



BRIEF PROGRAM OVERVIEW

August 24-26, 2021

Tuesday, August 24

Check in: 4:00pm
Dinner: 6:00pm – 7:30pm
Introduction: 7:45pm – 8:00pm
Talks: 8:00pm – 9:00pm

Wednesday, August 25

Breakfast: 8:00am – 9:00am
Talks: 9:00am – 10:20am
Coffee Break: 10:20am – 10:40am
Talks: 10:40am – 12:00pm
Lunch: 12:30pm – 2:00pm
Free time: 2:00pm – 6:00pm
Dinner: 6:00pm – 7:30pm
Talks: 7:30pm – 8:30pm
Mingle: 8:30pm

Thursday, August 26th

Breakfast: 8:00am – 9:00am
Check out: 10:00am

DETAILED PROGRAM OVERVIEW

August 24-26, 2021

Tuesday, August 24th

EVENING SESSION: 8:00PM – 9:00PM

Intellectual Property Revisited

Filippo Mezzanotti

TBD

Scientific Prizes and the Extraordinary Growth of Scientific Topics

Ching Jin

Fast growing scientific topics have famously been key harbingers of the new frontiers of science, yet, large-scale analyses of their genesis and impact are rare. We investigated one possible factor connected with a topic's extraordinary growth: scientific prizes. Our longitudinal analysis of nearly all recognized prizes worldwide and over 11,000 scientific topics from 19 disciplines indicates that topics associated with a scientific prize experience extraordinary growth in productivity, impact, and new entrants. Relative to matched non-prizewinning topics, prizewinning topics produce 40% more papers and 33% more citations, retain 55% more scientists, and gain 37% and 47% more new entrants and star scientists, respectively, in the first five-to-ten years after the prize. Funding do not account for a prizewinning topic's growth. Rather, growth is positively related to the degree to which the prize is discipline-specific, conferred for recent research, or has prize money. These findings reveal new dynamics behind scientific innovation and investment.

Connecting biomedical laboratory science to studies of science and scholarship

Thomas Stoeger

Biomedical research currently concentrates on a small set of genes that have already been heavily investigated more than 30 years ago. This concentration cannot be explained by current experimental possibilities or the physiological importance of individual genes. Yet, this concentration leaves more than half of the human genes ignored by small-scale laboratory science, which constitutes most of current-day biomedical research. Shockingly, this also extends to frequently investigated topics of direct societal relevance such as COVID-19 or Alzheimer's disease. Hoping that this situation can be remedied by a more complete understanding of science, and by laboratory science becoming directly informed by studies of science and scholarship, I will dare to establish a "laboratory of the overlooked". While I believe that such an endeavor is quite likely to fail, I will try to establish such a laboratory in a way that seems most promising to succeed and to create further opportunities for my trainees and serve as a potential proof-of-feasibility to other laboratories. Consequently, my laboratory will be investigating overlooked aspects of already heavily investigated biological phenomena such as human aging. However, the number of pressing research questions outnumbers by at least two orders of magnitude the number of questions that the most productive laboratories have historically been able to resolve over the course of their existence. Consequently, I will in parallel strive to connect studies of science and scholarship more directly toward biomedical laboratory sciences, and to make the mutual exposure mutually rewarding. As parts of this effort, I will in the next few years continue to integrate over 100 different bibliometric and biomedical knowledge-sources to allow others and my laboratory to probe how science progresses as an

activity that depends not only on human scientists and prior knowledge, but also depends on chemical and physical properties of molecules, and the interactions of these molecules.

Technological Complementarity: the Division of Know-how and the Future of Work

Moh Hosseinioun

By applying bundles of technologies, workers put the embedded knowledge of a diversity of human experts to use. Here, we examine the relationship between such bundles of technologies used by human labor. To do so, we measure complementarity between technologies which can explain part of wages unexplained by other factors.

Wednesday, August 25th

MORNING SESSION #1: 9:00AM – 10:20AM

Reshaping the research enterprise

B. Ian Hutchins

Biomedical research has entered a phase of “hypercompetition”; current organization and policies are straining under this pressure. Scientists seeking to remain competitive depend on publishing in artificially scarce space in prestigious journals. I transitioned from neuroscience into computational science policy in order to build analytical tools that can be used to inform policy responses to these challenges. At the same time, I designed my methods such that they can also be used by scientists to communicate the impact of their work regardless of publication venue, in order to help alleviate undue career stresses for researchers. I will discuss the network science behind my measure of scientific influence, and the machine learning system I developed to predict early stages of bench-to-bedside translation, a crucial applied goal of biomedical research. These quantitative measures provide researchers with two additional ways to showcase their work’s influence, and are increasingly driving policy-making discussions as the research enterprise responds to modern pressures.

Algorithms and Shepherds in AI-Driven Scientific Discovery

Lav Varshney

Artificial Intelligence (AI) is becoming a prominent cognitive support tool in scientific discovery. Here we discuss four vignettes of our own work in AI-driven discovery: (1) discovering active phytochemicals in herbs and spices for different health conditions from a network science representation of traditional knowledge; (2) discovering new formulations of sustainable building materials that are stronger and require half the carbon emissions as existing formulations using variational autoencoders; (3) discovering laws of music theory in the same human-interpretable form as textbooks using a novel group-theoretic and information-theoretic technique called information lattice learning; and (4) showing that rates of technological innovation are explained in part by the structure of online friendship networks in universities. We close by discussing the role of human shepherds in AI-driven science, and how AI may impact human autonomy.

Accelerating science with human versus alien artificial intelligences

Jamshid Sourati

Data-driven artificial intelligence models fed with published scientific findings have been used to create powerful prediction engines for scientific and technological advance, such as the discovery of novel materials with desired properties and the targeted invention of new therapies and vaccines. These AI

approaches typically ignore the distribution of human prediction engines—scientists and inventors—who continuously alter the landscape of discovery and invention. As a result, AI hypotheses are designed to substitute for human experts, failing to complement them for punctuated collective advance. We show that incorporating the distribution of human expertise into self-supervised models by training on inferences cognitively available to experts dramatically improves AI prediction of future human discoveries and inventions. Moreover, by tuning our model to avoid the crowd, it generates scientifically promising “alien” hypotheses unlikely to be imagined or pursued without intervention, which we demonstrate hold promise to not only accelerate but also punctuate scientific advance.

A novel approach to symmetry-aware machine learning

Nima Dehmamy

The reductionist view of the world has been very successful in disciplines such as physics. Reductionism requires us to simplify and minimize the number of variables needed to describe the behavior of a system. In physics, such simplifications lead to the idea of symmetry: that the laws of physics stay the same if the system is moved, rotated, or transformed in other ways. In machine learning, this reductionism can lead to a significant reduction of parameters to be learned. As a result, symmetry-aware neural networks have become an active area of research in recent years. Yet, existing methods can only encode known symmetries into neural networks and are difficult to implement. We take a step toward simplifying the process of encoding symmetries by introducing a novel and easy-to-implement neural network architecture, which we call L-conv. We prove that L-conv can approximate many other symmetry encoding architectures. Additionally, the simplicity of L-conv allows us to learn symmetries from data. Lastly, we find that L-conv is closely related to models used in physics. This connection could allow for using mature physics tools in ML and to use ML to make more general and flexible physics models.

Quantifying punctuated record-breaking dynamics

Yian Yin

As the fundamental engine of human society and economic growth, science and technology not only advance in breadth, opening up more novel areas for exploration, they also progress in depth, developing increasingly sophisticated solutions to accomplish specific tasks. Yet our quantitative understanding of the record-breaking dynamics in science and technology remains limited. Here we collect large-scale datasets of technology progress across various domains to shed light on the emergence of record-breaking innovations. We find a universal punctuated dynamics of innovation record-breaking where short periods of clustered breakthroughs and long periods of stasis appear alternatively, highlighting a short-term predictability but long-term uncertainty in technology breakthrough. We further develop a simple mathematical framework that not only allows us to probe quantitatively the punctuated pattern, but also offers a new mechanistic tool to study its relationship with other contingency factors.

Learning, Fast and Slow: The Returns to Experience and Team Size for High-Impact Innovation in the United States between 1836 and 1975

Christopher Esposito

We study changes in the returns that U.S. patent inventors received from experience and collaboration between 1836 and 1975. Our study combines two novel datasets which allow us to trace the careers of inventors, record instances of collaboration, and measure the impact of patents on subsequent invention across a long period of U.S. history. We show that the returns to experience, measured by the change in the likelihood that inventors create high-impact invention as their careers progress, have been negative since the 1920s, while the returns to collaboration, measured as the change in the likelihood of inventing a high-impact invention when collaborating with larger teams, became positive and significant in the 1920s. We develop a model to interpret these findings. The model proposes that teams “learn” quickly by

pooling together the knowledge of their co-inventors while inventors learn slowly by accumulating experience. When the knowledge frontier is rapidly expanding, inventors are not able to learn quickly enough to keep up with the new ideas that are being introduced, creating a negative association between experience and impact. To test the model, we examine whether the returns to experience and collaboration vary across knowledge fields that are advancing at different rates, finding that (a) the average level of experience inventors is lower in fast-advancing knowledge fields, (b) the returns to experience are more negative in fast-advancing fields, (c) the average size of inventor teams is larger in fast-advancing fields, and (d) the returns to team size are greater in fast-advancing fields.

MORNING SESSION #2: 10:40AM – 12:00PM

TBD

Luis Bettencourt

TBD

TBD

Hyejin Youn

TBD

Abandonment of Innovation and Emergence of Fragility in Robust Ecosystems

Binglu Wang

Despite extensive studies on diffusion of innovations, our knowledge about the reverse process—abandonment of innovations—remains limited. Here, we analyze two large-scale datasets, each capturing detailed socio-temporal patterns that trace the entire lifecycle of innovations. By analyzing 2.5 M scholars studying in 2651 scientific fields and 3.5 M individuals using 994 mobile handsets, we find that, in contrast to the Poissonian dynamics commonly assumed in the abandoning process of the lifecycle, the abandoning probability increases with the number of past abandonments, revealing that a bandwagon effect characterizes the abandonment of innovations. We examine the social networks underneath the two systems through co-authorships and mobile communication records, finding that a preferential abandonment mechanism at a network level is responsible for generating the observed effect. Most importantly, we show analytically that the presence of preferential abandonment induces a structural collapse in heterogeneous system, where networked systems that were thought to be robust undergo a novel phase transition. We test the theoretical predictions systematically in our datasets, obtaining broadly consistent empirical support. Together these results demonstrate that the collapse of real systems follows reproducible but fundamentally different dynamics than what traditional theoretical frameworks predicted. Our findings suggest that preferential abandonment and the structural collapse it induces may be a generic property that prevails in the declining phase of the innovation lifecycle.

Estimating the Impact of Artificial Intelligence across Scientific Disciplines

Jian Gao

The ongoing artificial intelligence (AI) revolution has the potential to change almost every line of work. Despite the enormous efforts devoted to understanding the impact of AI in labor, we know relatively little about how AI might impact science. By integrating a variety of data sources including millions of publications, university course syllabi, research funding, and their inter-linkages, here we attempt to

estimate the benefits of AI in science. Our analyses reveal several findings: (1) while research disciplines overall increasingly use more AI and have larger potential AI benefits during the past two decades, there is an abrupt increase in the direct use of AI since 2015, especially new AI capabilities. (2) A field's potential AI benefits can predict its future direct use of AI, and new AI capabilities are increasingly applicable for direct use. (3) Despite discipline-level heterogeneity, every discipline has some subfields that can benefit from AI, suggesting that AI benefits are pervasive across disciplines. (4) The education of AI in courses is overall misaligned with AI benefits in science, while there is a well alignment between AI funding investment and AI scientific benefits. (5) There are disparities of AI benefits across gender and race, where women and under-represented minorities are disproportionately present in disciplines of smaller AI benefits. Our work may help deepen our understanding of AI's important policy implications for the future of science.

TBD

Zander Furnas

TBD

EVENING SESSION: 7:30PM – 8:30PM

TBD

Agnes Horvat

TBD

TBD

David Schonthal

TBD

TBD

Suman Maity

TBD

TBD

Fengli Xu

TBD