

Visual Analytics & Learning Analytics

Katy Börner, Indiana University @katycns

Doktorandenkolloquium "Education & Technology"
Fakultät für Erziehungswissenschaften und Medienzentrum der TU Dresden, Germany

July 5, 2018



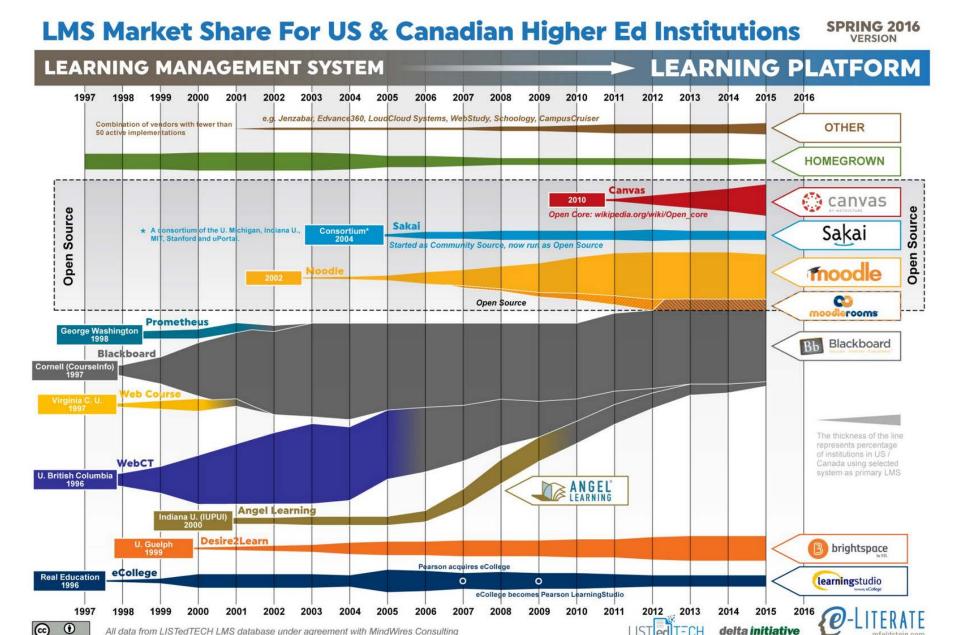
Visual Analytics & Learning Analytics

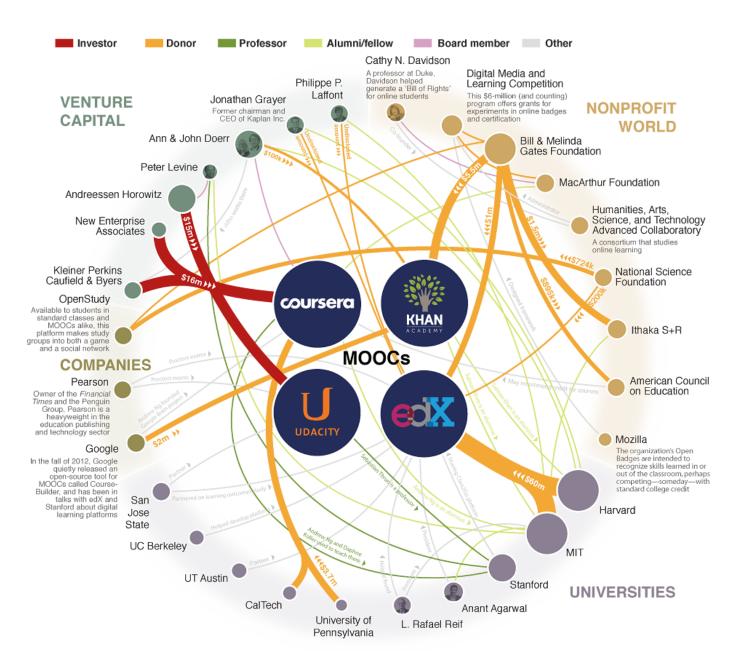
Outline:

Context
Visual Analytics
Learning Analytics



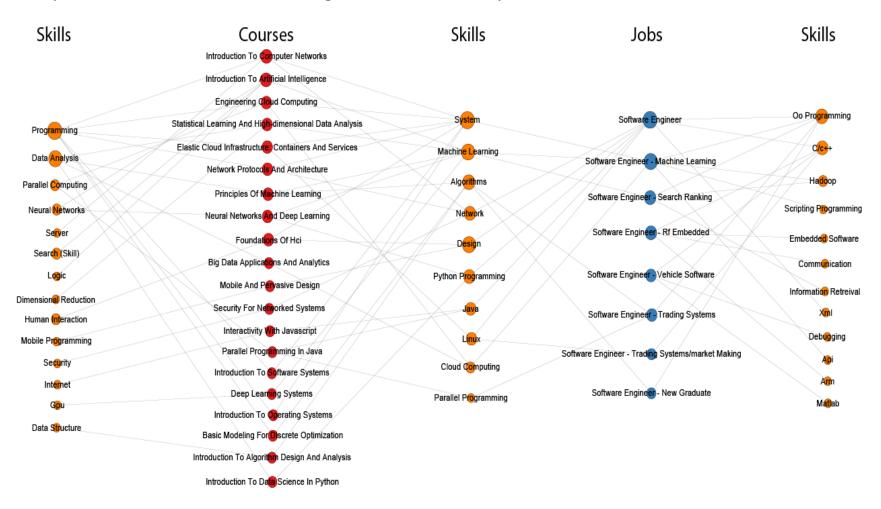
Context





IU Data Science Program: Courses, Skills & Jobs

Katy Börner, Michael Ginda & Xiaozhong Liu, Indiana University

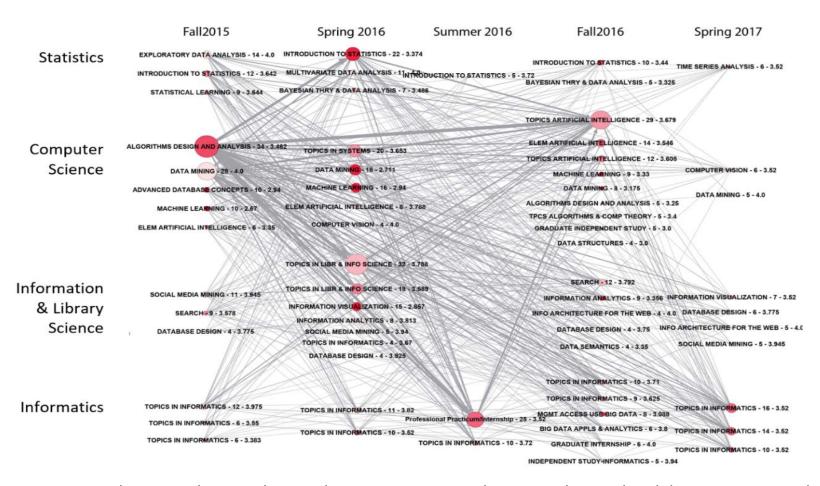


Exemplary set of IU Data Science courses, 'Software Engineering' jobs, and associated skills.

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).

IU Data Science Program: Student Course Transition Network

Michael Ginda, Kayla Scroggins & Katy Börner, Indiana University

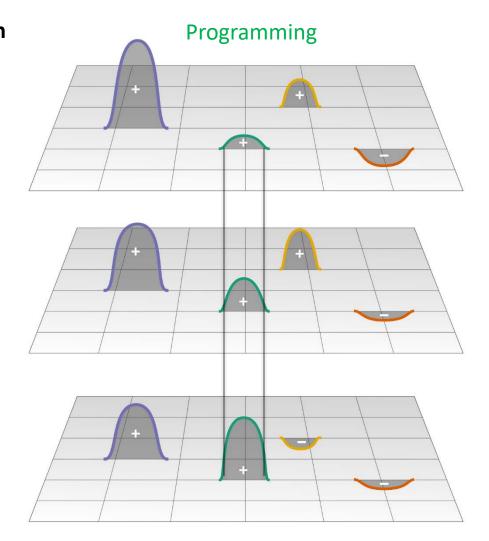


Empower students, teachers, and curriculum committee members to understand and discuss current and desirable student cohorts, key course trajectories, or the (gatekeeper) role that specific courses play. Vertically, courses are arranged into four groups based on the department offering the course. Within 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.

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.

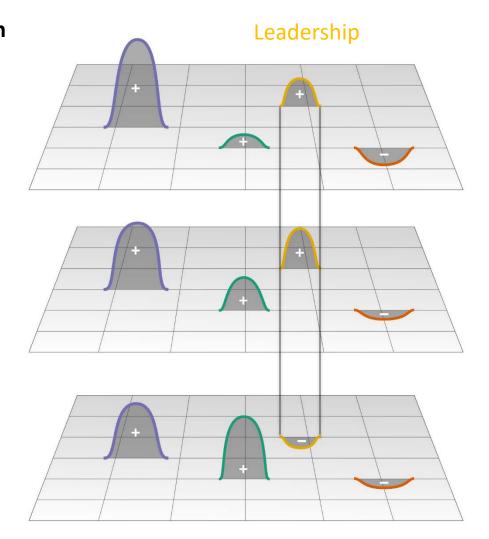
- Rapid change of STEM knowledge
- Increase in tools, Al
- Social skills (project management, team leadership)
- Increasing team size



Katy Börner, Olga Scrivner, Mike Gallant, Shutian Ma, Xiaozhong Liu, Keith Chewning, Lingfei Wu and James A. Evans

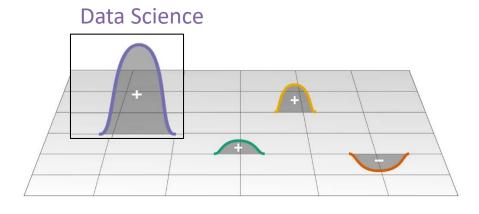
Need to study the (mis)match and temporal dynamics of S&T progress, education and workforce development options, and job requirements.

- Rapid change of STEM knowledge
- Increase in tools, Al
- Social skills (project management, team leadership)
- Increasing team size

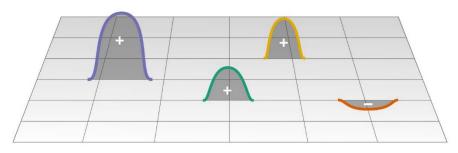


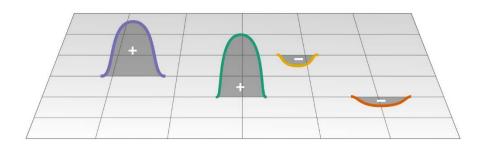
Katy Börner, Olga Scrivner, Mike Gallant, Shutian Ma, Xiaozhong Liu, Keith Chewning, Lingfei Wu and James A. Evans

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



- Rapid change of STEM knowledge
- Increase in tools, Al
- Social skills (project management, team leadership)
- Increasing team size

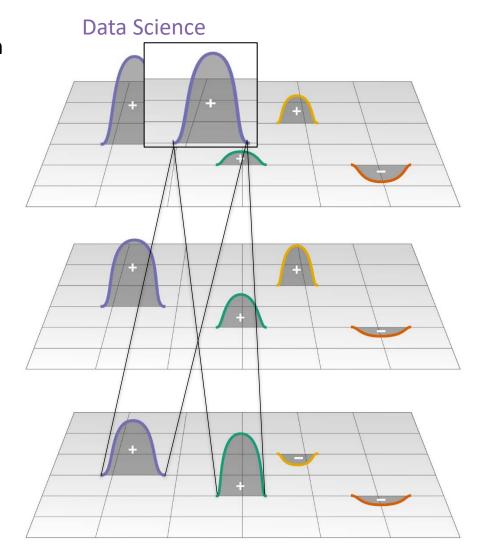




Katy Börner, Olga Scrivner, Mike Gallant, Shutian Ma, Xiaozhong Liu, Keith Chewning, Lingfei Wu and James A. Evans

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

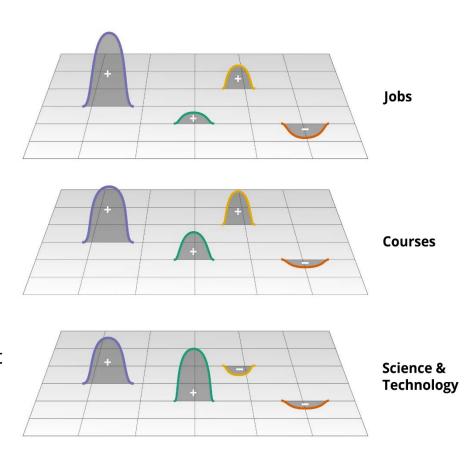
- Rapid change of STEM knowledge
- Increase in tools, Al
- Social skills (project management, team leadership)
- Increasing team size



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?



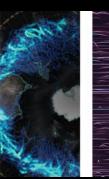
What is ROI of my time, money, compassion?

#SacklerModVisST









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







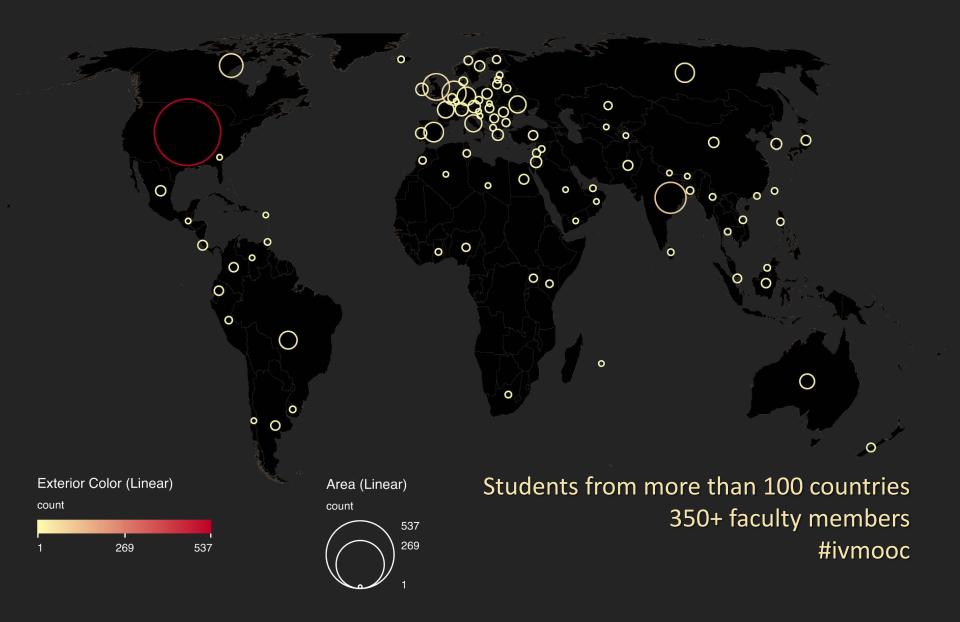


Visual Analytics - IVMOOC



Register for free: http://ivmooc.cns.iu.edu

The Information Visualization MOOC ivmooc.cns.iu.edu



Data Visualization Literacy

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

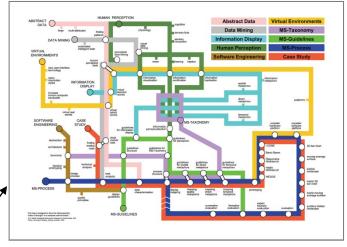
- literacy (ability to read and write text, e.g., in titles, axis labels, legend),
- visual literacy (ability to find, interpret, evaluate, use, and create images and visual media), and
- data literacy (ability to read, create, and communicate data).

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



Terabytes of data



Find your way

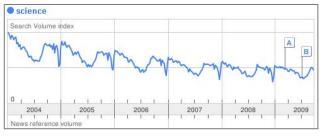
Descriptive &

Predictive

Models



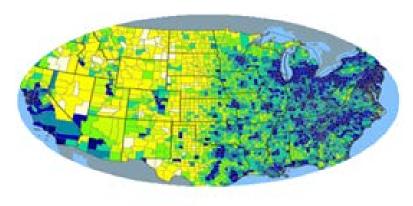
Find collaborators, friends



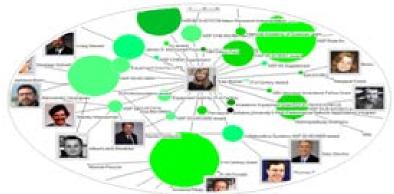
Identify trends

Different Levels of Abstraction/Analysis

Macro/Global Population Level



Meso/Local Group Level

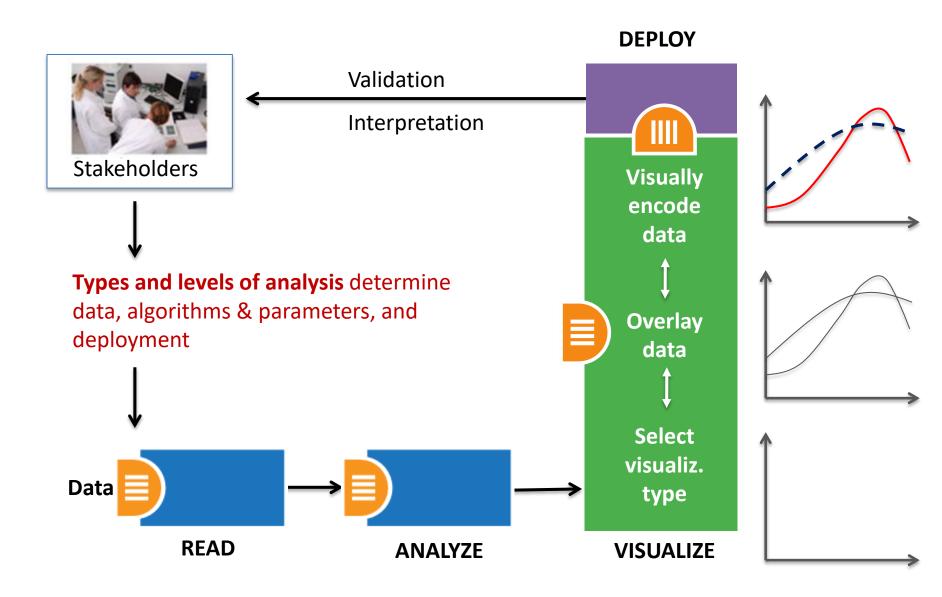


Micro Individual Level

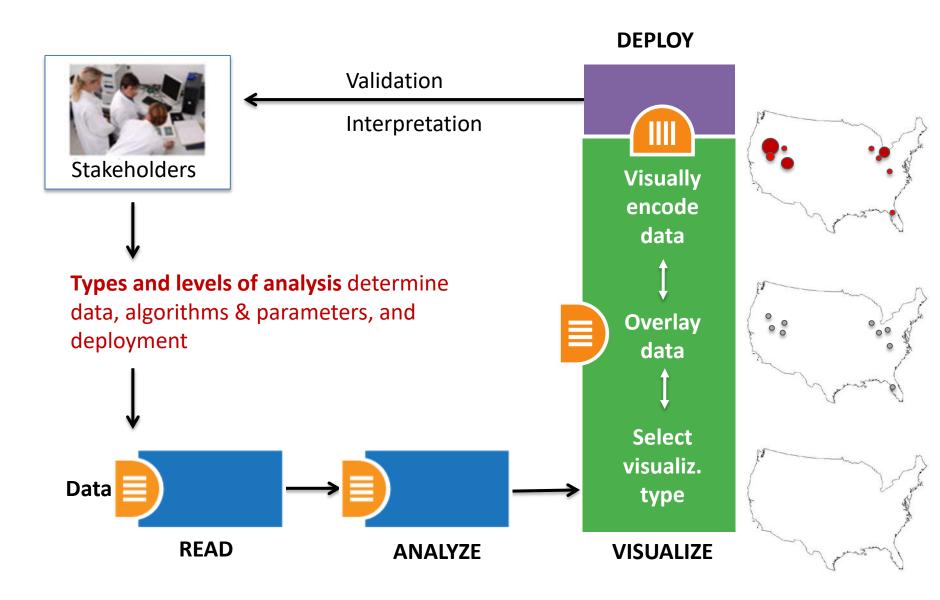


LEVELS Tasks MICRO: Individual Level MESO: Local Level MACRO: Global Level about 1-1,000 records about 1,001-100,000 records more than 100,000 records page 10 page 6 page 8 nitan dipin dipin **TYPES** Statistical Analysis page 44 Number Productivity of scientists of Russian Knowledge versus life sciences Cartography population research and R&D costs page 135 versus GNP. page 105 page 103 WHEN: Temporal Analysis Key events page 48 Visualizing Increased in the decisiontravel and development making communication of the video processes speeds tape recorder page 95 page 83 page 85 WHERE: Geospatial Analysis page 52 Cell phone Victorian Ecological usage in poetry in footprint of Milan, Italy Europe countries page 109 page 137 page 99 WHAT: Evolving **Topical Analysis** patent Product space holdings Evolving page 56 showing of Apple co-export Computer, networks in patterns of Inc. and nanotechnology countries Jerome page 139 Technology Design Research Street page 93 Lemelson A PROPERTY OF THE PARTY OF THE WITH WHOM: Network Analysis page 60 World World-wide Electronic and Finance scholarly new media art Corporation collaboration networks network networks Atlas of Knowledge page 133 page 87 page 157 Anyone Can Map

Needs-Driven Workflow Design

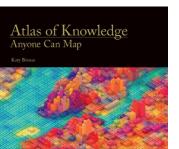


Needs-Driven Workflow Design



Visualization Framework

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

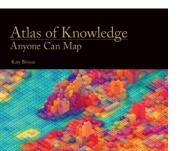


Visualization Framework

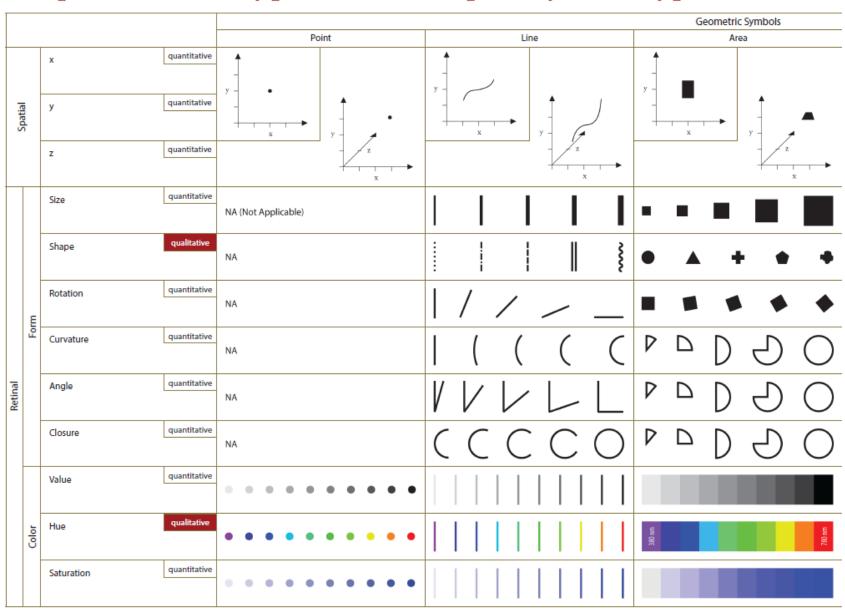
Basic Task Types								
Bertin, 1967	Wehrend & Lewis, 1996	Few, 2004	Yau, 2011	Rendgen & Wiedemann, 2012	Frankel, 2012	Tool: Many Eyes	Tool: Chart Chooser	Börner, 2014
selection	categorize			category				categorize/ cluster
order	rank	ranking					table	order/rank/ sort
	distribution	distribution					distribution	distributions (also outliers, gaps)
	compare	nominal comparison & deviation	differences		compare and contrast	compare data values	comparison	comparisons
		time series	patterns over time	time	process and time	track rises and falls over time	trend	trends (process and time)
		geospatial	spatial relations	location		generate maps		geospatial
quantity		part-to- whole	proportions		form and structure	see parts of whole, analyze text	composition	compositions (also of text)
association	correlate	correlation	relationships	hierarchy		relations between data points	relationship	correlations/ relationships

Visualization Framework

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



Graphic Variable Types Versus Graphic Symbol Types



Graphic Variable Types Versus Graphic Symbol Types Linguistic Symbols Pictorial Symbols Text, Numerals, Punctuation Marks Text quantitative quantitative See Stepped Relief Map, pages 53-54 See Elevation Map, See Helahts of the Principal NA (Not Applicable) Mountains ..., page 67 Shape See also Life in Los Angele Text Text Rotation Text Text quantitative Curvature Angle Some table cells are left blank to encourage quantitative quantitative Value quantitative Saturation Linguistic Symbols Pictorial Symbols Spacing quantitative Granularity Pattern quantitative Orientation quantitative Gradient Blur quantitative Transparency quantitative Shading quantitative Stereoscopic Depth background quantitative Speed quantitative Rhythm Blinking point Blinking area Blinking volume Blinking text Blinking icons Blinking line Blinking surface

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%).



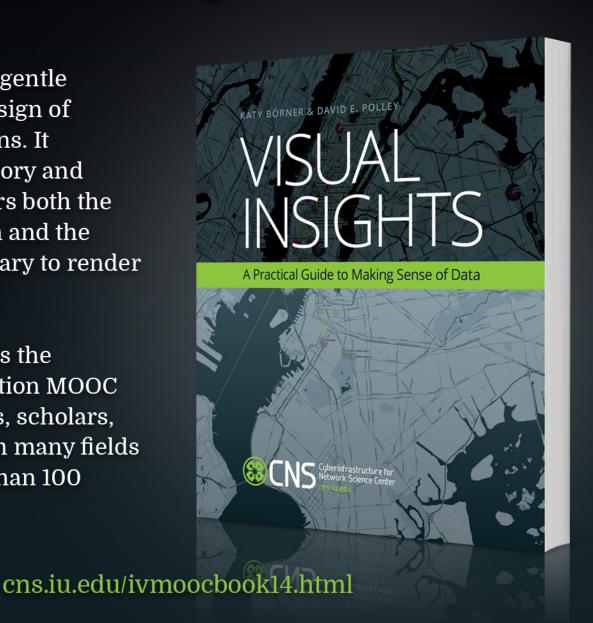


The IVMOOC Companion Textbook

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.

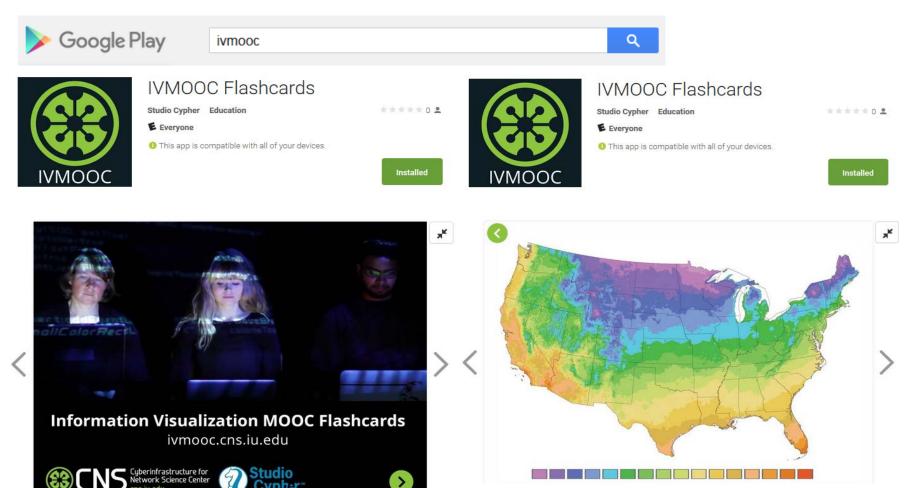
The book accompanies the Information Visualization MOOC that attracted students, scholars, and practitioners from many fields of science and more than 100 different countries.

http://ivmooc.cns.iu.edu



IVMOOC App

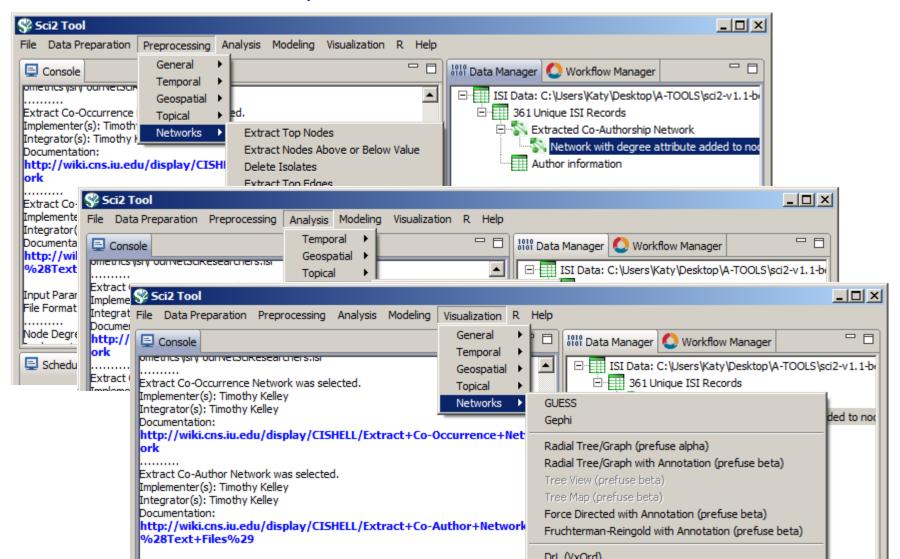
The "IVMOOC Flashcards" app can be downloaded from Google Play and Apple iOS stores.





Sci2 Tool Interface Components Implement Vis Framework

Download tool for free at http://sci2.cns.iu.edu



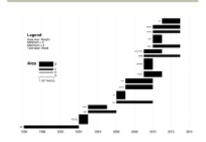
Load One File and Run Many Analyses and Visualizations

Times Cited	Publication Year	City of Publisher	Country	Journal Title (Full)	Title	Subject Category	Authors
12	2011	NEW YORK	USA	COMMUNICATI ONS OF THE ACM	Plug-and-Play Macroscopes	Computer Science	Borner, K
18	2010	MALDEN	USA	CTS-CLINICAL AND TRANSLATIONA L SCIENCE	Advancing the Science of Team Science	Research & Experimental Medicine	Falk-Krzesinski, HJ Borner, K Contractor, N Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B
13	2010	WASHINGTON	USA		A Multi-Level Systems Perspective for the Science of Team Science	Cell Biology Research & Experimental Medicine	Borner, K Contractor, N Falk- Krzesinski, HJ Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B

Statistical Analysis—p. 44

Location	Count	# Citations
Netherlands	13	292
United States	9	318
Germany	11	36
United Kingdom	1	2

Temporal Burst Analysis-p. 48



Geospatial Analysis—p. 52

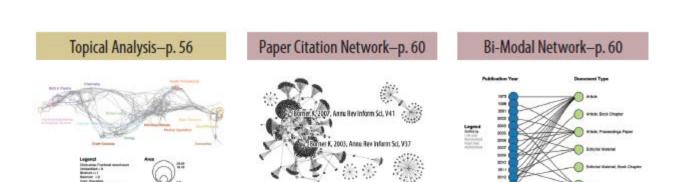


Geospatial Analysis-p. 52



Load **One** File and Run **Many** Analyses and Visualizations

Times Cited	Publication Year	City of Publisher	Country	Journal Title (Full)	Title	Subject Category	Authors
12	2011	NEW YORK	USA	COMMUNICATI ONS OF THE ACM	Plug-and-Play Macroscopes	Computer Science	Borner, K
18	2010	MALDEN	USA	CTS-CLINICAL AND TRANSLATIONA L SCIENCE	Advancing the Science of Team Science	Research & Experimental Medicine	Falk-Krzesinski, HJ Borner, K Contractor, N Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B
13	2010	WASHINGTON	USA	TRANSLATIONA	A Multi-Level Systems Perspective for the Science of Team Science	Cell Biology Research & Experimental Medicine	Borner, K Contractor, N Falk- Krzesinski, HJ Fiore, SM Hall, KL Keyton, J Spring, B Stokols, D Trochim, W Uzzi, B



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?

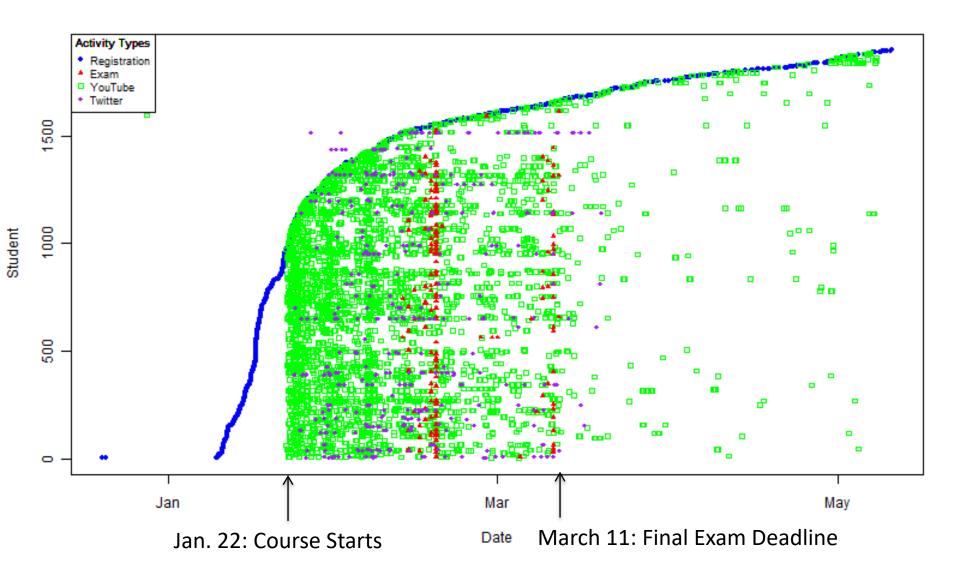


Visualizing IVMOOC Data

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

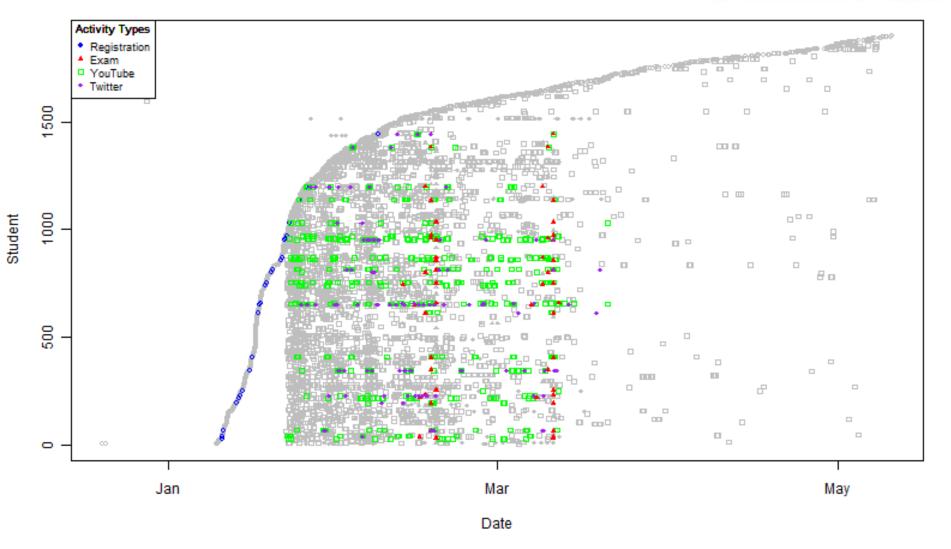






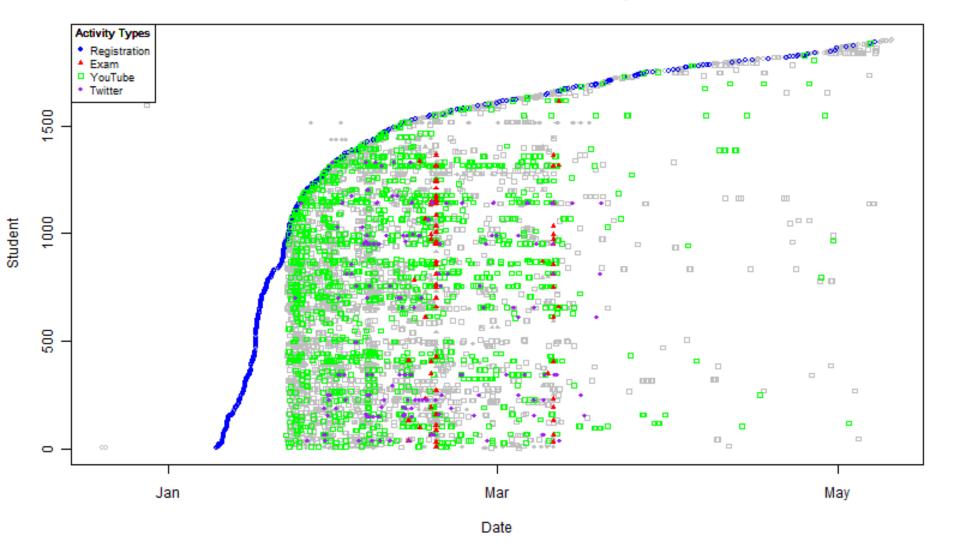


IVMOOC Student Activity (Achievement Badge)



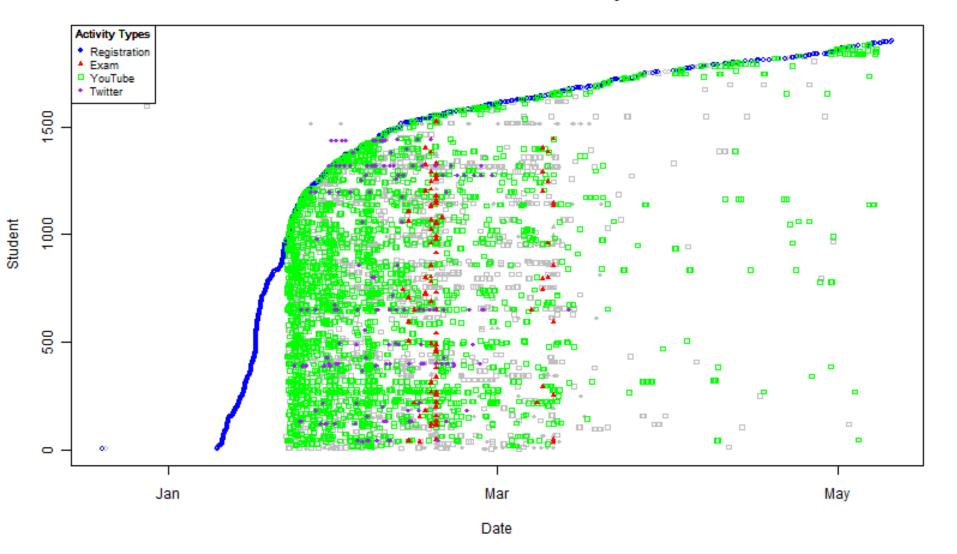
1215 male students557 female students

Female IVMOOC Student Activity



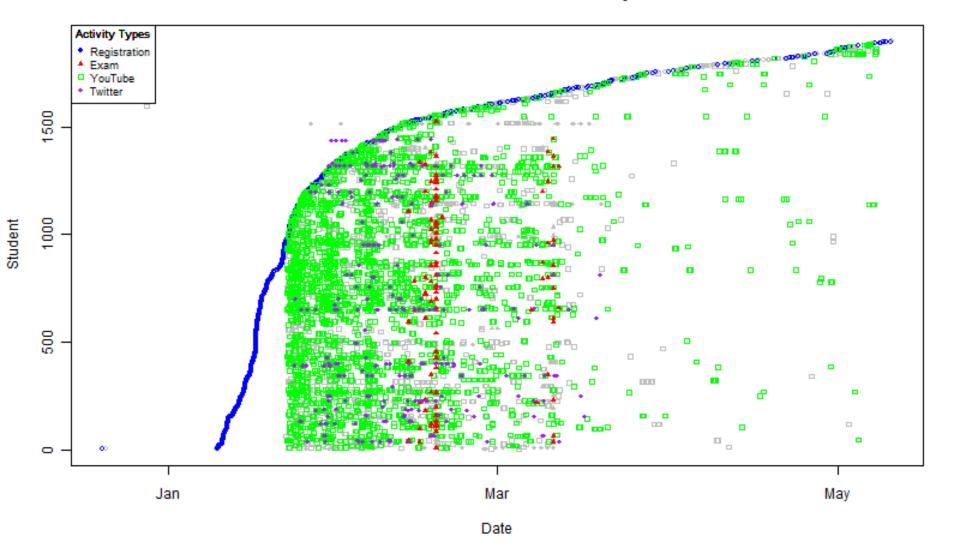
1215 male students557 female students

Male IVMOOC Student Activity



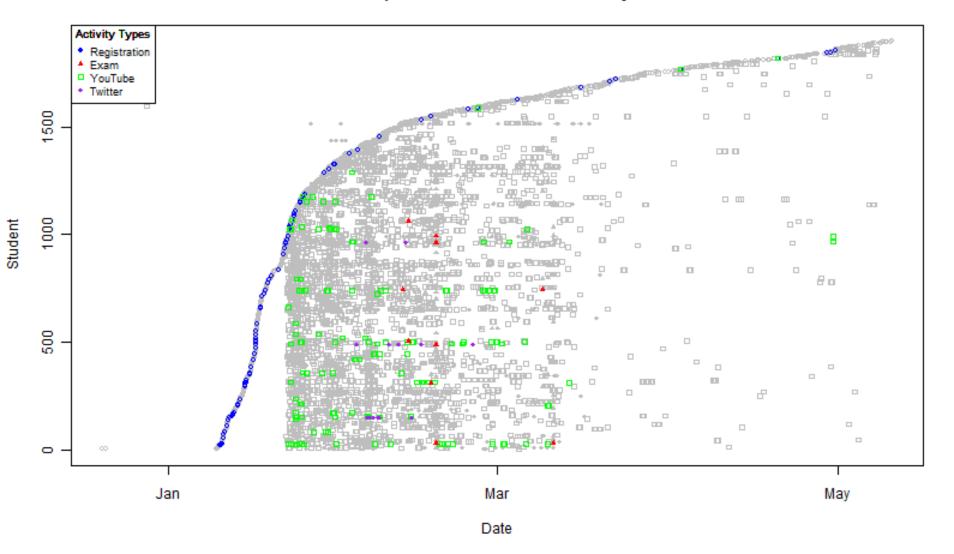


Novice IVMOOC Student Activity



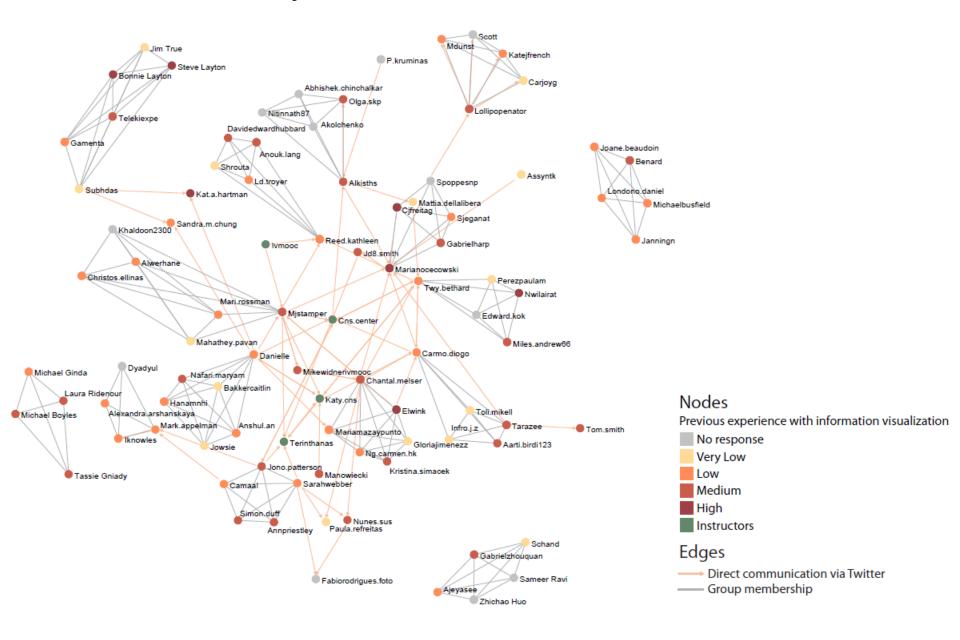


Expert IVMOOC Student Activity





Student Client Projects: All Interactions





Student Engagement and Performance

Learning Analytics

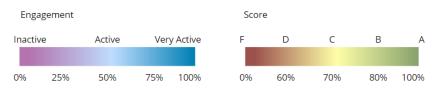
IVMOOC 2015 Student Group Engagement and Scores

	Pre-Course	Week 1	Week 2	Week 3	Week 4	Midterm	Week 5	Week 6	Week 7	Week 8	Week 9	Final	Curr. Score
IVMOOC	26.05%	38.32%	31.32%	29.96%	27.1%	28.34%	31.07%	24.28%	16.86%	18.23%	13.08%	13.41%	20.87%
Z637-29374	33.01%	52.91%	49.89%	59.22%	50.89%	82.56%	65.04%	49.99%	39.59%	61.63%	54.91%	82.25%	82.4%
Z637-32593	25.08%	54.54%	43.58%	50.67%	53.63%	77.67%	65.7%	59.48%	52.19%	65.71%	47.27%	72.59%	75.13%
Z637-33781	29.33%	55.38%	49.26%	62.18%	77.47%	85%	87.4%	69.8%	55.56%	57.6%	45.69%	70.89%	77.94%

IVMOOC 2015 Student Group Engagement for Midterm

	Midterm	Final	Curr. Score	Overall Engagemer
Student 198	100%	85.33%	92.67%	30.34%
Student 210	100%	84%	92%	33.91%
Student 242	97.14%	98.67%	97.9%	55.89%
Student 265	95.71%	92%	93.86%	82.64%
Student 216	95.71%	24%	59.86%	34.92%
Student 257	94.29%	98.67%	96.48%	68.25%
Student 264	94.29%	89.33%	91.81%	80.47%
Student 262	94.29%	85.33%	89.81%	79.65%

Legends



Description

The heat map visualization is a representation of student engagement (magenta to blue color scale) and performance (red to green color scale) throughout a course. The visualization has two levels. The top level provides an overview of engagement and performance for groups of students, while the bottom level provides a detailed break out of student engagement statistics for individuals with an identified group.

Custom interactive visualizations of IVMOOC student engagement and performance data, explore functionality online at http://goo.gl/TYixCn

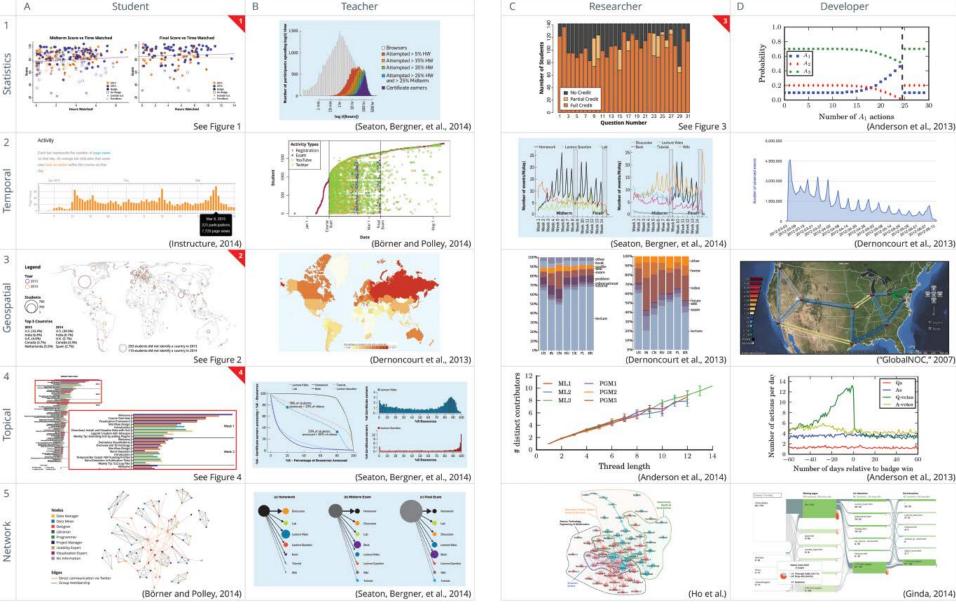


Figure 1: Analysis types vs. user needs.

Emmons, Light, and Börner. "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).

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

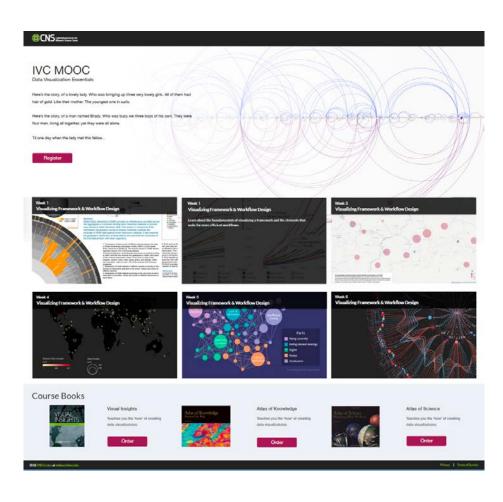
Covers:

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

Tools: Tableau, Gephi, BI,

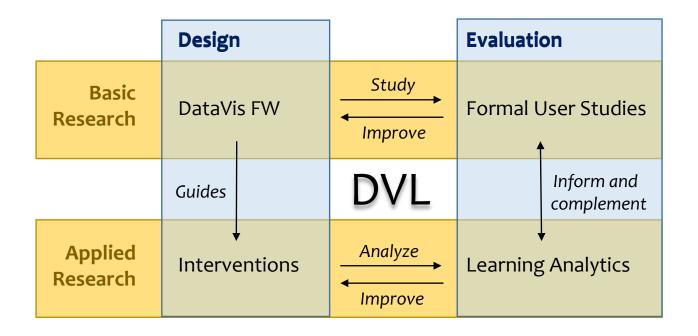
Industry case studies such as

- 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.

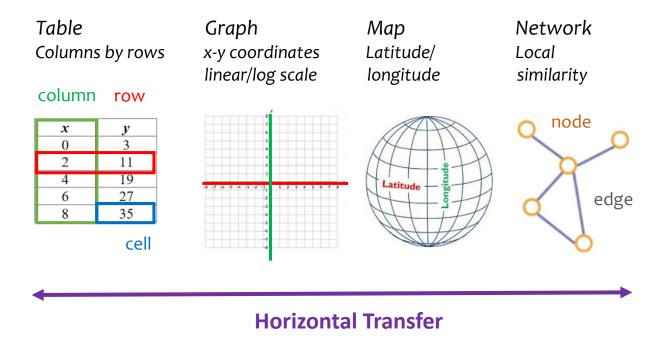


Systematic study of how different student cohorts learn best—using Mechanical Turk formal user studies and extensive learning analytics.

Optimization of Data Visualization Framework and Learning Modules.



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



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

Insights Trends, clusters, outliers

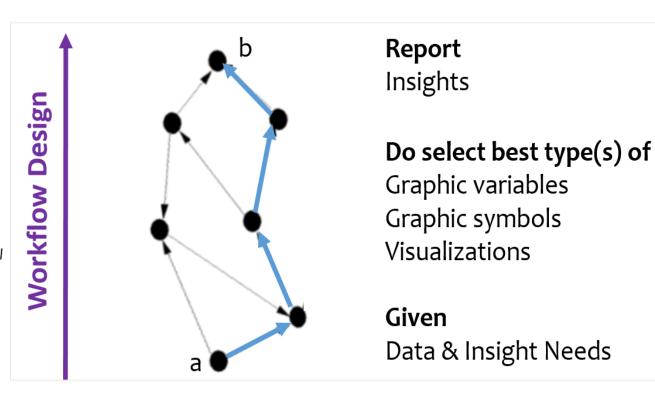
Interactivity
Zoom, filter, details-on-demand

Graphic variables
Position, color, size, shape code

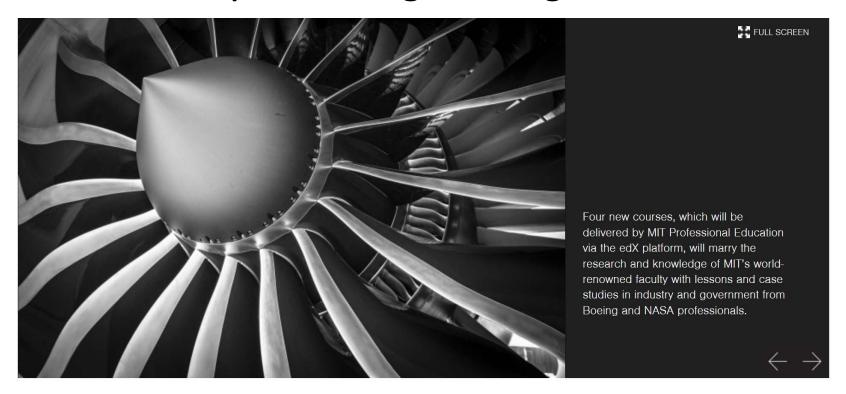
Graphic symbols Geometric, linguistic, or pictorial

Visualizations Table, graph, map, network

Insight Needs & Data



MIT xPRO: Systems Engineering



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

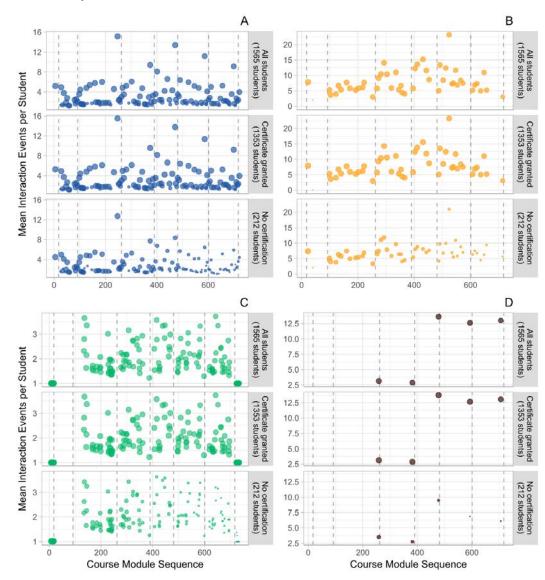
Improving Return on Investment in Education: Measuring, Visualizing, and Optimizing Learner Trajectories

Michael C. Richey, Michael Ginda, Mark Cousino, Katy Börner

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%.



Improving Return on Investment in Education: Measuring, Visualizing, and Optimizing Learner Trajectories

Michael C. Richey, Michael Ginda, Mark Cousino, Katy Börner

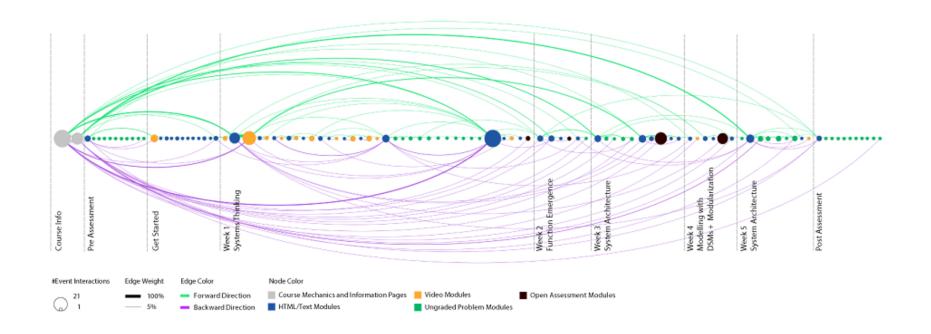


Figure 1: Learner path overlaid on linear sequence of course modules. Linear, temporal sequence of learning modules accessed by a high performing student plotted from left (first) to right (last) with dividing lines for pre and post but also week 1-5 modules.

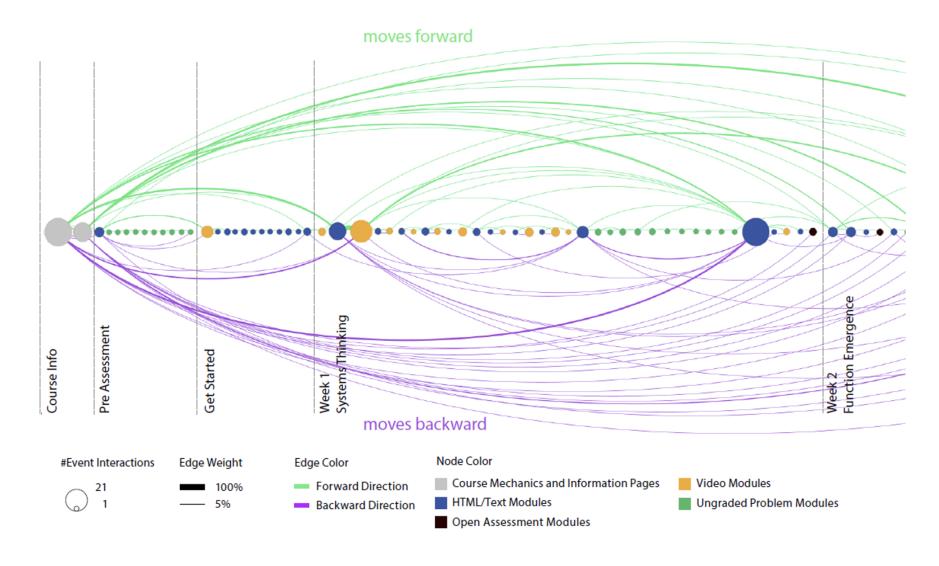
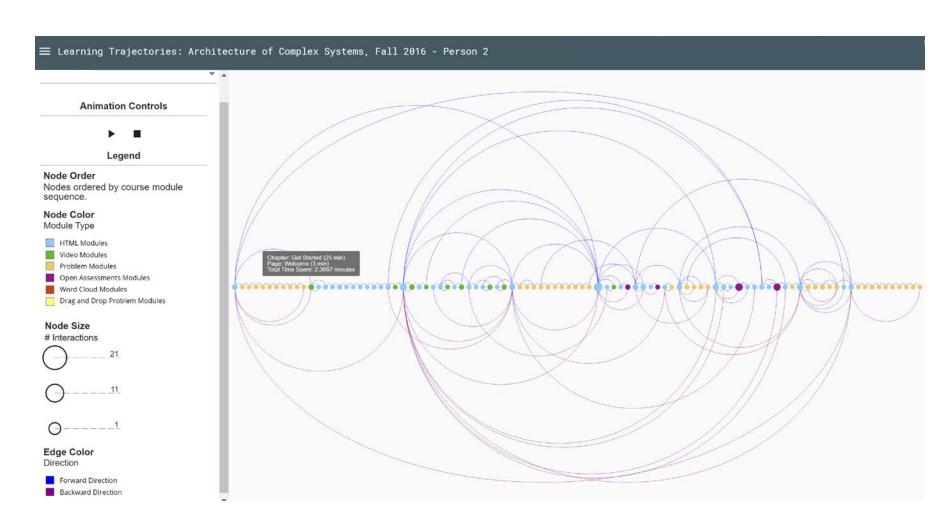


Figure 2. Zoom into learner path overlaid on linear sequence of course modules. Linear, temporal sequence of course modules used by a high performing student; plotted from left (first) to right (last) with dividing lines for different module sections.

Improving Return on Investment in Education: Measuring, Visualizing, and Optimizing Learner Trajectories

Michael C. Richey, Michael Ginda, Mark Cousino, Katy Börner



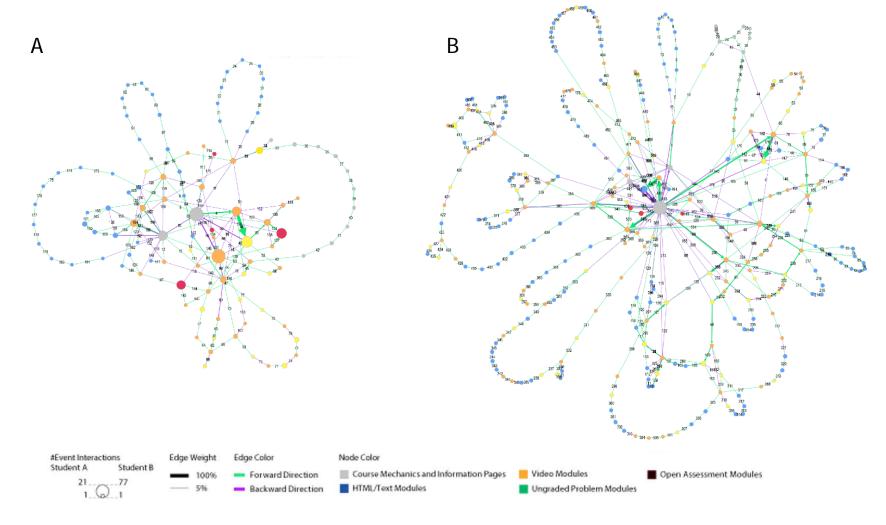


Figure 3: Learner path overlaid on force-directed layout of used course modules. Learner path of a students with high (left) and low (right) performance scores overlaid on force-directed layout of course modules.

MIT xPRO: Additive Manufacturing



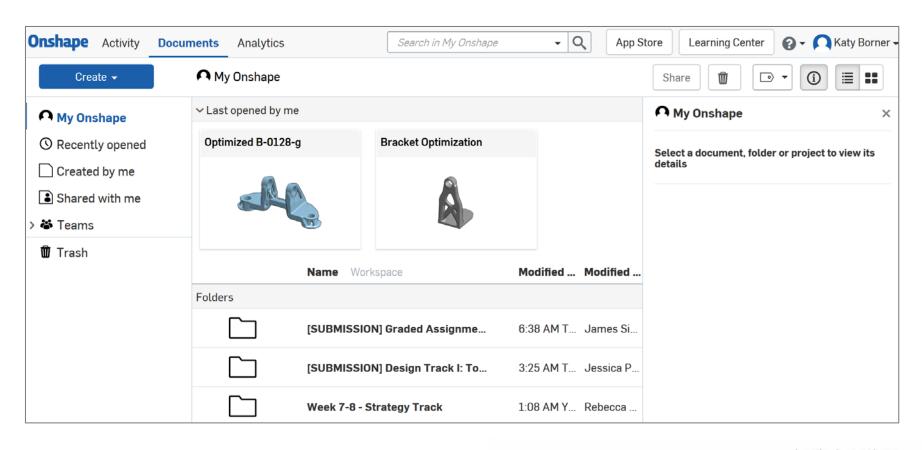
HOME

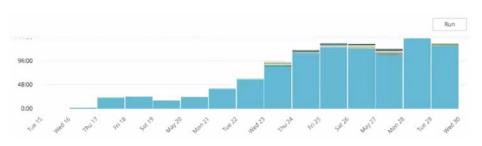
ENTERPRISE



https://additivemanufacturing.mit.edu

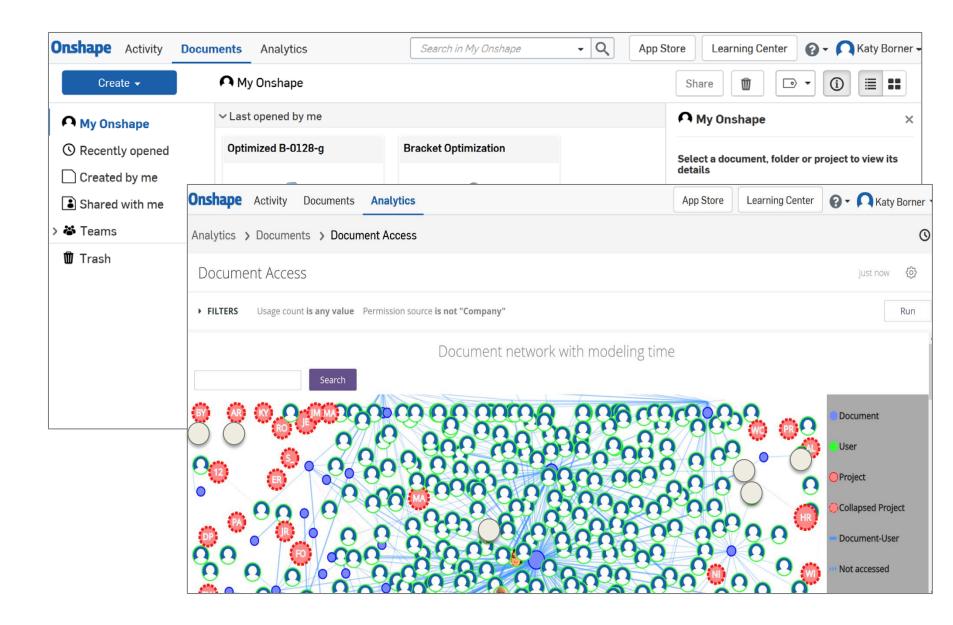
Students use Onshape to practice what they learned







Students use Onshape to practice what they learned



Additive Manufacturing

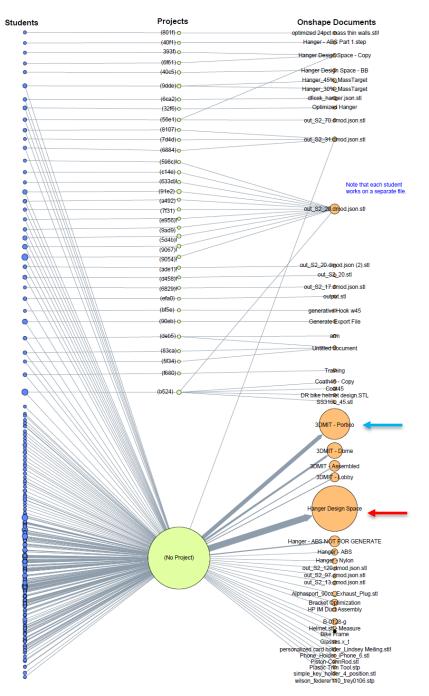
3-modal network of all Students (blue), Teams (green), Documents (orange) used in course by May 31, 2018.

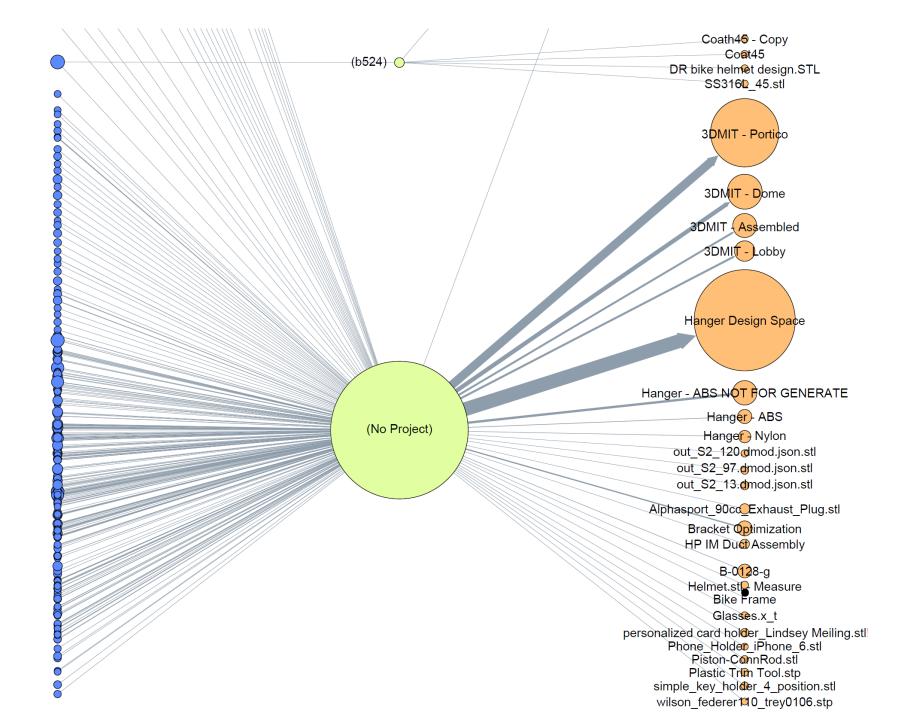
Area size represents the total time associated with a given node in the modeling software.

- 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.









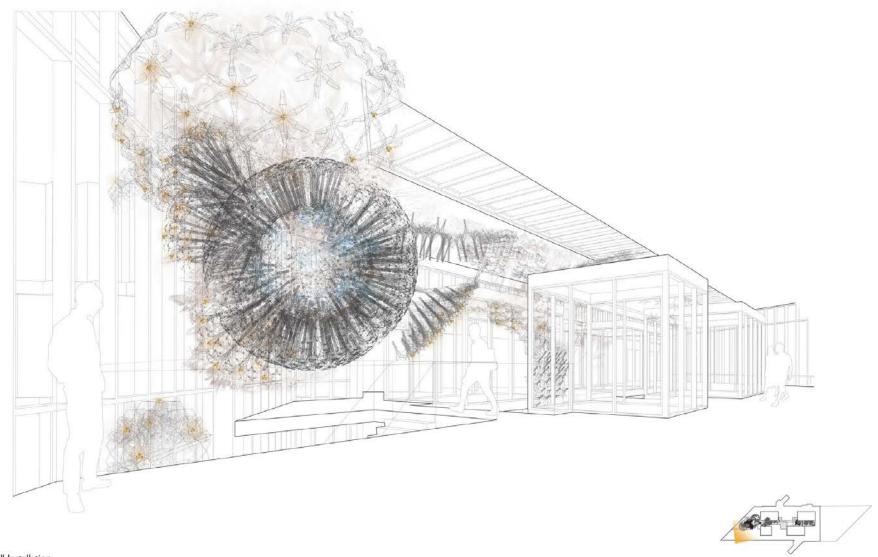
Embracing Human and Machine Intelligence Symbiosis

Visualizing the Internet of Things (IoT)

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



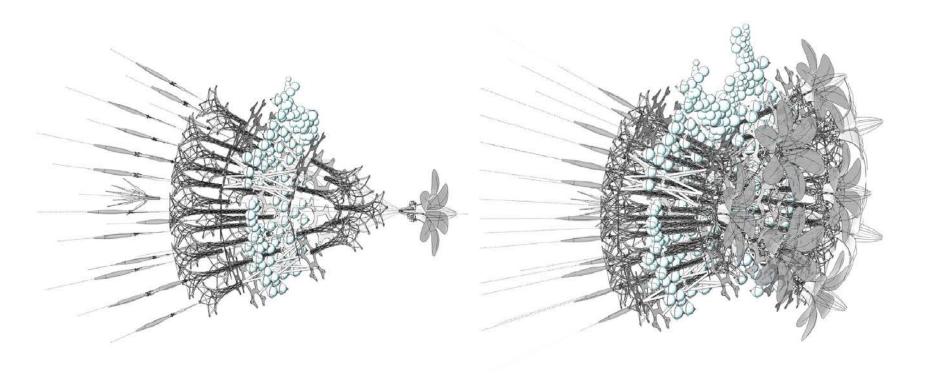




Luddy Hall Installation Indiana University Bloomington April 29 2017

UPPER ATRIUM

Philip Beesley • Living Architecture Systems







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

Börner, Katy, Chen, Chaomei, and Boyack, Kevin. (2003). **Visualizing Knowledge Domains.** In Blaise Cronin (Ed.), *ARIST*, Medford, NJ: Information Today, Volume 37, Chapter 5, pp. 179-255.

http://ivl.slis.indiana.edu/km/pub/2003-borner-arist.pdf

Shiffrin, Richard M. and Börner, Katy (Eds.) (2004). **Mapping Knowledge Domains**. *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl_1).

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

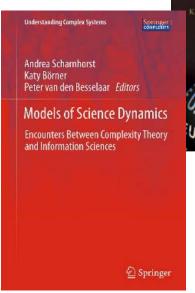
Börner, Katy (2010) Atlas of Science: Visualizing What We Know. The MIT Press.

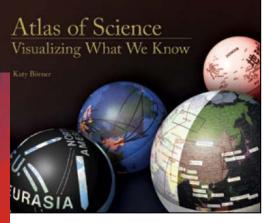
http://scimaps.org/atlas

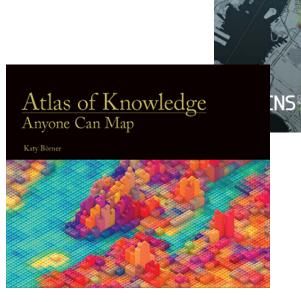
Scharnhorst, Andrea, Börner, Katy, van den Besselaar, Peter (2012) **Models of Science Dynamics**. Springer Verlag.

Katy Börner and David E Polley (2014) **Visual Insights: A Practical Guide to Making Sense of Data**. The MIT Press.

Börner, Katy (2015) **Atlas of Knowledge: Anyone Can Map**. The MIT Press. http://scimaps.org/atlas2









All papers, maps, tools, talks, press are linked from http://cns.iu.edu/presentations.html

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