2019 IRIS Summit — Agenda

September 26-27

Insitute for Social Research
426 Thompson St., Ann Arbor, MI

Thursday, Sept. 26

8:30 - 9 a.m.
Room 1450

Continental Breakfast

9 - 9:30 a.m.
Room 1430

Summit Opening — Jason Owen-Smith, IRIS Executive Director; Rebecca Cunningham, Interim Vice President for Research, University of Michigan

9:30 - 11 a.m.
Room 1430

Policy Plenary Session, "Perspectives on Making the Case for Research Universities"
Panelists will approach the topic of responding to demands for accountability from a variety of perspectives, with the goal of developing indicators that can demonstrate the impact research universities can have. Universities face unrelenting pressure to quantify their impact, which raises challenges in taking a more holistic and long-term view of applied research impacts. In addition, universities and academic medical centers must quantify to Congress and the public the outputs of increases to research funding in an era when new drugs and treatments can take many years to be developed. In addition, understanding the effects of involvement in research by students is increasingly critical, particularly in understanding how this may affect later outcomes such as degree completion and subsequent earnings.
Speakers:
- Matt Owens, Vice President for Federal Relations, Association of American Universities
- Sheila Martin, Vice President for Economic Development and Community Engagement, Association of Public and Land-grant Universities,
- Steve Heinig, Director, Science Policy, Association of American Medical Colleges
- Facilitator: Cynthia Wilbanks, Vice President for Government Relations, University of Michigan

11 - 11:15 p.m.

Break

11:15 - 12:15 p.m.
Room 1430

IRIS 101: Overview for Data Providers, Researchers and Policy Makers
This session will orient participants to the work that goes into the production of reports and a research dataset, and how this can be used to support policy making. The session will include an introduction of the technical team and will provide an overview of the data production and development processes. IRIS staff will also will share details about the IRIS UMETRICS research dataset production process as well as the 2019 research data release. Finally, representatives from IRIS member institutions will provide examples of how IRIS data can be leveraged for policy purposes.
Speakers:
- Kevin Bjorne, Technical Director, IRIS
- Natsuko Nicholls, Research Manager, IRIS

- Beth Uberseder, Research Support Specialist, IRIS
- Kimberly Griffin, Director of Electronic Research Administration, Northwestern University, IRIS Board of Directors
- Andrew Sallans, Assistant Director for Electronic Research Administration, University of Virginia

12:15 - 1:30 p.m.
Room 1430

Lunch & State of IRIS, Jason Owen-Smith, Executive Director, IRIS
Jason Owen-Smith will provide an update on the activities of the past year and will discuss the opportunities and challenges facing IRIS as we think about future directions.

1:30 - 2:30 p.m.
Data Track
Room 1430

Breakout Session 1
"Canonical Linkages of Employees and Research Grants"
IRIS has linked administrative records of employees and awards to federal grant programs, university research, scholarly publications, and patents. This session provides an overview of our linkage work including its goals, methods, and some applications. What do we gain from linking IRIS data to other sources? How do we clean, harmonize, and match data between sources? And how can administrators, researchers, and policymakers leverage the benefits of linked data to better understand university research and its impact?
Speakers:
- Natsuko Nicholls, Research Manager, IRIS
- Matthew VanEseltine, Research Investigator, IRIS
- Raphael Ku, Research Support Associate, IRIS

Research Track
Room 6050

Research Talks
"Postdoc Career Paths and Labor Market Outcomes: Evidence from UMETRICS-enhanced Census Data"
Gerald Marschke, Associate Professor of Economics, and Holden Diethorn, Ph.D Candidate in Economics, State University of New York at Albany
We describe our use of UMETRICS and machine learning to predict the postdoc status of workers in the Census's employee-employer Longitudinal Employer-Household Dynamics database. This is part of a larger project to leverage UMETRICS data to build a new, comprehensive database that will enable researchers to measure the labor market outcomes of STEM PhD graduates and postdocs: what occupations and industries they work in, how much they earn, and how their careers develop over time. We will present preliminary evidence of career outcomes of STEM postdocs using these new data.

"Grant Collaboration Networks"
Jacob Fisher, Research Investigator, IRIS
This presentation will focus on collaboration networks, meaning networks of people who work on several grants together. This presentation will show how these networks are constructed and will present preliminary results about characteristics of grant funding for university employees over time.

"IRIS Authority Control for Connecting Publication Data to UMETRICS"
Jinseok Kim, Research Assistant Professor, IRIS
Research and patent publication data can enable us to measure the impact of funded research in UMETRICS on science production and collaboration, and technological innovation. The most important prerequisite of realizing such a potential is, however, to correctly identify researchers and organizations recorded in ambiguous names in publication data. This talk reports an overview and progress of authority control framework developed by IRIS to address the name disambiguation challenge at scale using state-of-the-art machine learning techniques.

Policy Track *Room 6080*	<u>"How Graduate Schools Can Incorporate IRIS Data Into Analytic Strategies"</u> Graduate schools employ a diverse set of analytic strategies to recruit, retain, and understand the experiences of their students. At the University of Michigan, the Rackham Graduate School established its own institutional research office to provide insights about graduate programs. The office works with a range of traditional institutional data (admissions, enrollment, degrees, finances, and career outcomes) while also leveraging its position to advance knowledge of graduate education through studies such as the Michigan Doctoral Experience Study (MDES). After summarizing the University of Michigan's work, we discuss ways in which IRIS data might prove useful to addressing pressing questions about graduate education. <u>Speakers:</u> • <u>John Gonzalez, Director of Institutional Research, Rackham Graduate School, University of Michigan</u> • <u>Allyson Flaster, Assistant Research Scientist, Inter-university Consortium for Political and Social Research</u> • <u>Jason Owen-Smith, Executive Director, IRIS</u>

2:30 - 3:30 p.m.	**<u>Breakout Session 2</u>**
Data Track *Room 1430*	<u>"The Evolution of Job Title Classification: A Prototype"</u> IRIS has developed a new and improved method of employee classification that focuses on the impact of students and trainees. This new method of classification will allow IRIS to create more student focused metrics of impact in current and future reports. During this session we will compare and contrast the new methodology of classification with the current labor function method of classification and discuss what this means for our members and the reports IRIS produces. <u>Speakers:</u> • <u>Kevin Bjorne, Technical Director, IRIS</u> • <u>Robert Truex, Data Manager, IRIS</u> • <u>Elissa Irhamy, Research Associate, IRIS</u>

Research Track *Room 6050*	<u>Research Talks</u> <u>"Measuring the Efficiency of the Research Enterprise"</u> <u>Jason Coupet, Assistant Professor of Public Administration, North Carolina State University</u> Our project combines IRIS UMETRICS data with AUTM data to construct a profile of university research performance. We then use Data Envelopment Analysis (DEA), to construct a measure of university research productivity and identify strategic groups. Our work is ongoing, but we posit that benchmarking techniques such as what our study puts forth can aid the strategic decision making of universities <u>"The Network Dynamics of Interdisciplinary Research"</u> <u>Doug Guilbeault, PhD student, Communications, University of Pennsylvania</u> How do novel categories emerge in scientific teams? We use simulations to show that increasing the size of a research team can unexpectedly increase the similarity of the category systems produced by teams of the same size and show how we can predict patterns of category formation in the abstracts associated with decades of grants won by research teams in the IRIS UMETRICS data. <u>"A New Geography for Research: Describing Research Heterogeneity via Classification of Expenditure Profiles"</u> <u>Elan Segarra, PhD student, Economics, University of Wisconsin</u> It is vitally important that science of science investigators consider and account for heteroge-

neity when studying the research process and its outcomes. The transactional nature of the IRIS UMETRICS data provides a story for the type of research being conducted within each project. This paper investigates how expenditure profiles vary across classical research divisions, such as discipline and motivation, and examines how expenditure profiles can be used to construct novel research classifications.

"Quasi-experimental approaches to evaluating clinical research training programs"
Elias Samuels, Michigan Institute for Clinical and Health Research
The CTSA Consortium contains over 60 hubs, each of which dedicate funding for early-career faculty research and training through Clinical Research KL2 Scholars programs. There is a growing need for efficient approaches to conducting evaluations of these programs that can be replicated across CTSA hubs. This study utilized a variety of data sources and statistical methods to demonstrate approaches to evaluating one of these programs in a rigorous and reproducible way.

"Quantifying dynamics of failure across science, startups and security"
Yian Yin, Graduate Research Assistant, Complex Systems, Northwestern University
Human achievements are often preceded by repeated attempts that initially fail, yet little is known about the mechanisms governing the dynamics of failure. Here we develop a simple one-parameter model that mimics how successful future attempts build on those past, which predicts that agents who share similar characteristics and learning strategies may experience fundamentally different outcomes following failures. The model makes several empirically testable predictions on dynamics of failure, which are systematically verified on large-scale datasets from three disparate domains (science, startups and security).

Policy Track *Room 6080*	"Telling Stories With IRIS Data" Last year IRIS was approached by the University of Michigan's Vice President for Communications to help identify stories of impact on the west side of the state. Mining the IRIS vendor data unearthed an important and heartwarming story about the impact of research. A professor of kinesiology studied the effects of pediatric treadmills in helping infants with Down's Syndrome acquire greater mobility. He purchased treadmills from a manufacturer on the west side of the state, Carlin's Creations; the dollars from research allowed the company to expand. This panel brings together university communications professionals and the PI of the study to discuss the mechanics of telling compelling stories based on IRIS data. Speakers: • Alex Piazza, Communications Manager, U-M Office of Research • Dale Ulrich, Professor of Kinesiology, U-M • Dan Meisler, Communications Coordinator, IRIS

3:30 - 3:45 p.m.	**Break**

3:45 - 5:15 p.m. Room 1430	**Data Plenary Session.** "Opportunities for Data Aggregation" In this plenary discussion, the presenters will highlight the benefits of aggregating data, including more powerful analyses and greater confidentiality protections. There are a variety of ways to think about aggregation of data for IRIS member universities, including state-level aggregation, combining IRIS members who hold AAU or APLU memberships, and the like. Data aggregation does carry some particular challenges; the presenters will provide examples of such issues and discuss how they might be overcome. Finally, the panelists will highlight an example of one such aggregation for three institutions in the state of Michigan, discussing how such an aggregated report might be leveraged for the good of all involved.

5:15 - 6:15 p.m.   **Group photo and reception in Atrium**

Friday, Sept. 27

8:30 - 9 a.m.   **Continental Breakfast**
Room 1450

9 - 9:15 a.m.   **Welcome to Day 2** — Jason Owen-Smith, IRIS Executive Director
Room 1430

9:15 - 10:45 a.m.   **Research Plenary Session**. "New Findings from Research with IRIS Data"
Room 1430   IRIS co-founders Bruce Weinberg and Julia Lane present an overview of current work by IRIS researchers. They will highlight research findings from their respective teams at The Ohio State University and New York University and also provide insight on future directions and opportunities for the IRIS dataset, including current and potential collaborations.
Speakers:
- Julia Lane, Professor, Wagner School for Public Policy, New York University
- Bruce Weinberg, Professor of Economics, The Ohio State University

11 a.m. - noon   **Breakout Session 3**
Data Track   "An IRIS Project to Better Match and Report Grant Data: The Award ID Pilot"
Room 1430   During the 2018 IRIS Summit, the award identity project was introduced to IRIS members. The project is an effort to improve and expand award data submitted to IRIS. After a year of development and refinement, IRIS is ready to share the finalized version of the new data file that IRIS is asking all university members to prepare and submit as part of the normal submission process. This year we will be discussing information about the amount of effort to produce the file, what types of data are expected, and how it will be used with examples of visualizations and report impact.
Speakers:
- Kevin Bjorne, Technical Director, IRIS
- Amy Butchart, Lead Programmer, IRIS
- Chris Allan, Business Systems Analyst, Institute for Social Research, University of Michigan

Research Track   Research Talks
Room 6050   "University R&D Funding Sources & Innovation Outcomes"
Alex Xi He, Assistant Professor of Finance, University of Maryland
While the majority of university research is funded by the federal government, a large fraction of researchers receive funding from private for-profit firms. We use IRIS UMETRICS data and linked data on publications, startups, and student placements to comprehensively investigate how funding sources affect basic research and commercial innovations in the short and longer run.

"Antecedents and Consequences of Sponsored Teaming among STEM Doctoral Graduates"
Kevin Kniffin, Assistant Professor, SC Johnson College of Business, Cornell University
Teamwork among academic researchers has been increasing substantially across STEM fields for the past several decades. On the one hand, research has demonstrated that teamwork tends to pay dividends to members through direct rewards (higher salary) and indirect benefits (higher impact). On the other hand, there is also evidence that team size can become too large to be effective (e.g., coordination problems tend to increase). With the benefit of IRIS data, this talk explore questions relating to team size, team composition, and team performance.

"How Funding Facilitates Interdisciplinary Research for Early Career Scientists"
Ran Xu, Assistant Professor, Department of Allied Health Sciences, University of Connecticut
Scientists often need to conduct interdisciplinary research that combines knowledge and concepts from multiple disciplines to fully understand the mechanisms and find solutions. However, there are also considerable risks associated with conducting interdisciplinary research, especially for early career scientists. It is vital to gain a better understanding of the institutional factors (e.g. funding, infrastructure) that facilitate/inhibit the production of interdisciplinary research, as well as the mechanisms through which they function. This project employs IRIS and ProQuest data to investigate these factors, using interdisciplinary neuroscience dissertations as a case study.

Policy Track *Room 6080*	"Using IRIS Data for Resource Allocation and Accountability" Oversight bodies – OMB, OSTP, Appropriators, Authorizers – have potential use cases for IRIS data that are different than those of the funding agencies or the research universities. Since many federal research programs supporting basic research are organized by disciplinary fields and subfields, potential use cases would ideally focus on evaluating differential scientific, technical and workforce impacts of subfields. Identifying paths of influence is often more relevant than quantification of benefits. Similarly, tools that enable international comparisons of research portfolios would be of great utility. Speaker: Michael Holland, Vice Chancellor for Science Policy and Research Strategies at the University of Pittsburgh, IRIS Board of Directors
Noon - 1:15 p.m. Room 1430	**Lunch & "Novel Approaches to Working with Administrative Data"** Administrative data of all stripes can provide a wealth of information, and may be utilized to respond to an incredibly diverse array of questions and outcomes. Panelists in this session will highlight just a few very interesting approaches. MariJane Selvaggio will discuss the work Equifax Workforce Solutions is doing with many universities as part of its Graduate Outcomes Project. Wan-Ying Chang of the National Center for Science & Engineering Statistics, will share information about the linkages between IRIS UMETRICS data and the Survey of Earned Doctorates. Shawn Klimek of the U.S. Census Bureau will share updates on the Post Secondary Education Outcomes project. Finally, Tim McKay of the University of Michigan will describe his work with the SEISMIC project that aims to help students be more successful in difficult gateway STEM courses. Speakers: • MariJane Selvaggio, Account Executive, Equifax Workforce Solutions • Wan-Ying Chang, Mathematical Statistician, National Center for Science and Engineering Statistics • Shawn Klimek, Assistant Center Chief for Business Research, U.S. Census Postsecondary Education Outcomes • Timothy McKay, Professor of Physics, University of Michigan, Sloan Equity and Inclusion in STEM Introductory Courses (SEISMIC) project

1:15 - 2 p.m.   <u>**Closing Session, Track Summaries**</u>
Room 1430

Data Track	Kevin Bjorne, Technical Director, IRIS
Research Track	Natsuko Nicholls, Research Manager, IRIS
Policy Track	Dan Meisler, Communications Coordinator, IRIS
Closing Remarks	Jason Owen-Smith, Executive Director, IRIS