Visual Analytics & Learning Analytics
in support of Data-Driven Decision Making

Katy Börner, Indiana University  @katycns

SFI Colloquium on “The Complexity of Educational Ecosystems”
Santa Fe Institute, Santa Fe, New Mexico

June 4, 2018
Visual Analytics & Learning Analytics in support of Data-Driven Decision Making

Outline:

Context
Data Driven Decision Making
Visual Analytics
Learning Analytics

Embracing Human and Machine Intelligence Symbiosis
Context
Exemplary set of IU Data Science courses, ‘Software Engineering’ jobs, and associated skills.
Job data was retrieved from LinkedIn and CareerBuilder and course data come from the IU course list. As can be seen, there are many skills (in orange) that are exclusively associated with courses or jobs; however, the skills in the middle interlink courses (in red) to jobs (in blue).
Empower students, teachers, and curriculum committee members to understand and discuss current and desirable student cohorts, key course trajectories, or the (gatekeeper) role that specific courses play. Vertically, courses are arranged into four groups based on the department offering the course. Within each vertical grouping, the nodes are sorted by the total enrollment for the course with highest values on top. Node size encodes number of students enrolled; node color denotes overall GPA for the course.
Science & Technology vs. Education/Training vs. Jobs

Need to study the **(mis)match** and **temporal dynamics** of S&T progress, education and workforce development options, and job requirements.

**Challenges:**
- Rapid change of STEM knowledge
- Increase in tools, AI
- Social skills (project management, team leadership)
- Increasing team size
Science & Technology vs. Education/Training vs. Jobs
Katy Börner, Olga Scrivner, Mike Gallant, Shutian Ma, Xiaozhong Liu, Keith Chewning, Lingfei Wu and James A. Evans

Need to study the **(mis)match** and **temporal dynamics** of S&T progress, education and workforce development options, and job requirements.

**Challenges:**
- Rapid change of STEM knowledge
- Increase in tools, AI
- Social skills (project management, team leadership)
- Increasing team size
Science & Technology vs. Education/Training vs. Jobs
Katy Börner, Olga Scrivner, Mike Gallant, Shutian Ma, Xiaozhong Liu, Keith Chewning, Lingfei Wu and James A. Evans

Need to study the (mis)match and temporal dynamics of S&T progress, education and workforce development options, and job requirements.

Challenges:
- Rapid change of STEM knowledge
- Increase in tools, AI
- Social skills (project management, team leadership)
- Increasing team size
Science & Technology vs. Education/Training vs. Jobs
Katy Börner, Olga Scrivner, Mike Gallant, Shutian Ma, Xiaozhong Liu, Keith Chewning, Lingfei Wu and James A. Evans

Need to study the (mis)match and temporal dynamics of S&T progress, education and workforce development options, and job requirements.

Challenges:
• Rapid change of STEM knowledge
• Increase in tools, AI
• Social skills (project management, team leadership)
• Increasing team size
Science & Technology vs. Education/Training vs. Jobs
Katy Börner, Olga Scrivner, Mike Gallant, Shutian Ma, Xiaozhong Liu, Keith Chewning, Lingfei Wu and James A. Evans

Study results are needed by:

- **Students:** What jobs will exist in 1-4 years? What program/learning trajectory is best to get/keep my dream job?
- **Teachers:** What course updates are needed? What curriculum design is best? What is my competition doing? How much timely knowledge (to get a job) vs. forever knowledge (to be prepared for 80 productive years) should I teach? How to innovate in teaching and get tenure?
- **Employers:** What skills are needed next year, in 5 years? Who trains the best? What skills does my competition list in job advertisements? How to hire/train productive teams?

What is ROI of my time, money, compassion?
Modeling and Visualizing Science and Technology Developments
National Academy of Sciences Sackler Colloquium, December 4-5, 2017, Irvine, CA

Rankings and the Efficiency of Institutions
H. Eugene Stanley | Albert-László Barabási | Lada Adamic | Marta González | Kaye Husbands Fealing | Brian Uzzi | John V. Lombardi

Higher Education and the Science & Technology Job Market
Katy Börner | Wendy L. Martinez | Michael Richey | William Rouse | Stasa Milojevic | Rob Rubin | David Krakauer

Innovation Diffusion and Technology Adoption
William Rouse | Donna Cox | Jeff Alstott | Ben Shneiderman | Rahul C. Basole | Scott Stern | Cesar Hidalgo

Modeling Needs, Infrastructures, Standards
Paul Trunfio | Sallie Keller | Andrew L. Russell | Guru Madhavan | Azer Bestavros | Jason Owen-Smith

nasonline.org/Sackler-Visualizing-Science
Data Driven Decision Making
Map of Scientific Collaborations from 2005-2009

Stream of Scientific Collaborations Between World Cities - Olivier H. Beauchesne - 2012

Computed Using Data from Elsevier's Scopus
The Emergence of Nanotechnology

Mapping the Nano Revolution

The emergence of nanotechnology has been one of the major scientific-technological revolutions in the last decade and is led to a structural reorganization of major fields of science. Price (1965) showed that fields of science and their development can be mapped using aggregated citations among the journals in the fields and their relevant environments.

The frames to the right show the evolving journal citation network for the years 1998-2003. Distances are proportional to cosine values between the citation patterns of the respective journals. Textual descriptions of key events during the development of Nanotechnology are given below each frame. Most notably, leading papers in Science and Nature catalyzed the breakthrough around 2000.

Changing Roles of Different Journals

The interdisciplinarity of a journal can be measured using betweenness centrality (BC)—journals that occur on many shortest paths between other journals in a network have higher BC values than those that do not. In the maps, sizes of nodes are proportional to the betweenness centrality of the respective journal in the citation networks.

From being a specialist journal in applied physics, the journal Nanotechnology obtained a high BC value in the years of the transition, so 2001. This is preceded by the "intervention" of Science. After the transition, the new field of nanotechnology is established, new journals such as Nano Letters published by the influential American Chemical Society take the lead, and a new speciality structure with low BC value journals results.

An animated sequence of this evolution is at: http://www.leydesdorff.net/journals/nanotech.

References


Pulse of the Nation: U.S. Mood Variations Inferred From Twitter

Mood Variations
A number of interesting trends can be observed in the data. First, overall daily variations can be seen in the graph, with the early morning and late evening having the highest level of tweeting. Second, geographic variations can be observed from the graph, with a higher density of tweets west, which is consistent with these states being in the front lines.

Weekly Variations
Weekly trends can be observed as well, with weekends much lower than weekdays.

About the Data and Visualization
The tweets were collected using over 30 million tweets (Sep 2009 – Aug 2009) collected by MIT-BIDS researchers, represented as density preserving cartograms. The mood of each tweet was inferred using the NRC Word List (Bradley, M., & Lang, P. J. Affective norms for English words (ANEW): Stimuli, instructions manual, and affective ratings. T. Technical report. 5.1, The Center for Research in Psychophysiology, University of California). The location data was taken from the U.S. Census Bureau, and the U.S. map was taken from National Geographic. The location’s mood inferred using the Google Maps API, and mapped into countries using PostGIS and U.S. county maps from the U.S. National Atlas. Mood scores were adjusted using Glor’s Brewer 2.

About Cartograms
A cartogram is a map in which the mapping variable (in this case, the number of tweets) is substituted for the true land area. Thus, the geometry of the actual map is altered so that the shape of each region is maintained as much as possible, but the area is scaled in order to be proportional to the number of tweets that originate in the region. The result is a density indicating map. The cartograms in this work were generated using the Carto software by Mark E. J. Newman.

Northeastern University
College of Computer and Information Science
Center for Complex Network Research

HARVARD UNIVERSITY

http://www.cos.nwu.edu/home/sml/jay/twittermood

© 2010 Alan Mislove, Sune Lehmann, Yong-Yeol Ahn, Jukka-Pekka Onnela, and James Niels Rosenquist - 2010
Check out our Zoom Maps online!

Visit scimaps.org and check out all our maps in stunning detail!
Earth – Cameron Beccario
Smelly Maps – Daniele Quercia, Rossano Schifanella, and Luca Maria Aiello – 2015
Visual Analytics - IVMOOC
Register for free: http://ivmooc.cns.iu.edu
The Information Visualization MOOC
ivmooc.cns.iu.edu

Students from more than 100 countries
350+ faculty members
#ivmooc
Data Visualization Literacy

Data visualization literacy (ability to read, make, and explain data visualizations) requires

- **literacy** (ability to read and write text, e.g., in titles, axis labels, legend),
- **visual literacy** (ability to find, interpret, evaluate, use, and create images and visual media), and
- **data literacy** (ability to read, create, and communicate data).

Being able to “read and write” data visualizations is becoming as important as being able to read and write text. Understanding, measuring, and improving data and visualization literacy is important for understanding STEAM developments and to strategically approach global issues.
Different Question Types

Find your way

Descriptive & Predictive Models

Find collaborators, friends

Terabytes of data

Identify trends

Terabytes of data

Find your way

Descriptive & Predictive Models
Different Levels of Abstraction/Analysis

Macro/Global
Population Level

Meso/Local
Group Level

Micro
Individual Level
<table>
<thead>
<tr>
<th>LEVELS</th>
<th>MICRO: Individual Level</th>
<th>MESO: Local Level</th>
<th>MACRO: Global Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>about 1–1,000 records</td>
<td>about 1,001–100,000 records</td>
<td>more than 100,000 records</td>
</tr>
<tr>
<td>TYPES</td>
<td>Statistical Analysis</td>
<td></td>
<td>Science and Society in Equilibrium</td>
</tr>
<tr>
<td></td>
<td>Knowledge Cartography</td>
<td></td>
<td>Number of scientists versus population and R&amp;D costs versus GNP.</td>
</tr>
<tr>
<td></td>
<td>page 44</td>
<td></td>
<td>page 103</td>
</tr>
<tr>
<td>WHEN:</td>
<td>Temporal Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Visualizing decision-making processes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>page 48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHERE:</td>
<td>Geospatial Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cell phone usage in Milan, Italy</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>page 52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHAT:</td>
<td>Topical Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Evolving patent holdings of Apple Computer, Inc. and Jerome Lemelson</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>page 56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WITH WHOM:</td>
<td>Network Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>World Finance Corporation network</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>page 60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

See Atlas of Science: Anyone Can Map, page 5
Needs-Driven Workflow Design

Stakeholders

Types and levels of analysis determine data, algorithms & parameters, and deployment

Validation
Interpretation

DEPLOY

Visually encode data
Overlay data
Select visualiz. type

READ
ANALYZE

Data

VISUALIZE
Needs-Driven Workflow Design

Types and levels of analysis determine data, algorithms & parameters, and deployment

Validation
Interpretation

DEPLOY

Visually encode data
Overlay data
Select visualiz. type

READ
ANALYZE

Data

Stakeholders
### Visualization Framework

<table>
<thead>
<tr>
<th>Insight Need Types</th>
<th>Data Scale Types</th>
<th>Visualization Types</th>
<th>Graphic Symbol Types</th>
<th>Graphic Variable Types</th>
<th>Interaction Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>page 26</td>
<td>page 28</td>
<td>page 30</td>
<td>page 32</td>
<td>page 34</td>
<td>page 26</td>
</tr>
<tr>
<td>categorize/cluster</td>
<td>nominal</td>
<td>table</td>
<td>geometric symbols</td>
<td>spatial position</td>
<td>overview</td>
</tr>
<tr>
<td>order/rank/sort</td>
<td>ordinal</td>
<td>chart</td>
<td>point</td>
<td>zoom</td>
<td>zoom</td>
</tr>
<tr>
<td>distributions</td>
<td>interval</td>
<td>graph</td>
<td>line</td>
<td>search and locate</td>
<td>search and locate</td>
</tr>
<tr>
<td>(also outliers, gaps)</td>
<td>ratio</td>
<td>map</td>
<td>area</td>
<td>filter</td>
<td>filter</td>
</tr>
<tr>
<td>comparisons</td>
<td></td>
<td>network layout</td>
<td>surface</td>
<td>details-on-demand</td>
<td>details-on-demand</td>
</tr>
<tr>
<td>trends</td>
<td></td>
<td></td>
<td>volume</td>
<td>history</td>
<td>history</td>
</tr>
<tr>
<td>(process and time)</td>
<td></td>
<td></td>
<td>linguistic symbols</td>
<td>extract</td>
<td>extract</td>
</tr>
<tr>
<td>geospatial</td>
<td></td>
<td></td>
<td>text</td>
<td>link and brush</td>
<td>link and brush</td>
</tr>
<tr>
<td>compositions</td>
<td></td>
<td></td>
<td>numerals</td>
<td>projection</td>
<td>projection</td>
</tr>
<tr>
<td>(also of text)</td>
<td></td>
<td></td>
<td>punctuation marks</td>
<td>distortion</td>
<td>distortion</td>
</tr>
<tr>
<td>correlations/relationships</td>
<td></td>
<td></td>
<td>pictorial symbols</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>images</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

See *Atlas of Science: Anyone Can Map*, page 24
## Visualization Framework

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>selection</td>
<td>categorize</td>
<td></td>
<td></td>
<td></td>
<td>category</td>
<td></td>
<td></td>
<td></td>
<td>categorize/cluster</td>
</tr>
<tr>
<td>order</td>
<td>rank</td>
<td>ranking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>order/rank/sort</td>
</tr>
<tr>
<td>distribution</td>
<td>distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>distributions (also outliers, gaps)</td>
</tr>
<tr>
<td>compare</td>
<td>nominal</td>
<td>differences</td>
<td></td>
<td></td>
<td>compare and contrast</td>
<td></td>
<td></td>
<td></td>
<td>comparisons</td>
</tr>
<tr>
<td></td>
<td>comparison</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>trends (process and time)</td>
</tr>
<tr>
<td>time series</td>
<td>patterns</td>
<td>time</td>
<td></td>
<td></td>
<td>process and time</td>
<td></td>
<td></td>
<td></td>
<td>geospatial</td>
</tr>
<tr>
<td></td>
<td>over time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>compositions (also of text)</td>
</tr>
<tr>
<td>quantity</td>
<td>part-to-whole</td>
<td>proportions</td>
<td></td>
<td></td>
<td>form and structure</td>
<td></td>
<td></td>
<td></td>
<td>correlations/relationships</td>
</tr>
<tr>
<td>association</td>
<td>correlate</td>
<td>correlation</td>
<td></td>
<td></td>
<td>relationships</td>
<td></td>
<td></td>
<td></td>
<td>relationship</td>
</tr>
</tbody>
</table>


Visualization Framework

<table>
<thead>
<tr>
<th>Insight Need Types</th>
<th>Data Scale Types</th>
<th>Visualization Types</th>
<th>Graphic Symbol Types</th>
<th>Graphic Variable Types</th>
<th>Interaction Types</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>categorize/cluster</em></td>
<td><em>nominal</em></td>
<td><em>table</em></td>
<td><em>geometric symbols</em></td>
<td><em>spatial position</em></td>
<td><em>overview</em></td>
</tr>
<tr>
<td><em>order/rank/sort</em></td>
<td><em>ordinal</em></td>
<td><em>chart</em></td>
<td><em>point</em></td>
<td><em>retinal position</em></td>
<td><em>zoom</em></td>
</tr>
<tr>
<td><em>distributions</em></td>
<td><em>interval</em></td>
<td><em>graph</em></td>
<td><em>line</em></td>
<td><em>form</em></td>
<td><em>search and locate</em></td>
</tr>
<tr>
<td>(also outliers, gaps)</td>
<td><em>ratio</em></td>
<td><em>map</em></td>
<td><em>area</em></td>
<td><em>color</em></td>
<td><em>filter</em></td>
</tr>
<tr>
<td><em>comparisons</em></td>
<td></td>
<td><em>network layout</em></td>
<td><em>surface</em></td>
<td><em>optics</em></td>
<td><em>details-on-demand</em></td>
</tr>
<tr>
<td><em>trends</em> (process and time)</td>
<td></td>
<td></td>
<td><em>volume</em></td>
<td><em>motion</em></td>
<td><em>history</em></td>
</tr>
<tr>
<td><em>geospatial</em></td>
<td></td>
<td></td>
<td><em>linguistic symbols</em></td>
<td></td>
<td><em>extract</em></td>
</tr>
<tr>
<td><em>compositions</em></td>
<td></td>
<td></td>
<td><em>text</em></td>
<td></td>
<td><em>link and brush</em></td>
</tr>
<tr>
<td>(also of text)</td>
<td></td>
<td></td>
<td><em>numerals</em></td>
<td></td>
<td><em>projection</em></td>
</tr>
<tr>
<td><em>correlations/relationships</em></td>
<td></td>
<td></td>
<td><em>punctuation marks</em></td>
<td></td>
<td><em>distortion</em></td>
</tr>
</tbody>
</table>

See *Atlas of Science: Anyone Can Map*, page 24
## Graphic Variable Types Versus Graphic Symbol Types

<table>
<thead>
<tr>
<th>Spatial</th>
<th>Quantitative</th>
<th>Geometric Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td></td>
<td><img src="image" alt="Point Graph" /></td>
</tr>
<tr>
<td>y</td>
<td></td>
<td><img src="image" alt="Point Graph" /></td>
</tr>
<tr>
<td>z</td>
<td></td>
<td><img src="image" alt="Point Graph" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Form</th>
<th>Quantitative</th>
<th>Geometric Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>NA (Not Applicable)</td>
<td><img src="image" alt="Symbol Representations" /></td>
</tr>
<tr>
<td>Shape</td>
<td>Qualitative</td>
<td><img src="image" alt="Symbol Representations" /></td>
</tr>
<tr>
<td>Rotation</td>
<td>Quantitative</td>
<td><img src="image" alt="Symbol Representations" /></td>
</tr>
<tr>
<td>Curvature</td>
<td>Quantitative</td>
<td><img src="image" alt="Symbol Representations" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Retinal</th>
<th>Quantitative</th>
<th>Geometric Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angle</td>
<td>NA</td>
<td><img src="image" alt="Symbol Representations" /></td>
</tr>
<tr>
<td>Closure</td>
<td>Quantitative</td>
<td><img src="image" alt="Symbol Representations" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Value</th>
<th>Quantitative</th>
<th>Geometric Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>Qualitative</td>
<td><img src="image" alt="Symbol Representations" /></td>
</tr>
<tr>
<td>Hue</td>
<td>Qualitative</td>
<td><img src="image" alt="Symbol Representations" /></td>
</tr>
<tr>
<td>Saturation</td>
<td>Quantitative</td>
<td><img src="image" alt="Symbol Representations" /></td>
</tr>
<tr>
<td>Graphic Variable Types Versus Graphic Symbol Types</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Symbol</strong></td>
<td><strong>Geometric Symbols</strong></td>
<td><strong>Linguistic Symbols</strong></td>
</tr>
<tr>
<td><strong>Point</strong></td>
<td><strong>Line</strong></td>
<td><strong>Area</strong></td>
</tr>
<tr>
<td><strong>X</strong></td>
<td>Quantitative</td>
<td></td>
</tr>
<tr>
<td><strong>Y</strong></td>
<td>Quantitative</td>
<td></td>
</tr>
<tr>
<td><strong>Z</strong></td>
<td>Quantitative</td>
<td></td>
</tr>
<tr>
<td><strong>Size</strong></td>
<td>Quantitative</td>
<td></td>
</tr>
<tr>
<td><strong>Shape</strong></td>
<td>Qualitative</td>
<td></td>
</tr>
<tr>
<td><strong>Rotation</strong></td>
<td>Quantitative</td>
<td></td>
</tr>
<tr>
<td><strong>Curvature</strong></td>
<td>Quantitative</td>
<td></td>
</tr>
<tr>
<td><strong>Angle</strong></td>
<td>Quantitative</td>
<td></td>
</tr>
<tr>
<td><strong>Closure</strong></td>
<td>Quantitative</td>
<td></td>
</tr>
<tr>
<td><strong>Value</strong></td>
<td>Quantitative</td>
<td></td>
</tr>
<tr>
<td><strong>Color</strong></td>
<td>Qualitative</td>
<td></td>
</tr>
<tr>
<td><strong>Saturation</strong></td>
<td>Quantitative</td>
<td></td>
</tr>
</tbody>
</table>

Some table cells are left blank to encourage future exploration of combinations.

---

<table>
<thead>
<tr>
<th>Graphic Variable Types Versus Graphic Symbol Types</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Geometric Symbols</strong></td>
</tr>
<tr>
<td><strong>Point</strong></td>
</tr>
<tr>
<td><strong>Spacing</strong></td>
</tr>
<tr>
<td><strong>Granularity</strong></td>
</tr>
<tr>
<td><strong>Pattern</strong></td>
</tr>
<tr>
<td><strong>Orientation</strong></td>
</tr>
<tr>
<td><strong>Gradient</strong></td>
</tr>
<tr>
<td><strong>Blur</strong></td>
</tr>
<tr>
<td><strong>Transparency</strong></td>
</tr>
<tr>
<td><strong>Shading</strong></td>
</tr>
<tr>
<td><strong>Stereo/Depth</strong></td>
</tr>
<tr>
<td><strong>Speed</strong></td>
</tr>
<tr>
<td><strong>Velocity</strong></td>
</tr>
<tr>
<td><strong>Rhythm</strong></td>
</tr>
</tbody>
</table>

---

37
Course Schedule

Part 1: Theory and Hands-On

- **Session 1** – Workflow Design and Visualization Framework
- **Session 2** – “When:” Temporal Data
- **Session 3** – “Where:” Geospatial Data
- **Session 4** – “What:” Topical Data

Mid-Term

- **Session 5** – “With Whom:” Trees
- **Session 6** – “With Whom:” Networks
- **Session 7** – Dynamic Visualizations and Deployment

Final Exam

Part 2: Students work in teams on client projects.

Final grade is based on Homework and Quizzes (10%), Midterm (20%), Final (30%), Client Project (30%), and Class Participation (10%).
The IVMOOC Companion Textbook

This textbook offers a gentle introduction to the design of insightful visualizations. It seamlessly blends theory and practice, giving readers both the theoretical foundation and the practical skills necessary to render data into insights.

The book accompanies the Information Visualization MOOC that attracted students, scholars, and practitioners from many fields of science and more than 100 different countries.

http://ivmooc.cns.iu.edu

cns.iu.edu/ivmoocbook14.html
IVMOOC App

The “IVMOOC Flashcards” app can be downloaded from Google Play and Apple iOS stores.
Sci2 Tool Interface Components Implement Vis Framework
Download tool for free at http://sci2.cns.iu.edu
Load **One** File and Run **Many** Analyses and Visualizations

<table>
<thead>
<tr>
<th>Times Cited</th>
<th>Publication Year</th>
<th>City of Publisher</th>
<th>Country</th>
<th>Journal Title (Full)</th>
<th>Title</th>
<th>Subject Category</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>2011</td>
<td>NEW YORK</td>
<td>USA</td>
<td>COMMUNICATIONS OF THE ACM</td>
<td>Plug-and-Play Macroscopes</td>
<td>Computer Science</td>
<td>Borner, K</td>
</tr>
<tr>
<td>18</td>
<td>2010</td>
<td>MALDEN</td>
<td>USA</td>
<td>CTS-CLINICAL AND TRANSLATIONAL SCIENCE</td>
<td>Advancing the Science of Team Science</td>
<td>Research &amp; Experimental Medicine</td>
<td>Falk-Krzesinski, HJ</td>
</tr>
</tbody>
</table>
Load **One** File and Run **Many** Analyses and Visualizations

<table>
<thead>
<tr>
<th>Times Cited</th>
<th>Publication Year</th>
<th>City of Publisher</th>
<th>Country</th>
<th>Journal Title (Full)</th>
<th>Title</th>
<th>Subject Category</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>2011</td>
<td>NEW YORK</td>
<td>USA</td>
<td>COMMUNICATIONS OF THE ACM</td>
<td>Plug-and-Play Macroscopes</td>
<td>Computer Science</td>
<td>Borner, K</td>
</tr>
<tr>
<td>18</td>
<td>2010</td>
<td>MALDEN</td>
<td>USA</td>
<td>CTS-CLINICAL AND TRANSLATIONAL SCIENCE</td>
<td>Advancing the Science of Team Science</td>
<td>Research &amp; Experimental Medicine</td>
<td>Falk-Krzesinski, HJ</td>
</tr>
</tbody>
</table>

Co-author and many other bi-modal networks.
Learning Analytics
Learning Analytics

**Empowering Teachers:** How to make sense of the activities of thousands of students? How to guide them?

**Empowering Students:** How to navigate learning materials and develop successful learning collaborations across disciplines and time zones?

**Empowering Researchers:** How do people learn? What pedagogy works (in a MOOC) and when?

**Empowering MOOC Platform Designers:** What technology helps and what hurts?
Visualizing IVMOOC Data

Data was collected from different sources:
• 1,901 students registered via GCB (1215 male/557 female)
• 52,557 slide downloads from our server
• 18,893 video views via YouTube
• 193 accounts made 730 tweets
• 134 students took 183 exams in GCB
• 674 remarks on 215 different forum threads in Drupal
• 64 students submitted projects via Drupal
Learning Analytics

**Empowering Teachers:** How to make sense of the activities of thousands of students? How to guide them?

**Empowering Students:** How to navigate learning materials and develop successful learning collaborations across disciplines and time zones?

**Empowering Researchers:** How do people learn? What pedagogy works (in a MOOC) and when?

**Empowering MOOC Platform Designers:** What technology helps and what hurts?
Visualizing IVMOOC Data

Data was collected from different sources:

- 1,901 students registered via GCB (1215 male/557 female)
- 52,557 slide downloads from our server
- 18,893 video views via YouTube
- 193 accounts made 730 tweets
- 134 students took 183 exams in GCB
- 674 remarks on 215 different forum threads in Drupal
- 64 students submitted projects via Drupal
Student Registration and Activity

- Jan. 22: Course Starts
- March 11: Final Exam Deadline
Student Registration and Activity
Student Registration and Activity

1215 male students
557 female students
Student Registration and Activity

1215 male students
557 female students
Student Registration and Activity

Novice IVMOOC Student Activity

Activity Types:
- Registration
- Exam
- YouTube
- Twitter
Student Registration and Activity

Expert IVMOC Student Activity

Activity Types:
- Registration
- Exam
- YouTube
- Twitter

Date:
- Jan
- Mar
- May
Student Client Projects: All Interactions
Custom interactive visualizations of IVMOOC student engagement and performance data, explore functionality online at [http://goo.gl/TYixCn](http://goo.gl/TYixCn)
Figure 1: Analysis types vs. user needs.

Next Generation IVMOOC

**Instructor:** Victor H. Yngve Distinguished Professor Katy Börner & CNS Team, ISE, SICE, IUB

**Duration:** 6 weeks x 5 hours = 30 hours (3 CEUs)

**Format:** Online | Theory and Hands-on Instruction, Concept Questions, Graded Assignments, Case Studies, Discussions

**Start:** Sept 15, 2018

**Covers:**
Temporal, geospatial, topical (linguistic), network analyses and 60+ visualization types

**Tools:** Tableau, Gephi, BI,

**Industry case studies** such as
- Acting on customer complaints data.
- Improving communication/traffic flows.
- Understanding web page usage.
- Visualizing online shopping behavior.
- Optimizing supply chains.
- Reducing customer/supplier churn.
- Monitoring emerging R&D areas.
- Workforce development planning.
Next Generation IVMOOC

Systematic study of how different student cohorts learn best—using Mechanical Turk formal user studies and extensive learning analytics.

Optimization of **Data Visualization Framework** and **Learning Modules**.
Next Generation IVMOOC

Systematic study of how different student cohorts learn best—using Mechanical Turk formal user studies, e.g., to optimize horizontal transfer:

**Table**
Columns by rows

<table>
<thead>
<tr>
<th>column</th>
<th>row</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>y</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>19</td>
</tr>
<tr>
<td>6</td>
<td>27</td>
</tr>
<tr>
<td>8</td>
<td>35</td>
</tr>
</tbody>
</table>

**Graph**
x-y coordinates
linear/log scale

**Map**
Latitude/longitude

**Network**
Local similarity

Horizontal Transfer
Next Generation IVMOOC

Systematic study of how different student cohorts learn best—using **Learning Analytics** to optimize scaffolding and learning trajectories:

- **Insights**
  - Trends, clusters, outliers

- **Interactivity**
  - Zoom, filter, details-on-demand

- **Graphic variables**
  - Position, color, size, shape code

- **Graphic symbols**
  - Geometric, linguistic, or pictorial

- **Visualizations**
  - Table, graph, map, network

- **Insight Needs & Data**

**Workflow Design**

**Report**
- Insights

**Do select best type(s) of**
- Graphic variables
- Graphic symbols
- Visualizations

**Given**
- Data & Insight Needs
MIT, Boeing, NASA, and edX to launch online architecture and systems engineering program

Four-course program will train professionals in latest practices on models and methods to manage complex systems

https://sysengonline.mit.edu
MIT xPRO Course
“Architecture of Complex Systems” delivered via the edX platform in Fall 2016.

**1,611 Boeing engineers** registered; 1,565 were active and generated nearly **31 million click event records** while accessing videos, projects, and assessments. Some students generated over 100,000 separate events.

All but 255 engineers passed the course, resulting in a completion rate of 84.1%.
Figure 1: Learner path overlaid on linear sequence of course modules. Linear, temporal sequence of learning modules accessed by a high performing student plotted from left (first) to right (last) with dividing lines for pre and post but also week 1-5 modules.
Figure 2. Zoom into learner path overlaid on linear sequence of course modules. Linear, temporal sequence of course modules used by a high performing student; plotted from left (first) to right (last) with dividing lines for different module sections.
Improving Return on Investment in Education: Measuring, Visualizing, and Optimizing Learner Trajectories

Michael C. Richey, Michael Ginda, Mark Cousino, Katy Börner
Figure 3: Learner path overlaid on force-directed layout of used course modules. Learner path of a student with high (left) and low (right) performance scores overlaid on force-directed layout of course modules.
Additive Manufacturing for Innovative Design and Production

A 9-week online course on creating new products, processes, and business models using 3D printing.

https://additivemanufacturing.mit.edu
Students use Onshape to practice what they learned
Students use Onshape to practice what they learned
Additive Manufacturing

3-modal network of all Students (blue), Teams (green), Documents (orange) used in course by May 31, 2018.

Area size represents the total time associated with a given node in the modeling software.

- Top Student - 16.91 hours
- (b524) Project - 7.74 hours
- Hanger Design Space Document - 367.28 hours

Edge thickness denotes number of times an relationship occurred in the data.
Embracing Human and Machine Intelligence Symbiosis
Visualizing the Internet of Things (IoT)

Using large scale datasets, advanced data mining and visualization techniques, and substantial computing resources.

Work by Philip Beesley | www.philipbeesley.ca | www.lasg.ca
Amatria Unveiled by Andreas Bueckle et al. Data visualizations of sensor/actuator positions and types, energy and communication flows, and emergent behavior of smart environments.
References


http://www.pnas.org/content/vol101/suppl_1

http://scimaps.org/atlas


http://scimaps.org/atlas2
All papers, maps, tools, talks, press are linked from http://cns.iu.edu
These slides are at http://cns.iu.edu/presentations.html

CNS Facebook: http://www.facebook.com/cnscenter
Mapping Science Exhibit Facebook: http://www.facebook.com/mappingscience