BPM Institute (c2018) Fragmented business processes destroy value. Available at www.bpminstitute.org/resources/articles/fragmented-business-processes-destroy-value. Accessed October 29, 2018.
36 Haskins R, Margolis G (2014) *Show Me the Evidence: Obama's Fight for Rigor and Results in Social Policy* (Brookings Institution, Washington, DC).
37 Rouse WB, Lombardi JV, Craig DD (2018) Modeling research universities: Predicting probable futures of public vs. private and large vs. small research universities. *Proc Natl Acad Sci USA* 115:12582–12589.
38 Rouse WB (2016) *Universities as Complex Enterprises: How Academia Works, Why It Works These Ways, and Where the University Enterprise Is Headed* (Wiley, Hoboken, NY).
39 Shneiderman B (2018) Twin-Win Model: A human-centered approach to research success. *Proc Natl Acad Sci USA* 115:12590–12594.
40 Shneiderman B (2016) *The New ABCs of Research: Achieving Breakthrough Collaborations* (Oxford Univ Press, New York).
41 Madhavan G, Phelps CE, Rouse WB, Rappuoli R (2018) Vision for a systems architecture to integrate and transform population health. *Proc Natl Acad Sci USA* 115:12595–12602.
42 Sekara V, et al. (2018) The chaperone effect in scientific publishing. *Proc Natl Acad Sci USA* 115:12603–12607.
43 Ma Y, Uzzi B (2018) Scientific prize network predicts who pushes the boundaries of science. *Proc Natl Acad Sci USA* 115:12608–12615.
44 Zuckerman H (1977) *Scientific Elite: Nobel Laureates in the United States* (Free, New York).
45 National Academies of Sciences, Engineering, and Medicine (2018) *How People Learn II: Learners, Contexts, and Cultures* (National Academies, Washington, DC).
46 Tornatzky LG, Rideout EC (2014) Innovation U 2.0: Reinventing university roles in a knowledge economy. Available at https://ssti.org/report-archive/innovationu20.pdf. Accessed October 29, 2018.
47 Lawton D, et al. (2012) Online learning based on essential concepts and formative assessment. *J Eng Educ* 101:244–287.
48 Astronautics AIoAa (2012) Recruiting, retaining, and developing a world-class aerospace workforce. Available at https://www.aiaa.org/uploadedFiles/Issues_and_Advocacy/Education_and_Workforce/Aerospace%20Workforce-%20030112.pdf. Accessed October 29, 2018.
49 National Academy of Engineering (2017) NAE Grand Challenge: Advancing personalized learning. Available at www.engineeringchallenges.org/challenges/learning.aspx/ Accessed October 29, 2018.
50 Richey M (2014) A complex sociotechnical systems approach to provisioning educational policies for future workforce. *Proceedings of the Conference on Systems Engineering Research*, eds Madni AM, Boehm BW (CSER, Redondo Beach, CA), pp 857–864.
51 Auerswald PE (2017) *The Code Economy: A Forty-Thousand Year History* (Oxford Univ Press, New York), 1st Ed.
52 Beyer D (2016) *The Future of Machine Intelligence: Perspectives from Leading Practitioners* (O'Reilly Media, Sebastopol, CA).
53 Brynjolfsson E, McAfee A (2014) *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies* (W.W. Norton & Company, New York).
54 Ford M (2015) *Rise of the Robots: Technology and the Threat of a Jobless Future* (Basic Books, New York).
55 Rouse WB, Spohrer JC (February 8, 2018) Automating versus augmenting intelligence. *J Enterprise Transformation*, 10.1080/19488289.2018.1424059.
56 Roco MC, Bainbridge WS (2003) Overview converging technologies for improving human performance. *Converging Technologies for Improving Human Performance: Nanotechnology, Biotechnology, Information Technology and Cognitive Science*, eds Roco MC, Bainbridge WS (Springer, Dordrecht, The Netherlands), pp 1–27.
57 Milojević S, Radicchi F, Walsh JP (2018) Changing demographics of scientific careers: The rise of the temporary workforce. *Proc Natl Acad Sci USA* 115:12616–12623.
58 Martinez W (2018) How science and technology developments impact employment and education. *Proc Natl Acad Sci USA* 115:12624–12629.
59 Börner K, et al. (2018) Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy. *Proc Natl Acad Sci USA* 115:12630–12637.
60 Haskel J, Westlake S (2017) *Capitalism Without Capital: The Rise of the Intangible Economy* (Princeton Univ Press, Princeton).
61 Batty M (2013) *The New Science of Cities* (MIT Press, Cambridge, MA).

62 Keller S, Korkmaz G, Robbins C, Shipp S (2018) Opportunities to observe and measure intangible inputs to innovation: Definitions, operationalization, and examples. *Proc Natl Acad Sci USA* 115:12638–12645.
63 Jara-Figueroa C, Jun B, Glaeser EL, Hidalgo CA (2018) The role of industry-specific, occupation-specific, and location-specific knowledge in the growth and survival of new firms. *Proc Natl Acad Sci USA* 115:12646–12653.
64 Olmos LE, Çolak S, Shafiei S, Saberi M, González MC (2018) Macroscopic dynamics and the collapse of urban traffic. *Proc Natl Acad Sci USA* 115:12654–12661.
65 Khanna P (2016) *Connectography: Mapping the Future of Global Civilization* (Random House, New York).
66 Dash Nelson G, Rae A (2016) An economic geography of the United States: From commutes to megaregions. *PLoS One* 11:e0166083.
67 Lazer D, Kennedy R, King G, Vespignani A (2014) Big data. The parable of Google Flu: Traps in big data analysis. *Science* 343:1203–1205.
68 O'Neil C (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Broadway Books, New York).
69 Allison DB, Shiffrin RM, Stodden V (2018) Reproducibility of research: Issues and proposed remedies. *Proc Natl Acad Sci USA* 115:2561–2562.
70 Harris R (2017) *Rigor Mortis: How Sloppy Science Creates Worthless Cures, Crushes Hope, and Wastes Billions* (Basic Books, New York).
71 Munafò MR, et al. (2017) A manifesto for reproducible science. *Nature Human Behavior* 1:0021.
72 Erdemir A, Mulugeta L, Lytton WW (2015) The "not so simple" rules for credible practice of modeling & simulation in healthcare. Available at https://simtk.org/svn/cpms/doc/posters/fmd_2015.pdf. Accessed October 29, 2018.