# Evaluating and Optimizing Homework and Quizzes to Increase Learning Outcomes in the Information Visualization MOOC

**Principle Investigator:**    Dr. Katy Börner,
Victor H. Yngve Distinguished Professor, Information and Library Science,
Director, Cyberinfrastructure for Network Science Center (CNS)
School of Informatics and Computing, Indiana University

**Assistant Investigator:**    Michael Ginda
Assistant Instructor and Senior Research Analyst, CNS
School of Informatics and Computing, Indiana University

**Student Investigator:**    Andreas Bueckle
Assistant Instructor, Information and Library Science
School of Informatics and Computing, Indiana University

**Email Contact:**

katy@indiana.edu, mginda@indiana.edu, abueckle@indiana.edu

**Funding Level Requested:**  Phase I

**Duration of Funding Period:** Spring 2017 - Fall 2018

# Abstract

The Information Visualization MOOC (IVMOOC) is a graduate level course taught each spring since 2013. It is offered for credit as part of the Information and Library Science and Data Science curriculums at IU (Z637), and offered for free to members of the public. As part of efforts to update course materials, instructors need to evaluate course activities to improve learning outcomes of students taking the course. The proposed work will analyze quizzes and homework submissions from 500 students that have taken or will take the Information Visualization course between 2016 and 2018. Student data from 2016 and 2017 will be analyzed to understand how students use and perform on course activities. Insights gained from this analysis will be used to optimize course materials for the 2018 version of the course. The revised course homework assignments and quizzes will be implemented for the 2018 course and used by two parallel cohorts. Student engagement and performance will be analyzed for 2018 and compared with the 2016 and 2017 data. Results will be written up in scholarly publications that empower other teachers to evaluate the impact of assessment quizzes and hands-on homework activities on student engagement and performance in order to improve instructional materials and learning outcomes.

## Project Description

This project will analyze quizzes and homework submissions by 500 students that have taken or will take the IVMOOC between 2016 and 2018 to improve instructional materials and learning outcomes. Statistical analysis of the 2016 and 2017 will help understand engagement and performance patterns and trends and help improve instructional materials for the 2018 version of the course. Data collected in 2018 will be analyzed to understand the impact of the revised course materials on student engagement and performance. Novel pre- and post-questionnaires measuring student data visualization literacy will be designed and implemented. Results will be written up in a scholarly publications that empower other teachers to evaluate the impact of assessment quizzes and hands-on homework activities on student engagement and performance and to optimize instructional materials and resources so they meet the needs of the students taking the course.

**Course Details:** Since 2015, the Z637 Information Visualization course and the IVMOOC have been taught using the Canvas learning management system. For the first seven weeks of the course, students enrolled in either the free IVMOOC or Indiana University's course on Information Visualization (Z637) are assigned self-assessment quizzes and homework. Quizzes assess understanding of theory while homework assignments ask students to run specific data analysis and visualization workflows testing their practical skills to design visualizations. Quizzes mostly use multiple choice and short answer questions and are peer-reviewed according to detailed rubrics provided by course instructors. Students may retake the quizzes multiple times throughout the course to earn a higher grade. Homework assignments require students to apply data analysis and visualization skills learned in the weekly hands-on portion of the course with sample data sets taken from the real world. Students are allowed to help and collaborate on homework assignments solutions through the use of the discussion forums for each assignment. Submissions are reviewed and graded against a rubric by peers. The peer review grades are then checked by instructors for accuracy. Completion of quizzes and homework count towards students' Homework and Quiz (10%) grade in the course. Student's peer-review count towards students' Class Participation grade (10%).

**Data Details:** The data for this project will be collected from two sources: the Canvas Learning Management System (LMS) front-end and the Canvas Data Product, which is accessed through Amazon Redshift. Table 1 provides an overview of data use for the project by its type and source.

*Table 1*: *Project Data Resources*

|  | Canvas LMS Front-end | Canvas Data Product |
|---|---|---|
| Student Activity & Performance Data | Course Gradebook, Quiz attempts and responses, Homework submissions (e.g., visualizations, write-up, code) | Homework discussion posts, peer review responses, student web logs |
| Instructional Data | Assignment descriptions | Assignment rubrics, Quiz questions and items |

Data generated by students during the course provides measures of their engagement with course materials and activities and their comprehension of and ability to apply knowledge to generate data visualizations. Instructional materials will be evaluated in the context of student activity and performance data to help evaluate the quality of course resources in achieving learning objectives.

## Research Methodology

The project consist of four phases:

**Phase 1** applies a statistical analysis of student activity and performance on the assessment quizzes using data created by students during the 2016 and 2017 instances of IVMOOC. The data used in this phase is taken from Canvas' front end, and includes the course gradebook, and records all assessment quiz attempts and the course's two exams. The statistical analysis consists of descriptive analysis of student activity and performance on the courses assessment quizzes, item-analysis of quiz and exam questions, and comparative analysis of student quiz behavior and performance for relevant exams.

**Phase 2** uses the results of the first phase as a guide, to revise the course's pre- and post-questionnaires that collect relevant student demographics and measure students' data visualization literacy abilities at the start and conclusion of the course. The team will also develop new sets of a revised experimental quiz questions and items, and homework assignments for the first seven weeks of the course. The quiz questions will focuses on assessing comprehension conceptual understanding of data analysis and information visualization theory and application of this knowledge taught. Homework assignments will be redesigned to assess students' ability to apply data analysis techniques in combination with visualization methods using visual literacy and accessibility standards.

**Phase 3** of the project will implement both the new pre- and post-questionnaires, updated homework assignments and rubrics, and two sets of revised weekly quizzes in Canvas LMS for use in the 2018 instance of the IVMOOC. Students in the course will be divided students into two cohorts using Canvas project groups. During the course, students from the first cohort will be assigned to use revised quizzes and homework assignments for each week of the course; the second group will serve as a control and be assigned to use the 2017 quizzes and homework assignments during the course. Quizzes and homework assignments will be locked by passwords that are assigned to each group. Each cohort will be allowed to take each quiz three times each, and keep the highest score for their final grade. Homework assignments will be submitted once by each student. Homework assignments will be open to discussion by students in the group; homework submissions will be graded by peer review. Both groups of students will be asked to complete the midterm and final exams for the course.

**Phase 4** performs an item-analysis for evaluation of the revised quiz questions; combinations of correlation, ANOVA, and regression analysis comparing the students' performance on quizzes and homework assignments with pre- and post-questionnaire measurements of data visualization literacy abilities and demographic groups, and in comparison of performance to the 2016-17 courses. Results are written up as scholarly publications.

## Significance and impact the study may have upon undergraduate learning

Many large undergraduate courses but also graduate courses that are offered by IU use Canvas to provide students homework and quiz activities in support of higher student engagement and performance. These courses have an abundance of data and a need to evaluate the impact of different activities on student learning outcomes. The methods and protocols implemented for this project will be disseminated for reuse or modification by other instructors and departments at IU to help their own efforts to evaluate and improve their courses and student learning outcomes.

## Anticipated outcomes

The proposed project will analyze students' use of quizzes and homework to measure the impact of course activities on students' engagement in the course and improvement in students' performance in terms of data visualization literacy. Data taken from the 2016 and 2017 instances of the IVMOOC will be analyzed to understand how student use and perform on homework and quizzes. Results of this analysis will help optimize learning materials for the course. While the projects primary efforts are to improve the outcomes

for the course and to advance educational data science effort in SOIC and IU, the project's data collection protocol, data analytics methods, and results will be disseminated widely online and be written up as journal or conference publications. We also plan to organize workshop presentations on applying learning analytics analysis and controlled interventions to help improve courses and curriculums at Indiana University.

## Means by which you will measures the success of your project

Statistical analysis of 2016-2017 student engagement and performance data.

Development and implementation of optimized quiz and homework assignments for 2018 course; data collection using two comparable cohorts.

Documentation of research results and learning outcomes in scholarly publications.

Dissemination of results via publications, talks, and workshops.

## Publication of project results via a journal and/or conference paper

Data collection protocol, data analytics methods, and results will be shared via scholarly publications, an online site, and a workshop that empowers other instructors at IU and beyond to perform analyses of quiz and homework self-assessments data more efficiently.

## Previous research results when applicable

Emmons, Scott, Robert Light, and Katy Börner. 2016. "MOOC Visual Analytics: Empowering Teachers, Students, Researchers, and Developers of Massively Open Online Courses". *Journal of the Association for Information Science and Technology*. In Press.

Ginda, Michael, Naren Suri, Andreas Bückle, and Katy Börner. "Empowering Instructors in Learning Management Systems: Interactive Heat Map Analytics Dashboard".

# BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person.  **DO NOT EXCEED FIVE PAGES.**

NAME: Börner, Katy

eRA COMMONS USER NAME (credential, e.g., agency login): KBORNER

POSITION TITLE:  Victor H. Yngve Distinguished Professor of Information Science

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

| INSTITUTION AND LOCATION | DEGREE *(if applicable)* | Completion Date MM/YYYY | FIELD OF STUDY |
|---|---|---|---|
| University of Technology, Leipzig | ME | 1991 | Engineering/Electronics |
| University of Kaiserslautern | Ph.D. | 1997 | Computer Science |

## A.  Personal Statement

KATY BÖRNER is the Victor H. Yngve Distinguished Professor of Information Science at the Departments of Intelligent Systems Engineering and Information and Library Science, School of Informatics and Computing, Adjunct Professor at the Department of Statistics in the College of Arts and Sciences, Core Faculty of Cognitive Science, Research Affiliate of the Biocomplexity Institute, Fellow of the Center for Research on Learning and Technology, Member of the Advanced Visualization Laboratory, and Founding Director of the Cyberinfrastructure for Network Science Center (http://cns.iu.edu) at Indiana University. Visiting Professor, Royal Netherlands Academy of Arts and Sciences (KNAW), Amsterdam, The Netherlands (since 2012) and Visiting Professor and Mercator Fellow, Department of Computer Science and Applied Cognitive Science, University of Duisburg-Essen, Duisburg, Germany (since 2015). She is a curator of the Places & Spaces: Mapping Science exhibit (http://scimaps.org). Her research focuses on the development of data analysis and visualization techniques for information access, understanding, and management. She is particularly interested in the study of the structure and evolution of scientific disciplines; the analysis and visualization of online activity; and the development of cyberinfrastructures for large scale scientific collaboration and computation. Börner has authored or co-authored more than 200 peer-reviewed publications and has presented her sole author and collaborative work at more than 500 academic, industry, and government events including the World Economic Forum, TEDx, TTI/Vanguard, Good Experience Live (GeI), and the Cannes Film Festival. Börner became an American Association for the Advancement of Science (AAAS) Fellow in 2012.

## B.  Positions and Honors

| | |
|---|---|
| 1991-1994 | Research Assistant in BMBF funded projects GOSLER and FABEL, University of Technology Leipzig, Department of Mathematics and Computer Science |
| 1992-1995 | Ph.D. Fellowship Studienstiftung des Deutschen Volkes e.V., Germany, Altstipendiat Konrad Adenauer Stiftung e.V., Germany |
| 1992-1998 | Professional Services (Technical Development) |
| 1993 | Development Engineer, Horiba LTD, Automotive Instruments Development Department, Kyoto, Japan, July / Aug. 1993 |
| 1994 | Ph.D. Studies, Boston University, Department of Computer Science |
| 1994-1996 | Research Assistant in the BMBF funded project FABEL, University of Freiburg, Center for Cognitive Science |
| 1996 –1998 | Research Associate and Instructor, University of Bielefeld, Faculty of Technology |
| 1998-1999 | Visiting Assistant Professor, Computer Science Department, Indiana University |
| 1999 | Visiting Assistant Professor, School of Library and Information Science, Indiana University |

| | |
|---|---|
| 2000-2005 | Asst. Professor, School of Library and Information Science, Indiana University |
| 2000–Present | Core Cognitive Science Faculty, Indiana University |
| 2001–Present | Adjunct Associate Professor School of Informatics, Indiana University |
| 2003 | Trustees Teaching Award for the School of Library and Information Science, Indiana University |
| 2003-2004 | SBC Fellow (formerly Ameritech Fellow) |
| 2003-2004 | Pervasive Technology Laboratories Fellowship, Indiana University |
| 2003 | Outstanding Junior Faculty Award, Indiana University |
| 2005-2007 | Associate Professor School of Library and Information Science, Indiana University |
| 2007-2009 | Victor H. Yngve Associate Professor, School of Library and Information Science |
| 2009-Present | Adjunct Professor, School of Informatics and Computing, Indiana University |
| 2009-Present | Adjunct Professor, Department of Statistics, College of Arts and Sciences, Indiana University |
| 2009-2012 | Victor H. Yngve Professor of Information Science, School of Library and Information Science, Indiana University |
| 2013-Present | Victor H. Yngve Professor of Information Science, ILS, School of Informatics and Computing, Indiana University |
| 2012-Present | Visiting Professor at the Royal Netherlands Academy of Arts and Sciences, Amsterdam, The Netherlands. |
| 2015-Present | Visiting Professor (Gastprofessur) and Mercator Fellow in User-Centered Social Media, Department of Computer Science and Applied Cognitive Science, University of Duisburg-Essen, Duisburg, Germany. |

## C.  Contributions to Science

### 1. Data Analysis and Visualization Methods and Tools

Börner's research focuses on the development of data analysis and visualization techniques for improved information access, understanding, and management. In 2015, she published the Visualization Design and Sense Making (VDSM) framework that draws on prior work in psychology, cartography, statistics, information visualization, and other fields, see details and more than 300 relevant references in the *Atlas of Knowledge*. The VDSM framework was developed to empower the broadest spectrum of users from many disciplines of science to create and interpret visual representations of data that are useful and meaningful to them. VDSM has been tested and refined in Börner's 15 years of research, development, and teaching. The framework distinguishes, defines, and exemplifies the proper usage of different levels (micro to macro) and types of analysis and visualization (e.g., temporal, geospatial, topical, and network) together with different insight need types, data scale types, visualization types, graphic symbol types, graphic variable types, and interaction types.

- Börner, Katy. 2010. *Atlas of Science: Visualizing What We Know*. The MIT Press.
- Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. Cambridge, MA: The MIT Press.
- Börner, Katy, Chaomei Chen, and Kevin W. Boyack. 2003. "Visualizing Knowledge Domains". *Annual Review of Information Science & Technology* 37: 179-255.
- Börner, Katy, Richard Klavans, Michael Patek, Angela Zoss, Joseph R. Biberstine, Robert Light, Vincent Larivière, and Kevin W. Boyack. 2012. "Design and Update of a Classification System: The UCSD Map of Science". *PLoS One* 7 (7): e39464.

### 2. Mapping and Modelling Science

Science of science studies—or the study of science by scientific means—use large-scale scholarly datasets and advanced data analysis, modeling, and visualization techniques to model and map science at the individual, institutional, and global levels. Börner co-edited a special issue of *PNAS* on *Mapping Knowledge Domains* (101, Suppl. 1, April 2004) and made major contributions to the visualization of scientific domains (Börner, Chen, Boyack, 2003). She had a leadership role in the development and update of standards such as the UCSD Map of Science classification system that supports the science-location of scholarly records on a topic map of science—analogous to geolocation services that read an address and provide a latitude-longitude position on a geospatial map. The classification and science-location service can be used in return on investment analyses (e.g., to visually depict and compare and contrast science funding vs. scholarly publications/citations or college expenses vs. salaries); in support of resource allocation decisions (e.g., when examining the expertise profiles

and scholarly networks of different institutions); or to understand the diffusion of knowledge and expertise (e.g., by overlaying career trajectories over the landscape of science).

With colleagues, she proposed the very first computational model that co-evolves co-author and paper-citation networks (Börner, Maru, Goldstone, 2004). In 2011, she co-edited *Models of Science Dynamics* (Springer). In 2016, she co-organized the NSF-funded *Modelling Science, Technology, and Innovation Conference* (report is available at http://modsti.cns.iu.edu/report) and she is a co-organizer of the NAS Sackler Colloquium on *Modelling and Visualizing Science and Technology Developments* that will take place in October 2017. Börner is working on the third atlas, the *Atlas of Forecasts: Predicting and Broadcasting Science, Technology, and Innovation*.

- Börner, Katy, Jeegar Maru, and Robert Goldstone. 2004. "The Simultaneous Evolution of Author and Paper Networks". *Proceedings of the National Academy of Sciences of the United States of America* 101 (Suppl. 1): 5266-5273.
- Scharnhorst, Andrea, Katy Börner, and Peter van den Besselaar, eds. 2012. *Models of Science Dynamics: Encounters Between Complexity Theory and Information Science*. Springer Verlag.
- Börner, Katy and Staša Milojević, eds. 2016. Modeling Science, Technology and Innovation. NSF Conference Report, Indiana University.
- Börner, Katy, Bruce Edmonds, Stasa Milojevic, and Andrea Scharnhorst, eds. 2016. "Simulating the Processes of Science, Technology, and Innovation". *Special Issue of Scientometrics*.

### *3. Cyberinfrastructure Design*

Visualization design but also bibliometrics, scientometric and other science of science research is dominated by small teams using proprietary data, tools, and workflows. Since 18 years, Börner's team has been developing macroscopes (see the *Communications of the ACM* paper) that empower anyone to convert data into visual insights. The open-source macroscope tools read data from social media, news, or other sources; software modules analyze and layout this data; and output interactive visualizations. Leveraging the OSGi industry standard (http://osgi.org), Börner's team implemented the Cyberinfrastructure Shell (CIShell, http://cishell.org), which supports the integration of new and existing algorithms into simple yet powerful tools. Compiling such a custom tool is as easy as assembling a personal music library: the appropriate algorithm plug-ins are copied into the plug-in directory, and then appear in the tool menu, ready for use. To date, more than 190,000 users have downloaded macroscope tools. Custom visualization web services have been designed for the National Institutes of Health in the US, the Biotechnology and Biological Sciences Research Council (BBSRC) in the UK, and the VIVO research discovery semantic web application (http://vivoweb.org) run by more than 50 institutions. In addition, Börner's team has created a suite of open data sources including the Scholarly Database (http://sdb.cns.iu.edu) with 28 million paper, patent, and grant records. Sci2 and SDB are widely used by science of science researchers and science policy makers alike to analyze, model, and visualize the structure and dynamics of science in a scientifically rigorous and replicable manner.

- Börner, Katy. 2011. "Plug-and-Play Macroscopes". *Communications of the ACM* 54 (3): 60-69.
- Börner, Katy. 2014. "Plug-and-Play Macroscopes: Network Workbench (NWB), Science of Science Tool (Sci2), and Epidemiology Tool (Epic)". In *Encyclopedia of Social Network Analysis and Mining*, edited by Alhajj, Reda, and Jon Rokne, 1280-1290. Berlin: Springer Verlag.
- Light, Robert, David E. Polley, and Katy Börner. 2014. "Open Data and Open Code for Big Science of Science Studies". *Scientometrics* 101 (2): 1535-1551.

### 4. Improving Data Visualization Literacy

In the information age, data literacy (i.e., the ability to make sense of massive amount of data and information) is as important as being able to read and write text (Börner, Heimlich, et al., 2015).

As an avid teacher, Börner has been developing cutting-edge courses and innovative learning materials for more than 20 years. In Spring 2013, she taught one of the very first massively open online courses at Indiana University, the Information Visualization MOOC (IVMOOC, see http://ivmooc.cns.iu.edu). Now offered for free but also for Indiana University credits as part of the Online Data Science Program and other M.S. programs, this course attracts and empowers students from more than 100 countries each Spring. Using the theoretically grounded yet practically useful visualization framework published in the *Atlas of Knowledge* and workflows detailed in the *Visual Insights* textbook, students learn major temporal, geospatial, topical, and network

analysis and visualization techniques and they have the opportunity to collaborate on real-world projects for a variety of clients.

In addition to developing learning materials, Börner's team has been developing visual analytics tools that make teaching 1000s of students feasible and effective. Using student activity data, these tools help track and optimize the learning trajectories of students, to make sense of evolving social networks, learning accomplishments, social media activity, etc.

*Visual Insights,* co-authored with her former staff member David E. Polley, has been featured on the Amazon Best Sellers list for Library & Information Sciences.

- Börner, Katy, and David E. Polley. 2014. *Visual Insights: A Practical Guide to Making Sense of Data*. Cambridge, MA: The MIT Press.
- Börner, Katy, Joe E. Heimlich, Russell Balliet, and Adam V. Maltese. 2015. "Investigating Aspects of Data Visualization Literacy Using 20 Information Visualizations and 273 Science Museum Visitors". *Information Visualization* 1-16.
- Börner, Katy. 2015. *Atlas of Knowledge: Anyone Can Map*. Cambridge, MA: The MIT Press.


## 5. Educational Data Science

Börner is applying advanced data mining and analysis to improve student engagement and learning in formal and informal learning environments. She is the lead-PI on $3M Educational Data Science (EDS) proposal that aims to realize the new potential of big data and advanced learning analytics for the complex, high-stakes environment of higher education. The project proposal capitalizes on and extends existing IUB research strengths and external collaborations to perform research and development needed to understand and optimize the impact of different interventions on student success at IU and beyond. Theories and methods from Cognitive, Data, Learning, Economic, and Management Sciences will be used to develop, validate, and optimize models that capture details of key student choice points; behavioral and social factors; and the value of different types of interventions. If funded, the project will set up a secure, federated precision data infrastructure that integrates student engagement and performance before, during, and after IU; inside and outside of the classroom. Precision data and custom tools will be used to validate and optimize models, study the utility of different data sources, and understand the impact of interventions. Resulting models will support data-driven decision making by students, teachers, and leadership with the overall goal of improved student engagement, performance, and success. The full proposal can be accessed at http://ovpr.indiana.edu/ear1/final-proposals.

Prior work on her team's use of data analysis and visualization as a means to empower students, teachers, researchers, and platform developers is documented in

- Emmons, Scott, Robert Light, and Katy Börner. 2016. "MOOC Visual Analytics: Empowering Teachers, Students, Researchers, and Developers of Massively Open Online Courses". *Journal of the Association for Information Science and Technology*.