

# Presentations – Beyond the Numbers 2016

October 6-8, 2016

## [It's Out There — But Where?: Tracking Down, Sorting Through, and Utilizing Open International Economic Data](#)

Emily Alford, Indiana University

Jessica Huffman, Indiana University

## [Understanding the Census of Governments: Data for Public Finance](#)

Jennifer C. Boettcher, Georgetown University

## [The Future of Older Data: Zombie List, Moody's, Corporate Annual Reports \(ARs\)](#)

Jennifer C. Boettcher, Georgetown University

Greg Fleming, University of Chicago

Judith Nixon, Purdue University

## [The Importance of GIS Data Repositories](#)

Douglas Burns, University of North Texas

## [Introducing the Categorically LINKed Timeline \(CLINT\) database](#)

San Cannon, Federal Reserve Bank of Kansas City

## [Best Practices for Replicable and Accessible Data](#)

Andrew C. Chang, Board of Governors of the Federal Reserve System

Alison Raab Labonte, Board of Governors of the Federal Reserve System

## [Conducting Economic Research at the University of Missouri Research Data Center](#)

Jacob Cronin, United States Census

## [Further Developments of the Taxonomy of Terms and Concepts at the US Bureau of Labor Statistics](#)

Daniel Gillman, US Bureau of Labor Statistics

## [Rethinking How We Teach Macro: Bring FRED Into the Classroom](#)

Ryan Herzog, Gonzaga University

## [How to Deal With Some Commonly Encountered Financial Data Research Problems](#)

Todd Hines, Princeton University

## [ReplicationWiki - Improving Transparency in Economic Research](#)

Jan H. Höffler, University of Göttingen

## [Curating Data for Reproducibility](#)

Cynthia Hudson-Vitale, Washington University in St. Louis

**[Scanning for Scholarship: Digitization of State and Local Government Publications to Support New Frontiers of Original Research](#)**

Kris Kasianovitz, Stanford University Libraries

**[Organizing Economic Information: An Overview of Application and Reuse Scenarios of an Economics Knowledge Organization System](#)**

Andreas Oskar Kempf, ZBW - German National Library of Economics - Leibniz Information Centre for Economics

**[Teaching with Data and Interacting with FRED](#)**

Diego Mendez-Carbajo, Illinois Wesleyan University

**[Using Stata for Data Work](#)**

James Ng, University of Notre Dame

**[FRED Schema : Metadata for Aggregate Economic Time Series](#)**

Genevieve Podleski, Federal Reserve Bank of St. Louis

**[History of the Federal Reserve Board Statistical Releases](#)**

Sian Seldin, Research Library, Federal Reserve Board

**[Services for Storing and Sharing Research Data in Economics — Two Case Studies from Germany: ZBW Journal Data Archive and SowiDataNet](#)**

Ralf Toepfer, ZBW Leibniz Information Centre for Economics

**[Best Practice for International Economic Data Discovery: In Case of East Asian Countries](#)**

Jungwon Yang, Clark Library, University of Michigan

**[Are You Asking the Right Questions of Your Financial Data Providers?](#)**

James Zussman, RateWatch Scholar

## **It's Out There — But Where?: Tracking Down, Sorting Through, and Utilizing Open International Economic Data**

Emily Alford, Indiana University

Jessica Huffman, Indiana University

Librarians, data researchers, and other information professionals are benefiting more each day as an increasing amount of open economic data becomes available. While not always realized, this is especially true with open international economic information. This growing amount of open data is naturally causing the librarian's role to change. Less time is now spent selecting materials and subscription databases, and more time is spent consulting users on where to find this open data, how best to organize it, and how to really utilize it.

This typically involves very specific information requests, and the librarian must constantly be enhancing one's own education on these open resources — from World DataBank, UNdata, and Eurostat to International Monetary Fund Data and Statistics, and a variety of resources specifically providing international trade data. As Social Sciences Librarians at Indiana University, the presenters of this short session will share the main attractions of these and other similar resources, as well as brief tips and tricks to benefit the user.

Also discussed will be methods used by the presenters to promote and teach these resources, how the librarians personally stay polished on developments, and what further changes in open international economic data can be anticipated. Attendees will learn a great deal about applying open data in their daily reference and consultation sessions, as well as become more likely to embrace these resources in general.

## **Understanding the Census of Governments: Data for Public Finance**

Jennifer C. Boettcher, Georgetown University

Engaged citizens require an understanding of governments at all levels: federal, state, and local. The Census of Governments allows you to benchmark your community with other in that state, size, or configuration. It gives details on public employment, education, and operations. This resource reveals the financial transparency between all levels of government.

To understand a government is to understand the context in which a government operates. This workshop will review the Census of Governments and other government related resources that could help you find related data, understand the structure of governments, follow the flows of money, and monitor the fiscal strength of special districts, school districts, other local and state governments. This webinar will give you the confidence to help with promoting marketing plans to governments and preparing economic development plans for policy makers.

Attendees will leave with:

- an understanding of how the Census data are collected,
- a basic understanding of state and local government budgeting,
- how to match revenues and expenditures across communities,
- a review of the different reports and series including those relating to public employees and education,
- experience retrieving data from various websites,
- how to read the tables,
- how to find more current data,
- knowing what is NOT in the Census,
- how to place their communities in context, and
- ways to bring this knowledge into their communities

Building on her knowledge of the Economic Census and passion for government documents, Jennifer C. Boettcher, Georgetown University, will lead the workshop and provide attendees with a folder of worksheets, a glossary, and a list of ideas to introduce the Census of Governments and public finance to your town or library.

## **The Future of Older Data: Zombie List, Moody's, Corporate Annual Reports (ARSs)**

Jennifer C. Boettcher, Georgetown University

Greg Fleming (Regenstein Library, University of Chicago)

Judith Nixon (Humanities, Education & Social Sciences Library, Purdue University)

Today access to data is easy. It can be found over the internet or bought from copyright holders. What about information created before the internet?

This panel discussion will cover three business areas: Zombie List, a web-based tool for locating data that transformed into other formats; sources of company financial information prior to the Securities Act of 1933; and corporate annual reports (ARSs), printed annual reports mailed by publically owned companies to investors covering finances and current plans for publically owned companies.

Boettcher will monitor the panel and will introduction to the concepts behind the Zombie List: it is a work in progress intended to be a web-based tool for locating data that transformed into other formats. Many of the core business and economics reference titles used by librarians have been discontinued, sold, or have been recreated into new formats. Tracking data back to similar information to identify trends has been challenging for information professionals who try to create collections to meet the needs of the users. Unlike traditional library catalogs, Zombie List is not making a direct link through the publisher, but the content and context of that information. (10 minutes)

Nixon will present on the status of research libraries' collections of printed annual corporate reports. Although collecting ARSs was standard in university and college libraries with business schools, during the 1980s many libraries ceased collecting in print and purchased microfiche products and later digital products. Most libraries with print collections then discarded them, because they were bulky, low-use, and thinking they were probably duplicated in other collections. This left only a few research libraries with archived collections. Based on a combined index of twelve collections and an overlap analysis, Nixon will present findings on the uniqueness of these twelve collections and outline the strengths of several collections. (10 minutes)

Fleming will cover finding company financial information in the pre-SEC era: moving beyond the Moody's Manuals. Researchers searching for company financial data in the era prior to the enactment of the Securities Acts of 1933 and 1934 have long relied on the various manuals that were published by Moody, beginning in 1909. This presentation will review similar sources of information that predate those manuals and identify ways researchers can get access. Attendees will leave with a better understanding of the sources available and the types of data available from these sources. They will also learn about the challenges of constructing large data sets from these sources. (15 minutes)

The program will conclude with discussion and advice from the panelist addressing ways to preserve access to older published sources and simple steps that librarians can take so they contribute to the business history. (25 min)

## **When One Company Has the Data Our Researchers Need — Working With Corelogic to Purchase US Tax, Deed, and Non-Agency Backed Loan Data**

Adrienne Brennecke, St. Louis Federal Reserve Bank

Kris Kasianovitz, Stanford University Libraries

In this 45 minute presentation, we will talk about the process of working with CoreLogic, the primary provider of public record and non-agency backed loan data. This is not a typical vendor for academic libraries, which meant new acquisition territory had to be charted. We hope that our experience will aid other libraries and data centers in pursuing this extremely rich set of housing, mortgage and foreclosure data. We will describe the data and explain why CoreLogic was the only vendor for this data. We will discuss the negotiation of price and licensing terms and roadblocks and challenges we encountered and wrap up with explaining how researchers at Stanford University and the St. Louis Federal Reserve Bank are using this data.

## **The Importance of GIS Data Repositories**

Douglas Burns, University of North Texas

This short session will highlight the importance Geographic Information Systems (GIS) data repositories. Recent advances in technology have enabled GIS to mature as a software platform, but the availability and maintenance of GIS data is still a patchwork of resources at best – though improvements are happening across the nation. A number of current initiatives will be discussed, including efforts by the Federal government, industry and educational institutions. One example, specifically, will discuss the success of the Portal to Texas History as a digital repository at the University of North Texas; its success has helped create a nascent GIS repository partnership locally. Issues and challenges regarding GIS data repositories will also be addressed. For example, one recurring problem with GIS data is the sometimes ephemeral nature of the data, which makes it difficult to capture. Demand for GIS data repositories will only increase with time, so understanding their importance, accessibility and nuances will become increasingly important for GIS users across all disciplines, professional and non-professional alike.

## **Introducing the Categorically LINKed Timeline (CLINT) database**

San Cannon, Federal Reserve Bank of Kansas City

Economists and social science researchers use historical information to describe the present, predict the future, and understand the world around us. Sometimes empirical research involves an important understanding of that world, and the data that describe it, that is not easily derived from current information or readily available from simple web searches. For example, it can be hard to disentangle when data irregularities occur from changes in measures, important events, or market reactions. Without a deep understanding of economic history, it can be difficult to attribute measured effects to the appropriate event, especially without knowing what events might be relevant.

The motivation for developing the Categorically LINKed Timeline (CLINT) is to provide historical context of economic events since 1945 for users of any knowledge base. We are building this catalog to help people identify events that may have had some effect on measures of economic activity and attempting to link the events to the relevant descriptors. For example, the event “Hurricane Katrina” may have been a meteorological occurrence and a natural disaster but it affected measures of output, employment, and prices in 2005. An entry in the CLINT database for Hurricane Katrina would identify and categorize the event and the effects so that researchers can be aware and, if appropriate, take appropriate corrective measures for the relevant data.

This presentation will outline the project history, the current state of the CLINT database, the challenges we face in doing the categorization, and the opportunities for participation from the research community.

## **Best Practices for Replicable and Accessible Data**

Andrew C. Chang, Board of Governors of the Federal Reserve System

Alison Raab Labonte, Board of Governors of the Federal Reserve System

Sound economic scholarship and advances in policy decision-making rest upon accessible data and replicable results. In this set of presentations — two short 20-minute presentations, one by Andrew Chang, and one by Alison Raab Labonte — the presenters will discuss the challenges of improving replicability in economic research and ensuring proper citation, and thus findability, of data used in research.

First, Chang will discuss his research in which he found that more than 50% of a sample of economics research papers in 13 well-regarded journals were irreproducible, and explore recommendations on improving replication in economics research, including mandatory data and code files submissions to publications' data archives, notation on whether papers without replication files are exempt from a replication policy, required provision of details on how data analysis was conducted, and so on (Chang, Andrew C., and Phillip Li (2015). "Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say 'Usually Not'," Finance and Economics Discussion Series 2015-083. Washington: Board of Governors of the Federal Reserve System, <http://dx.doi.org/10.17016/FEDS.2015.083>).

Labonte will then shift the discussion to data citation recommendations, with a particular focus on citation of proprietary/vendor data citation and access; she will highlight some common missteps in citing vendor data, and make recommendations on best practices. She will underscore that data citation, like the citation of journal articles and books, is an important part of good research practice and should support data findability and reuse.

## **Conducting Economic Research at the University of Missouri Research Data Center**

Jacob Cronin, United States Census

The University of Missouri Research Data Center (MU RDC) is a joint project between the US Census Bureau and the University of Missouri and is part of the Federal Statistical Research Data Center (FSRDC) program. The MU RDC, housed at Ellis Library at the University of Missouri-Columbia, provides a secure environment where researchers with approved projects have access to numerous restricted-use federal datasets for research in economics. There are over forty different economic datasets that are available only through FSRDCs such as the MU RDC. Economic data are based on censuses of establishments and various surveys of establishments and firms. The Economic Censuses are the government's official five year measure of American business. These surveys attempt to cover all establishments within nine specified industries, including most establishments outside of agriculture. These micro datasets provide establishment-level information on employment and labor costs, sales and revenue, costs of materials and supplies, as well as a large variety of other measures of economic activity. Other datasets available to researchers include the Annual Survey of Manufacturers, the Annual Capital Expenditures Survey, and the Business Research & Development and Innovation Survey. Moreover, the MU RDC provides access to the Longitudinal Business Database, which links establishments over time, as well as the Longitudinal Employer-Household Dynamics Database that combines firm and worker information. The aim of this presentation is to review the economic datasets available to researchers as well as how to access them at the MU RDC.

## **Further Developments of the Taxonomy of Terms and Concepts at the US Bureau of Labor Statistics**

Daniel Gillman, US Bureau of Labor Statistics

The taxonomy and lexicon project at the US Bureau of Labor Statistics (BLS) began in summer 2013 with the goal to provide consistent access to BLS time series data and documents. A team consisting of experts from most of the offices within BLS was formed to conduct the work. The end result is the taxonomy will support searching for data and provide guidance for website redesign, while the lexicon will support tagging, cataloging, and searching for documents. The technical language used by BLS is sometimes misunderstood by users. For instance, users searching for data using the word “Cities” might not locate data labeled with the technical term “Metropolitan Statistical Areas.” The main problem is the identification and organization of all BLS technical terms into a useful terminology system. This includes identifying plain English words, such as “Cities,” users often employ in their searches for BLS data. The team has worked in phases. In the first two of these, identification of plain English words was completed, and a first draft taxonomy was built. The plain English were incorporated into the titles of the categories at the top levels of the taxonomy. Additionally, the taxonomy was divided into facets, one addressing the measures BLS produces, and the other supporting the so-called characteristics, or classifications, used to stratify or sub-divide the measures. Some of the most difficult work was in organizing the measures under the higher level categories. In the third phase, the taxonomy underwent quality improvement. The results of this work were given to the program offices in BLS for review. Cognitive testing of the taxonomy with people outside the agency is planned. The paper includes a detailed discussion of the work in each of the phases of development with special attention given to the quality improvement steps.

## **Rethinking How We Teach Macro: Bring FRED Into the Classroom**

Ryan Herzog, Gonzaga University

The macroeconomy and the tools to study it have changed greatly over the last decade. Unfortunately, the teaching of the macroeconomy hasn't. We need to rethink how we teach macroeconomics. FRED offers instructors and students an easy way to simultaneously learn about data analysis and principles of macroeconomics. FRED provides user-friendly graphing tools, dashboards, and widgets that make it easy for instructors to present the most current data and create interactive data assignments. I will provide sample lectures and assignments that utilize the most current FRED data for principles of macroeconomics. Further, I will present on advanced programming techniques (via Stata and R) that utilize the FRED API to automatically recreate graphs that can be directed linked in a presentation.

## **How to Deal With Some Commonly Encountered Financial Data Research Problems**

Todd Hines, Princeton University

When doing financial economics research there are many instances where a deeper understanding of the data is required in order to perform accurate and complete research.

This presentation will cover several problems that occur frequently in financial research and make recommendations on how to solve them. When applicable, specific sources will be recommended, with an emphasis placed on using widely- held databases and sources. Areas covered will include: explaining and recognizing header data, linking data sets; using industry code data; some major issues when doing non- US fundamentals research and constructing bond data sets.

Attendees will gain a better understanding of how to deal with a number of potential pitfalls commonly encountered when doing financial economics research.

## **ReplicationWiki - Improving Transparency in Economic Research**

Jan H. Höffler, University of Göttingen

In empirical economics, a twofold lack of incentives leads to chronic problems with replicability: For authors of empirical studies providing replicable material is not awarded in the same way as publishing new irreproducible studies is. Neither is authoring replication studies.

We set incentives for replicability and replication. We gave replication seminars at several faculties internationally. Integrating replication in the education of young scholars ensures that a big number of scientists get incentives to write replication studies: credit points and the prospect of publications of working papers already during their time as students.

Our wiki documents the results of our replications and of those found in the literature. It includes a database of more than 2000 empirical studies. For each study we provide information about the availability of material for replication. This helps to identify examples for courses focusing on empirical methods or on particular topics. We provide an overview of journals and their policies regarding data availability and publication of replications.

After exploring several dozen studies published in highly ranked journals, we have hardly determined any cases where we see replicability is fully ensured. We identified two main problems: Not all published results can be obtained from the replication material provided. Information about how the used data were obtained from the raw data is rarely sufficient.

The information we collected on data sources and their geographic origin indicates that the economic literature is dominated by US data. This can harm the process to determine economic policies adequate for other regions of the world as well as the learning process from knowledge that can be acquired from their experience.

Finally, we provide information on data and code archiving policies of economics journals and show how they are related with citations.

## **Curating Data for Reproducibility**

Cynthia Hudson-Vitale, Washington University in St. Louis

Research reproducibility and rigor are of increasing concern among many disciplines given growing federal, funder, publisher, and community pressures. When policy and social decisions are made upon research, the transparency and accessibility of the data and research are of significant importance to verify claims. To address this need, many organizations and institutions have developed services to support the robust curation of data so that replication may be achieved. This presentation will present a case study of services and policies currently in place at Washington University in St. Louis. Audience members will gain practical knowledge and 'how-to' on best practices for curating digital materials for the long term and specifically for replication. The aim of this presentation is to spread the adoption of appropriate curatorial practices for research reproducibility.

## **Scanning for Scholarship: Digitization of State and Local Government Publications to Support New Frontiers of Original Research**

Kris Kasianovitz, Stanford University Libraries

Libraries in the U.S. have been collecting state and local government publications for years well before the Internet age, through state depository programs, mailing lists, or visiting agency offices in person. Libraries are replete with rich historic collections of fifty-state publications. We have content, in some cases hundreds of years' worth, locked up in print containers, often with inadequate metadata. Through digitization projects, we can unlock these troves for our researchers and the public. Digitization can create new, previously untapped data sets at the state and local levels. However, there are some costs and roadblocks to accomplishing this.

This twenty-minute session will provide attendees with an overview of policy and technical issues involved with assembling a data set from digitized state and local government publications. Specifically, the presentation will cover: coordinating with multiple libraries and archives to assemble a complete run of documents; addressing inherent copyright issues with state government publications and working with HathiTrust to release these documents to the public; dealing with tabular data in pdfs; providing data curation support to researchers in order to preserve and archive this material in order to support data sharing and reproducible research. Two digitization projects will serve as case studies for this talk: State Banking Reports and the Municipal Yearbook.

## **Organizing Economic Information: An Overview of Application and Reuse Scenarios of an Economics Knowledge Organization System**

Andreas Oskar Kempf, ZBW - German National Library of Economics - Leibniz Information Centre for Economics

Recent changes in the overall information environment have led to an unprecedented opening of traditional information systems such as knowledge organization systems (KOS) like thesauri. In former times, they were mostly published only in print or as proprietarily formatted data files imported into library catalog silos. With the publication of thesauri on the web, they made a huge transition from formerly application-specific traditional documentation languages to highly flexible representation systems which could be easily reused by third parties. In 2009 the STW Thesaurus for Economics released by the ZBW — German National Library of Economics — Leibniz Information Centre for Economics was published on the web. In this way, it has become a publicly available web source, downloaded by individuals as well as institutions world-wide to organize economic information and research data. From 2010 to 2015, the STW was completely overhauled to adapt to the current terminology in economics and business studies.

In our presentation we would like to give an overview of how the STW, as a fine-grained and increasingly interoperable vocabulary, could be reused — as a whole or in parts. This mainly includes three different indexing scenarios, first by information professionals, second by users and third by computers, based on machine learning algorithms. In addition, we would like to present how the STW can be used for information retrieval support, inter-vocabulary mapping, and linked data. We would like to conclude by demonstrating how users can get involved in the development of the thesaurus, and thus contribute to a further internationalization of the terminology.

## **Teaching with Data and Interacting with FRED**

Diego Mendez-Carbajo, Illinois Wesleyan University

Data analysis skills are central to an economic education. These skills range from performing high-order statistical operations to mid-order skills such as building and reading data plots. All data analysis skills are used to build arguments and we posit that whereas instruction in statistics is delivered through a course-specific curriculum the efforts to build comfort and competency in working with numerical data are generally unfocused and vague. Moreover, we argue that in order for our students to be able to “work with data” in sophisticated and complex ways they also need to develop a spectrum of abilities, practices, and habits of mind related to information literacy. It is the effective integration of information literacy skills and numeracy that allows students to truly “think like an economist.”

This hands-on workshop will illustrate how to promote best practices and avoid common missteps in using economic data through a new method of instruction in the visualization of economic data and information based on the Federal Reserve Bank of St. Louis Federal Reserve Economic Database (FRED).

Workshop participants will sequentially take on the roles of students and instructors. By the end of the workshop participants will be able to identify new avenues to introduce data in the instruction of Economics and to design effective pedagogies reflecting three different information literacy frames. Starting with the notion of “research as inquiry” we will emphasize the strategic aspects of searching a data aggregator and highlight the value of information by reflecting on its social and historical context.

## **Using Stata for Data Work**

James Ng, University of Notre Dame

Stata is a leading statistical software package in the social sciences, including economics. Although not free, it has many of the hallmarks of open source software such as a user-contributed repository of add-on modules, an active community of users, and numerous third party-run online guides and tutorials. Stata arguably strikes perhaps the best balance between sophistication and usability among all statistical software packages.

This hands-on workshop will introduce participants to some of the ways Stata is used in empirical research in the social sciences. Participants will work through a series of exercises using data in commonly encountered formats. Many of the exercises will involve reproducing tables and graphs from scratch. Topics to be covered include reading data, cleaning data, manipulating data, combining data, and using the help system. Attention will be paid to reproducibility of results, which means that participants will be writing scripts in a do-file. Detailed notes will be provided to each participant for reference.

This workshop's target audience is social science librarians and other data service professionals. By the end of the workshop, participants should have gained enough familiarity with Stata to be able to start using it independently and to provide more in-depth help to their patrons who use Stata.

This is not a workshop in statistical methods, hence no knowledge of statistics is assumed. No knowledge of programming is required.

## **FRED Schema : Metadata for Aggregate Economic Time Series**

Genevieve Podleski, Federal Reserve Bank of St. Louis

Metadata schemas for data are generally highly complex and often focus on data transfer and/or storage. The St. Louis Fed's FRED Schema (FRED-S) was built for use with the FRED data site, but has broader use for user-centric description of economic time series data. This presentation will explain the development of FRED-S, how it differs from other major schemas used for data, including Dublin Core, DDI, and SDMX, and will preview the next steps in metadata implementation for FRED. Attendees will gain a deeper understanding of the organization of data within FRED and learn more about the challenges of descriptive metadata for time series.

## **History of the Federal Reserve Board Statistical Releases**

Sian Seldin, Research Library, Federal Reserve Board

Many people are familiar with the Federal Reserve Board's role in conducting monetary policy, supervising and regulating banking institutions, and maintaining financial stability. Few are aware of the Board's role as a statistical provider. Since 1913 the Board has published extensive statistical data on the U.S. economy and the banking system. The statistical releases are created from data collected from banks, by other Federal agencies, through surveys, and licensed from vendors. Four of the Board's statistical releases are principal economic indicators required by law. And all of them are freely available to the public!

The statistical releases present a wealth of time series data which economic historians and students can use in the study of topics of historical and economic importance. The challenge they present, though, is that the scope of individual statistical releases has changed over time, to address the ever-changing economic environment. This presentation discusses my research, found in the FEDS working paper, "[Federal Reserve Board Statistical Releases: a Publications History](#)" and FEDS Note "[History of the Federal Reserve Board Statistical Releases](#)" regarding the history of the Board's statistical releases, and provides guidance to help researchers understand these changes. It facilitates identification of releases and an understanding of their publication history. The session will also discuss how to locate the statistical releases in various print and digital sources, including the FRASER Federal Reserve Archive and the Federal Reserve Board public website.

## **Services for Storing and Sharing Research Data in Economics — Two Case Studies from Germany: ZBW Journal Data Archive and SowiDataNet**

Ralf Toepfer, ZBW Leibniz Information Centre for Economics

Data Sharing is rather the exception than the rule among economists: Approximately 80% of the researchers never share their data, while only 3% share data regularly in a comprehensive and transparent way. The main barrier to making data available openly seems to be the lack of formal recognition. Data sharing will only be widely adopted among researchers if sharing “pays” in form of reputation. Moreover research data were stored mainly on local, institutional servers and not at repositories combined with appropriate metadata.

Against this background my talk will give an overview about the researcher’s expectations and requirements concerning an information infrastructure for research data in economics.

Based on these requirements ZBW has developed and implemented services to enable researchers to store, share, preserve and publish their research data. As show cases I will present two services: 1) ZBW Journal Data Archive: A repository for research data used in scholarly journals and 2) SowiDataNet: A research data repository for self-contained archiving and distribution of research data in economics and social sciences. My talk will end with some lessons learned.

## **Best Practice for International Economic Data Discovery: In Case of East Asian Countries**

Jungwon Yang, Clark Library, University of Michigan

Due to the development of technology and the internet, researchers can easily access international economic data. The international data retrieved from these online resources, however, is often questionable for an academic purpose, because it is quite difficult to determine the reliability of the data. International governmental organizations (IGOs), such as the United Nations, the International Monetary Funds(IMF), the Organisation for Economic Cooperation and Development (OECD), and the World Bank, provide reliable national-level economic data to public. Subnational statistics and public opinion/survey data, however, are often not available from these IGOs databases.

Even if we discover subnational data from a national government agency or other venues, it is not easy to access because of language barriers, technical problems, and copyright and license issues. In this presentation, I will explain about the characteristics of economic data of IGOs. I will, also, introduce reliable data resources, where we can find subnational economic data, public opinion, and survey data of China, Korea and Japan.

## **Are You Asking the Right Questions of Your Financial Data Providers?**

James Zussman, RateWatch Scholar

You have an idea for new research. Now you need data to begin your analysis. How do you ensure the veracity of the data source?

Researchers are besieged with sources of financial data. But how do you know you're getting the best set of data for your research needs? Based on 7 years of researcher questions and real-life user input, this presentation will review key questions researchers need to ask their potential data providers. RateWatch is a leading source of banking interest rate and fee data and we've provided data sets to dozens of academic researchers as well as government agencies such as the FDIC, Federal Reserve Board of Governors and the NCUA. Through this experience, we have cultivated a list of the most common questions asked and will explain why they are helpful when exploring your options to acquire financial data.

Before you spend money to acquire data, we will present a go-to list of questions for your potential data providers and you will learn to identify red flag responses. Lastly, we will provide an overview of the data available through RateWatch Scholar and how questions we have been asked, and conversations we've had with researchers over the years, have improved our processes and led us to become the first choice for in-depth banking data.