## Workshop: Open Source Tools for Data Analysis and Visualization

**Katy Börner**
Victor H. Yngve Professor of Information Science
Director, Cyberinfrastructure for Network Science Center
School of Informatics and Computing, Indiana University, USA

*Science & Creativity Annual Conference (KOFAC)*
*Seoul, Korea*

*December 4, 2014*


All papers, maps, tools, talks, press are linked from http://cns.iu.edu
These slides are available at http://cns.iu.edu/docs/presentations

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

## **CNS Macroscopes** are used by hundreds of thousands around the globe

Our mission is to advance datasets, tools, and services for the study of biomedical, social and behavioral science, physics, and other networks. A specific focus is research on the structure and evolution of science and technology (S&T) and the communication of results via static and interactive maps of science. Learn more at cishell.org.

CNS Cyberinfrastructure for Network Science Center
cns.iu.edu

# AcademyScope

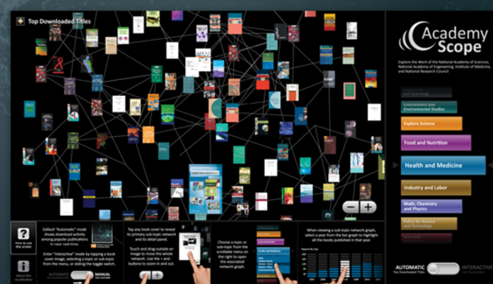
*AcademyScope* is a state-of-the-art, interactive touch-screen visualization developed by CNS in collaboration with the National Academy of Sciences.

Using a 55-inch, multi-touch screen, viewers can explore 20 years of reports published by the National Academy of Sciences, National Academy of Engineering, Institute of Medicine, and National Research Council.

Beginning in October 2014, the *AcademyScope* web application is available to the public through the National Academies Press website. Users can acccess the application through the "Browse by Topic" menu on the NAP homep-age (www.nap.edu), or via the "Browse Topics" button in the header of every interior page. The application can also be accessed directly at www.nap.edu/academy-scope.

Visit cns.iu.edu/interactive_displays to learn more about the design and programming.

CNS Cyberinfrastructure for Network Science Center cns.iu.edu




# Places & Spaces Exhibit

This exhibit aims to demonstrate the power of maps to navigate and make sense of physical places and abstract topic spaces. The tenth and final iteration of maps debuted at the University of Miami on September 4, 2014, where all 100 maps will remain in display through December 11, 2014.

Phase 2 of this unique exhibit is designed to bring Macroscope tools to public places to help exhibit visitors not only learn how to **read** science maps but how to **make** them.

See all the maps and more at the new scimaps.org.

PLACES & SPACES MAPPING SCIENCE

CNS Cyberinfrastructure for Network Science Center cns.iu.edu

# IVMOOC 2015

The Information Visualization MOOC provides an overview about the state of the art in information visualization, teaching the process of producing effective visualizations that take the needs of users into account.

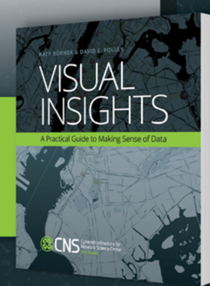
The inaugural IVMOOC, which launched in January 2013, attracted participants from more than 100 countries. It is one of the first MOOCs offered by IU and the first to offer an opportunity for students to work in teams with real clients. All registrants gain free access to the Scholarly Database and the Sci2 Tool.

The course can be taken for three Indiana University credits as part of the Online Data Science Program offered by the School of Informatics and Computing.

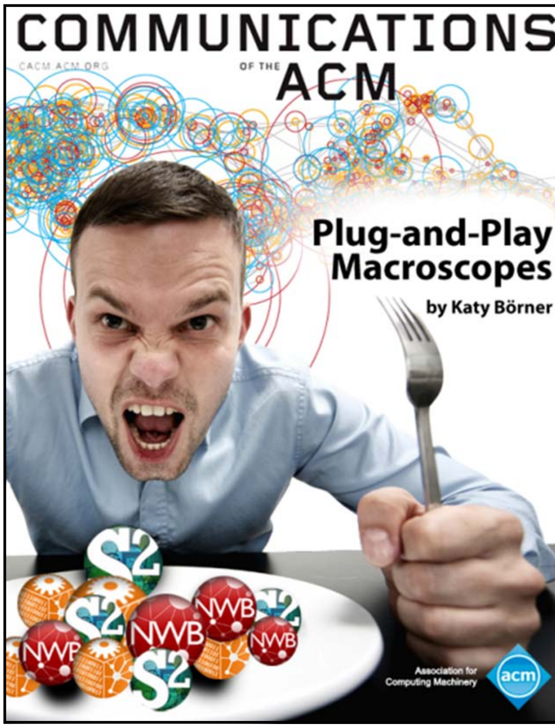
The course will return in January 2015. Learn more at ivmooc.cns.iu.edu.

CNS Cyberinfrastructure for Network Science Center cns.iu.edu


This IVMOOC companion 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.




## Plug-and-Play Macroscopes




Börner, Katy. (2011).
Plug-and-Play Macroscopes.
*Communications of the ACM, 54(3), 60-69.*

Video and paper are at
http://www.scivee.tv/node/27704

7

## Designing "Dream Tools"

Many of the best micro-, tele-, and macroscopes are designed by **scientists keen to observe and comprehend what no one has seen or understood before.** Galileo Galilei (1564–1642) recognized the potential of a spyglass for the study of the heavens, ground and polished his own lenses, and used the improved optical instruments to make discoveries like the moons of Jupiter, providing quantitative evidence for the Copernican theory.

Today, scientists **repurpose, extend, and invent new hardware and software** to create **"macroscopes"** that may solve both local and global challenges.

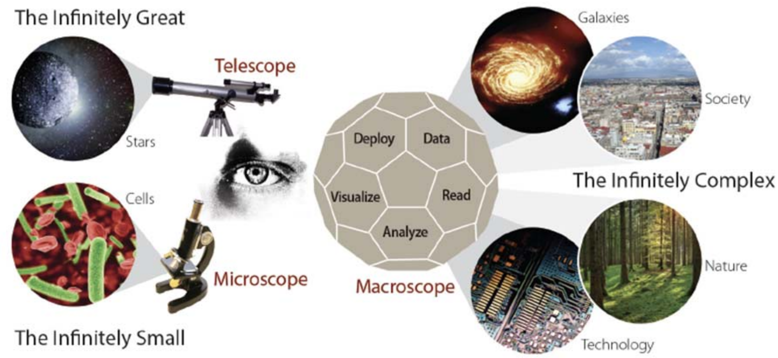
CNS Macroscope tools **empower** me, my students, colleagues, and more than 100,000 others that downloaded them.

8

# Macroscopes

Decision making in science, industry, and politics, as well as in daily life, requires that we make sense of data sets representing the structure and dynamics of complex systems. Analysis, navigation, and management of these continuously evolving data sets require a new kind of data-analysis and visualization tool we call a macroscope (from the Greek macros, or "great," and skopein, or "to observe") inspired by de Rosnay's futurist science writings.

Macroscopes provide a "vision of the whole," helping us "synthesize" the related elements and enabling us to detect patterns, trends, and outliers while granting access to myriad details. Rather than make things larger or smaller, **macroscopes let us observe what is at once too great, slow, or complex for the human eye and mind to notice and comprehend.**




9

# Plug-and-Play Macroscopes

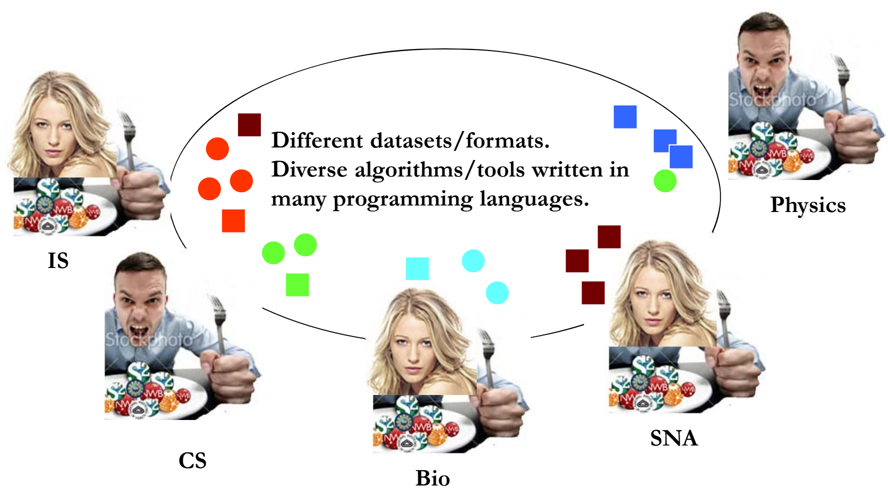
**Inspire computer scientists** to implement software frameworks that **empower domain scientists** to assemble their own continuously evolving macroscopes, adding and upgrading existing (and removing obsolete) plug-ins to arrive at a set that is truly relevant for their work—with little or no help from computer scientists.

While microscopes and telescopes are physical instruments, **macroscopes resemble continuously changing bundles of software plug-ins.** Macroscopes make it easy to select and combine algorithm and tool plug-ins but also interface plug-ins, workflow support, logging, scheduling, and other plug-ins needed for scientifically rigorous yet effective work.

They make it easy to share plug-ins via email, flash drives, or online. To use new plugins, simply copy the files into the plug-in directory, and they appear in the tool menu ready for use. No restart of the tool is necessary. **Sharing algorithm components, tools, or novel interfaces becomes as easy as sharing images on Flickr or videos on YouTube. Assembling custom tools is as quick as compiling your custom music collection.**

10

## Changing Scientific Landscape— Personal Observations




11

## Related Work

Google Code and SourceForge.net provide special means for developing and distributing software

- In August 2009, SourceForge.net hosted more than 230,000 software projects by two million registered users (285,957 in January 2011);
- In August 2009 ProgrammableWeb.com hosted 1,366 application programming interfaces (APIs) and 4,092 mashups (2,699 APIs and 5,493 mashups in January 2011)

Cyberinfrastructures serving large biomedical communities

- Cancer Biomedical Informatics Grid (caBIG) (http://cabig.nci.nih.gov)
- Biomedical Informatics Research Network (BIRN) (http://nbirn.net)
- Informatics for Integrating Biology and the Bedside (i2b2) (https://www.i2b2.org)
- HUBzero (http://hubzero.org) platform for scientific collaboration uses
- myExperiment (http://myexperiment.org) supports the sharing of scientific workflows and other research objects.

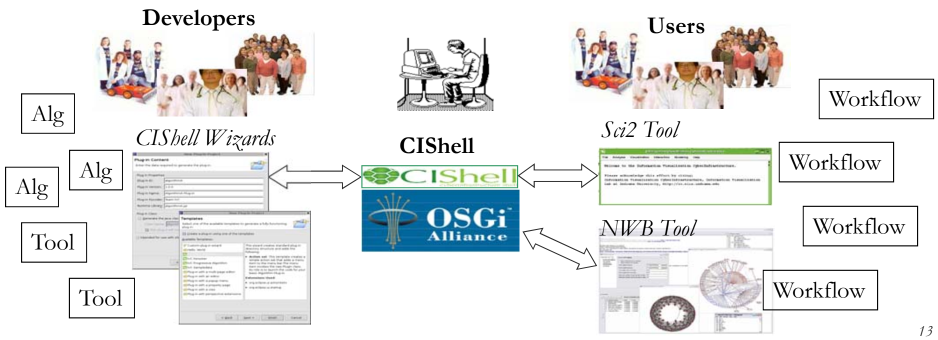
Missing so far is a **common standard** for

- the design of **modular, compatible algorithm and tool plug-ins** (also called "modules" or "components")
- that can be **easily combined into scientific workflows** ("pipeline" or "composition"),
- and packaged as **custom tools.**

12

## OSGi & CIShell

- CIShell (http://cishell.org) is an open source software specification for the integration and utilization of datasets, algorithms, and tools.
- It extends the Open Services Gateway Initiative (OSGi) (http://osgi.org), a standardized, component oriented, computing environment for networked services widely used in industry since more than 10 years.
- Specifically, CIShell provides "sockets" into which existing and new datasets, algorithms, and tools can be plugged using a wizard-driven process.






## CIShell Portal and Developer Guide
(http://cishell.org)

**Home**

Edit   Add

1 Added by Micah Linnemeier, last edited by Micah Linnemeier on Mar 16, 2011 (view change)

### About the Cyberinfrastructure Shell

The Cyberinfrastructure Shell (CIShell) is an open source, community-driven platform for the integration and utilization of datasets, algorithms, tools, and computing resources. Algorithm integration support is built in for Java and most other programming languages. Being Java based, it will run on almost all platforms. The software and specification is released under an Apache 2.0 License.

CIShell is the basis of Network Workbench, TexTrend, Sci² and the upcoming EpiC tool.

CIShell supports remote execution of algorithms. A standard web service definition is in development that will allow pools of algorithms to transparently be used in a peer-to-peer, client-server, or web front-end fashion.

### CIShell Features

#### A framework for easy integration of new and existing algorithms written in any programming language

Using CIShell, an algorithm writer can fully concentrate on creating their own algorithm in whatever language they are comfortable with. Simple tools are provided to then take their algorithm and

### Learn More...

- CIShell Papers
- CIShell Powered Tools
- Algorithms
- Plugins (coming soon)
- Misc. Tool Documentation
- CIShell Web Services (coming soon)
- Screenshots

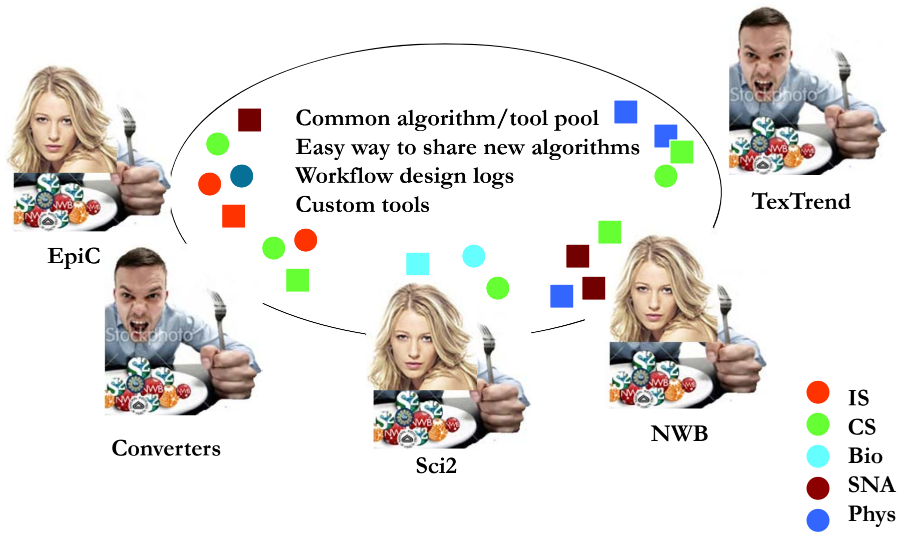
### Getting Started...

- Documentation & Developer Resources
- Download

### Getting Involved...

- Contact Us

# Changing Scientific Landscape— Personal Observations Cont.




15

# OSGi/CIShell Adoption

CIShell/OSGi is at the core of different CIs and a total of 169 unique plugins are used in the

- ***Information Visualization*** (http://iv.slis.indiana.edu),
- ***Network Science (NWB Tool)*** (http://nwb.slis.indiana.edu),
- ***Scientometrics and Science Policy ($Sci^2$ Tool)*** (http://sci.slis.indiana.edu), and
- ***Epidemics*** (http://epic.slis.indiana.edu) research communities.

Most interestingly, a number of other projects recently adopted OSGi and one adopted CIShell:

***Cytoscape*** (http://www.cytoscape.org) lead by Trey Ideker, UCSD is an open source bioinformatics software platform for visualizing molecular interaction networks and integrating these interactions with gene expression profiles and other state data (Shannon et al., 2002). **Bruce visits Mike Smoot in 2009**

***Taverna Workbench*** (http://taverna.sourceforge.net) lead by Carol Goble, University of Manchester, UK is a free software tool for designing and executing workflows (Hull et al., 2006). Taverna allows users to integrate many different software tools, including over 30,000 web services. **Micah, June 2010**

***MAEviz*** (https://wiki.ncsa.uiuc.edu/display/MAE/Home) managed by Shawn Hampton, NCSA is an open-source, extensible software platform which supports seismic risk assessment based on the Mid-America Earthquake (MAE) Center research.

***TEXTrend*** (http://www.textrend.org) lead by George Kampis, Eötvös University, Hungary develops a framework for the easy and flexible integration, configuration, and extension of plugin-based components in support of natural language processing (NLP), classification/mining, and graph algorithms for the analysis of business and governmental text corpuses with an inherently temporal component.

As the functionality of OSGi-based software frameworks improves and the number and diversity of dataset and algorithm plugins increases, the capabilities of custom tools will expand.

16

# Macroscopes for Kids: Learning Objectives

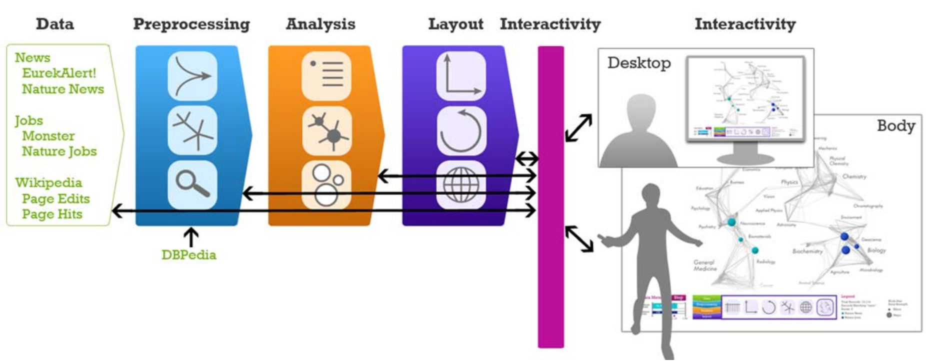
**Read Charts/Maps**
- Distinguish different data sources, types, amounts
- Understand different reference systems
- Understand distortion/projection
- Read visual languages

**Make Charts/Maps** – a la http://makezine.com but hands-on
- Identify type and level of analysis
- Design and execute workflows
- Invent effective visual languages
- Deploy (print, bring online)

Employ **READ<->MAKE feedback loop** and have fun with both!

17

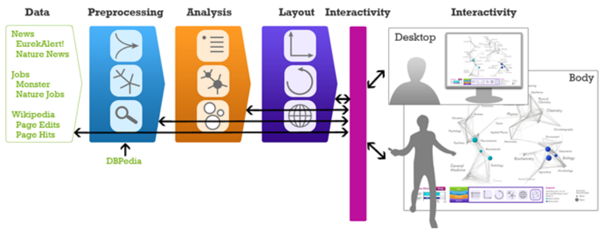
# Macroscopes for Kids: Setup




**Macroscope can empower children and others to**
- Select one or many live data streams and/or static datasets, e.g., Wikipedia and DBPedia,
- Preprocess data, e.g., to delete, filter, merge, extract networks, search data.
- Analyze data, e.g., to compute node and edge properties, identify bursts and clusters.
- Layout data using different reference systems, e.g., tables, charts, circular layout, network, geospatial maps, or science maps.

18

# Macroscopes for Kids: Setup cont.




**Macroscope can empower children and others to**

- Select a data record (e.g., rendered as dot in a scatter plot, dot on a map, or a node in a network) and to search for all other nodes with similar attributes and to request a change of their color, size, shape, or other visual attributes.
- Manipulate the layout, e.g., to sort tables, to change chart axes from linear to log scaling, change cartographic projections.
- To zoom and pan, filter, and request details, e.g., to open a Web browser with the page loaded.

At any moment in time, the complete workflow (currently active datasets, applied preprosessing, analyses, layouts, automatically compiled legend) is visible to facilitate learning and memorization of dataset, algorithm, and parameter choices and their impact on the design of meaningful visualizations.

19

## Visualization Framework

20

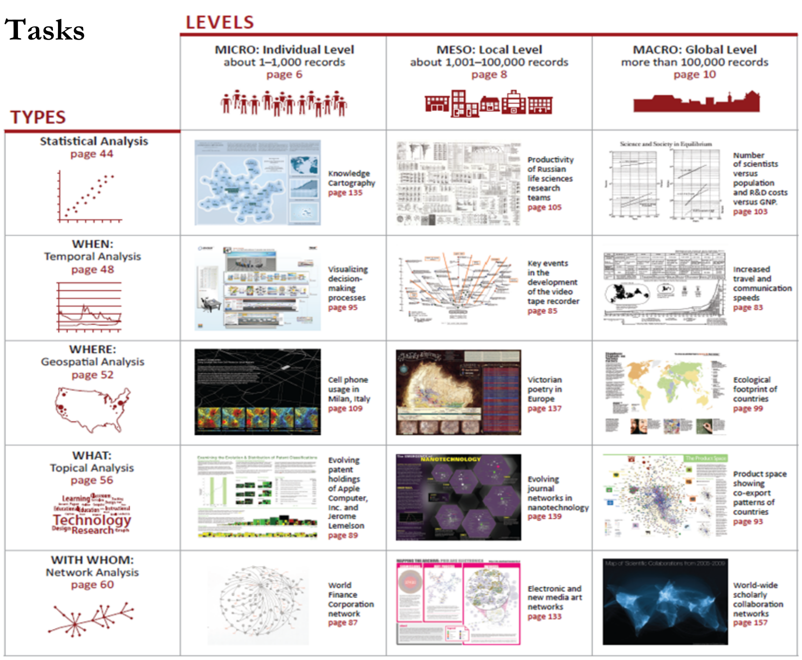
**Theoretically Grounded and Practically Useful Visualization Framework**
developed to empower the broadest spectrum of users to read and make data visualizations that are useful and meaningful to them.
The visualization framework was used to

- design the aforementioned study and
- develop plug-and-play macroscope tools that improve the data visualization literacy of researchers, practitioners, IVMOOC students, museum visitors, and others.

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

21

Tasks




22

## Workflow Design

**Interpret**
Validation and Interpretation, **page 72**

**Acquire**
User Needs Acquisition, **page 40**
Data Acquisition, **page 42**

**Deploy**
Human-Computer Interface, **page 70**
Interaction, **page 68**
Combination, **page 66**

**Analyze & Visualize**
Statistical Studies, **page 44**
Statistical Visualization Types, **page 46**
Temporal Studies—"When", **page 48**
Temporal Visualization Types, **page 50**
Geospatial Studies—"Where", **page 52**
Geospatial Visualization Types, **page 54**
Topical Studies—"What", **page 56**
Topical Visualization Types, **page 58**
Network Studies—"With Whom", **page 60**
Network Visualization Types, **page 62**
Studying Dynamics, **page 64**

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



## Types

**Insight Need Types** (page 26)
- categorize/cluster
- order/rank/sort
- distributions (also outliers, gaps)
- comparisons
- trends (process and time)
- geospatial
- compositions (also of text)
- correlations/relationships

**Data Scale Types** (page 28)
- nominal
- ordinal
- interval
- ratio

**Visualization Types** (page 30)
- table
- chart
- graph
- map
- network layout

**Graphic Symbol Types** (page 32)
- geometric symbols
  - point
  - line
  - area
  - surface
  - volume
- linguistic symbols
  - text
  - numerals
  - punctuation marks
- pictorial symbols
  - images
  - icons
  - statistical glyphs

**Graphic Variable Types** (page 34)
- spatial
  - position
- retinal
  - form
  - color
  - optics
  - motion

**Interaction Types** (page 26)
- overview
- zoom
- search and locate
- filter
- details-on-demand
- history
- extract
- link and brush
- projection
- distortion

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

# Types

| 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 |

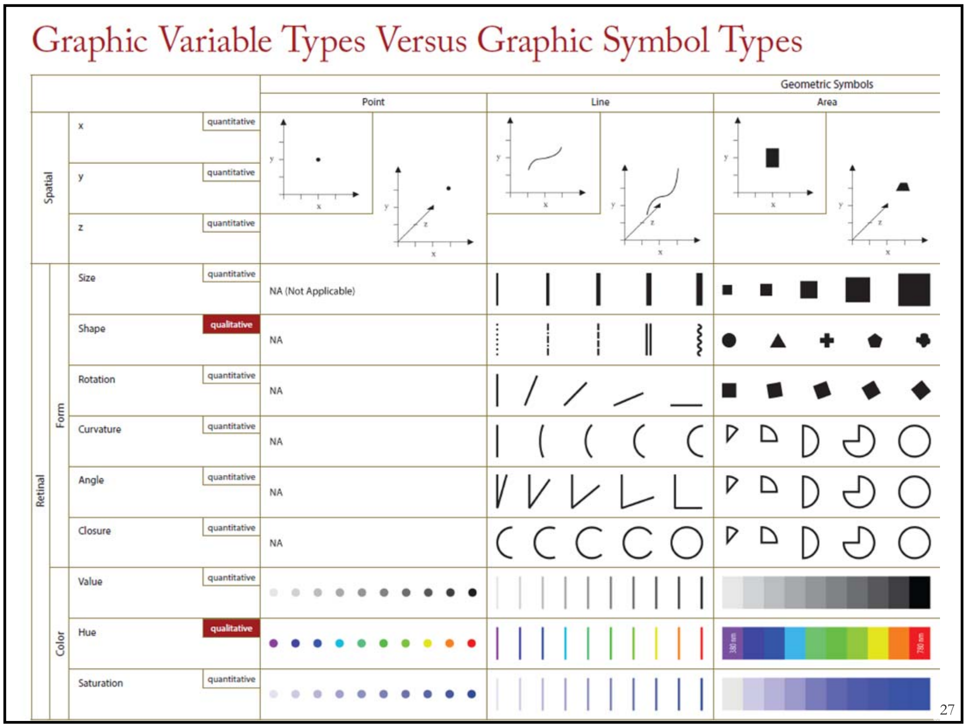
25

# Types

| 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<br>• order/rank/sort<br>• distributions (also outliers, gaps)<br>• comparisons<br>• trends (process and time)<br>• geospatial<br>• compositions (also of text)<br>• correlations/relationships | • nominal<br>• ordinal<br>• interval<br>• ratio | • table<br>• chart<br>• graph<br>• map<br>• network layout | • geometric symbols<br>point<br>line<br>area<br>surface<br>volume<br>• linguistic symbols<br>text<br>numerals<br>punctuation marks<br>• pictorial symbols<br>images<br>icons<br>statistical glyphs | • spatial<br>position<br>• retinal<br>form<br>color<br>optics<br>motion | • overview<br>• zoom<br>• search and locate<br>• filter<br>• details-on-demand<br>• history<br>• extract<br>• link and brush<br>• projection<br>• distortion |

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

26




## The Sci2 Tool:

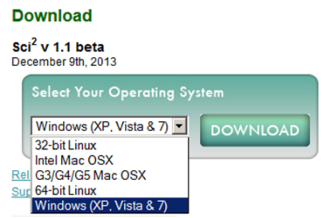
**A Plug-and-Play Macroscope that implements the Visualization Framework**

28

CNS Cyberinfrastructure for Network Science Center

## Software, Datasets, Plugins, and Documentation

- These slides
  http://cns.iu.edu/docs/presentations/2014-borner-opentoolstutorial-oecd.pdf
- Sci2 Tool Manual v0.5.1 Alpha, updated to match v1.0 Alpha tool release
  http://sci2.wiki.cns.iu.edu
- Sci2 Tool v 1.1 beta (Dec 9, 2013)
  http://sci2.cns.iu.edu
- Additional Datasets
  http://sci2.wiki.cns.iu.edu/2.5+Sample+Datasets
- Additional Plugins
  http://sci2.wiki.cns.iu.edu/3.2+Additional+Plugins




Make sure you have Java 1.6 (32-bit suffices) or higher installed or download from http://www.java.com/en/download. To check your Java version, open a terminal and run 'java -version'.

Some visualizations are saved as Postscript files. A free Postscript to PDF viewer is at http://ps2pdf.com and a free PDF Viewer at http://www.adobe.com/products/reader.html.

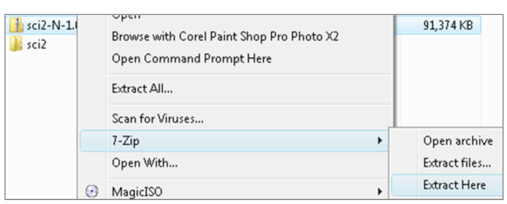
29

CNS Cyberinfrastructure for Network Science Center

## Install and Run Sci2


Sci2 Tool runs on Windows, Mac, and Linux.

Unzip.




Run /sci2/sci2.exe

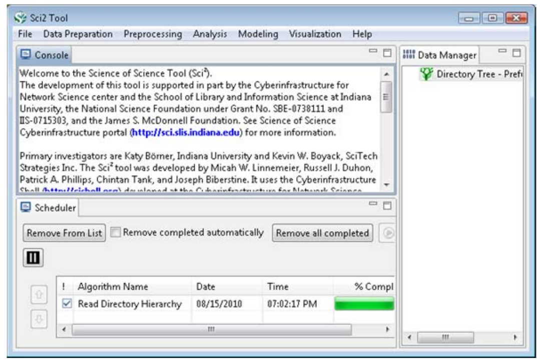
30

## Sci2 Tool Interface Components

See also http://sci2.wiki.cns.iu.edu/2.2+User+Interface

Use

- **Menu** to read data, run algorithms.
- **Console** to see work log, references to seminal works.
- **Data Manager** to select, view, save loaded, simulated, or derived datasets.
- **Scheduler** to see status of algorithm execution.




All workflows are recorded into a log file (see /sci2/logs/…), and can be re-run for easy replication. If errors occur, they are saved in a error log to ease bug reporting.

All algorithms are documented online; workflows are given in Sci2 Manual at http://sci2.wiki.cns.iu.edu

31

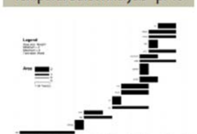
## Load **One** File and Run **Many** Analyses and Visualizations

| Times Cited | Publication Year | City of Publisher | Country | Journal Title (Full) | Title | Subject Category | Authors |
|---|---|---|---|---|---|---|---|
| 12 | 2011 | NEW YORK | USA | COMMUNICATIONS OF THE ACM | Plug-and-Play Macroscopes | Computer Science | Borner, K |
| 18 | 2010 | MALDEN | USA | CTS-CLINICAL AND TRANSLATIONAL 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 | SCIENCE TRANSLATIONAL MEDICINE | 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

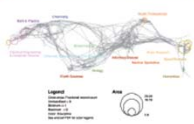
32

CNS Cyberinfrastructure for Network Science Center

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

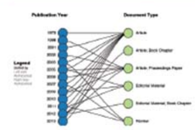
| Times Cited | Publication Year | City of Publisher | Country | Journal Title (Full) | Title | Subject Category | Authors |
|---|---|---|---|---|---|---|---|
| 12 | 2011 | NEW YORK | USA | COMMUNICATIONS OF THE ACM | Plug-and-Play Macroscopes | Computer Science | Borner, K |
| 18 | 2010 | MALDEN | USA | CTS-CLINICAL AND TRANSLATIONAL 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 | SCIENCE TRANSLATIONAL MEDICINE | 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 |

Topical Analysis–p. 56



Paper Citation Network–p. 60

Bi-Modal Network–p. 60



Co-author and many other bi-modal networks.

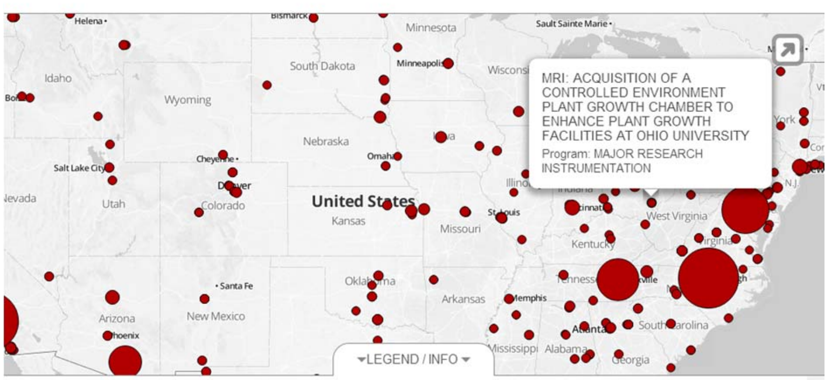
33

CNS Cyberinfrastructure for Network Science Center

NETE® AV ANALYZER | VISUALIZER

Proportional Symbol Map

Relationship between Projects and External Organizations - Larry E. Humes, Bernice A. Pescosolido; Generated by NETE March 5, 2014 | 9:34 AM EST

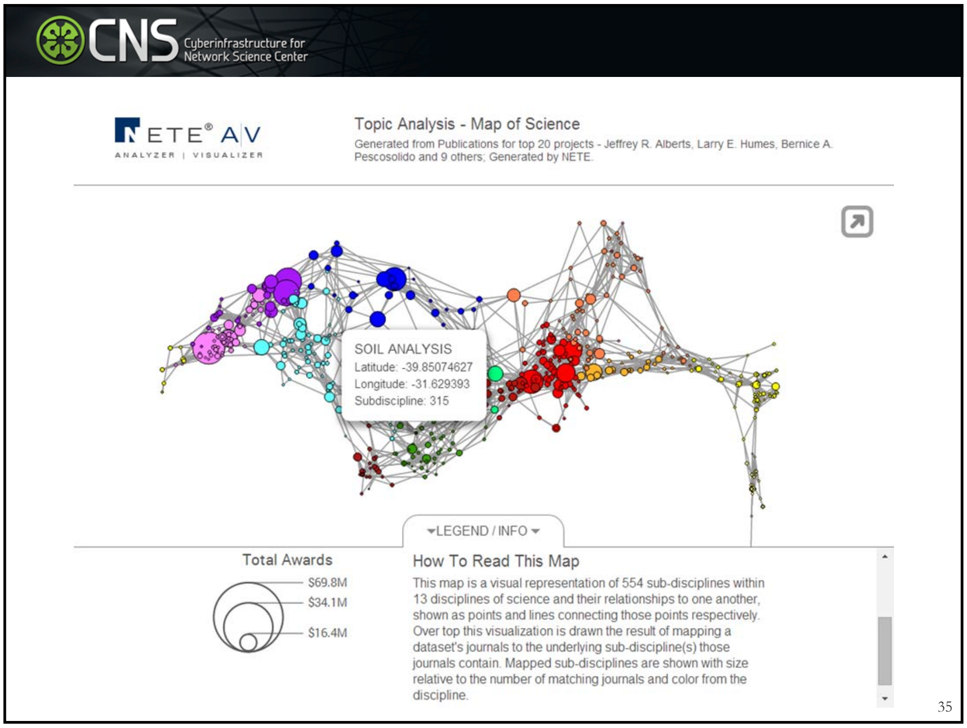



Amount Awarded


How To Read This Map

This proportional symbol map shows 52 U.S. states and other jurisdictions using the Albers equal-area conic projection with Alaska, Puerto Rico, and Hawaii inset. Each dataset record is represented by a circle centered at its geolocation. The area, interior color, and exterior color of each circle may represent numeric attribute values. Minimum and maximum data values are given in the legend.

34

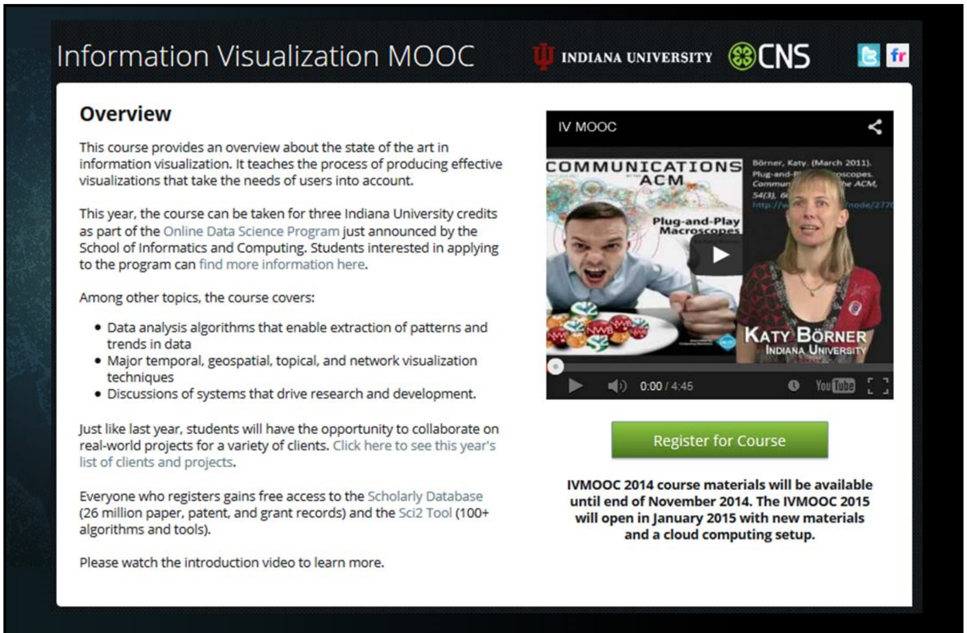



**Information Visualization MOOC**

**Teaches the**

**Visualization Framework and The Sci2 Tool**

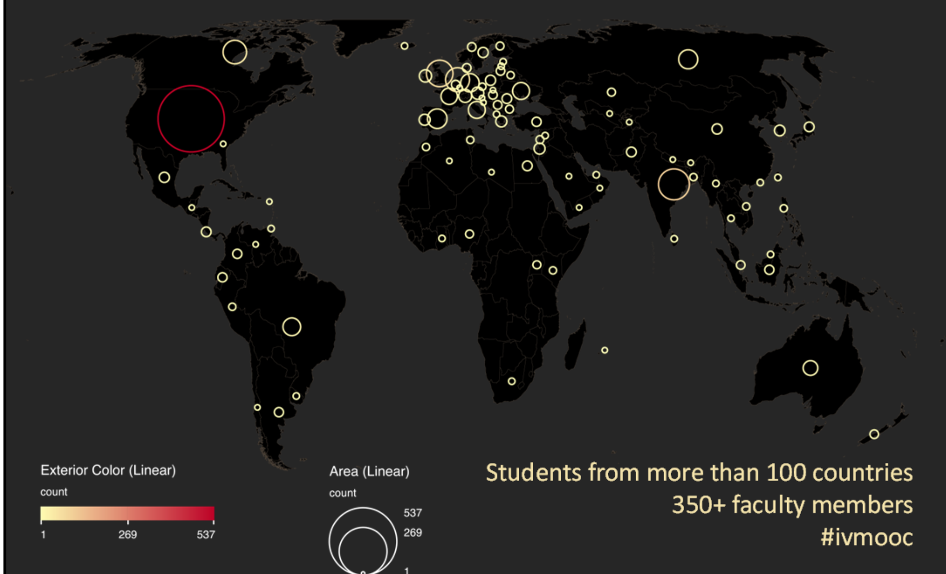
36




Register for free at http://ivmooc.cns.iu.edu. Class will restart in January 2015.

37

The Information Visualization MOOC
ivmooc.cns.iu.edu




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

38

## Course Schedule

- **Session 1** – Workflow design and visualization framework
- **Session 2** – “When:” Temporal Data
- **Session 3** – “Where:” Geospatial Data
- **Session 4** – “What:” Topical Data

**Mid-Term**


**Students work in teams with clients.**

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

**Final Exam**

Final grade is based on Midterm (**30%**), Final (**40%**), Client Project (**30%**).



## Theory Unit Structure

Each theory unit comprises:

- Examples of best visualizations
- Visualization goals
- Key terminology
- General visualization types and their names

---

- Workflow design
  - Read data
  - Analyze
  - Visualize

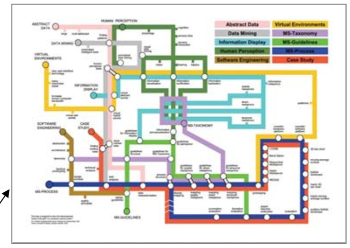
---

- Discussion of specific algorithms

## Different Question Types

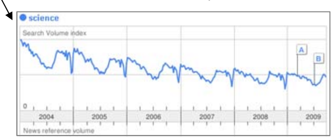
Terabytes of data

Descriptive & Predictive Models



Find your way

Find collaborators, friends

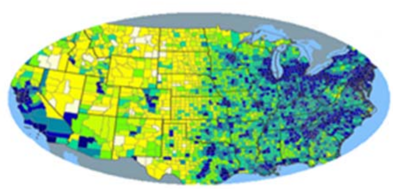



Identify trends

41

## Different Levels of Abstraction/Analysis

Macro/Global
Population Level



Meso/Local
Group Level

Micro
Individual Level

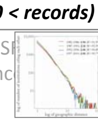
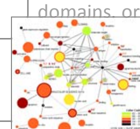
# Type of Analysis vs. Level of Analysis

| | *Micro/Individual (1-100 records)* | *Meso/Local (101–10,000 records)* | *Macro/Global (10,000 < records)* |
|---|---|---|---|
| ***Statistical Analysis/Profiling*** | Individual person and their expertise profiles | Larger labs, centers, universities, research domains, or states | All of NSF, all of USA, all of science. |
| ***Temporal Analysis (When)*** | Funding portfolio of one individual | Mapping topic bursts in 20 years of *PNAS* | 113 years of physics research |
| ***Geospatial Analysis (Where)*** | Career trajectory of one individual | Mapping a state's intellectual landscape | *PNAS* publications |
| ***Topical Analysis (What)*** | Base knowledge from which one grant draws. | Knowledge flows in chemistry research | VxOrd/Topic maps of NIH funding |
| ***Network Analysis (With Whom?)*** | NSF Co-PI network of one individual | Co-author network | NIH's core competency |

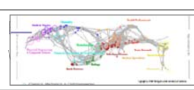
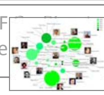
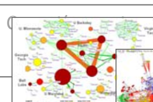
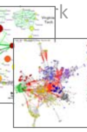
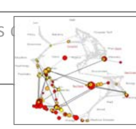
43

# Type of Analysis vs. Level of Analysis

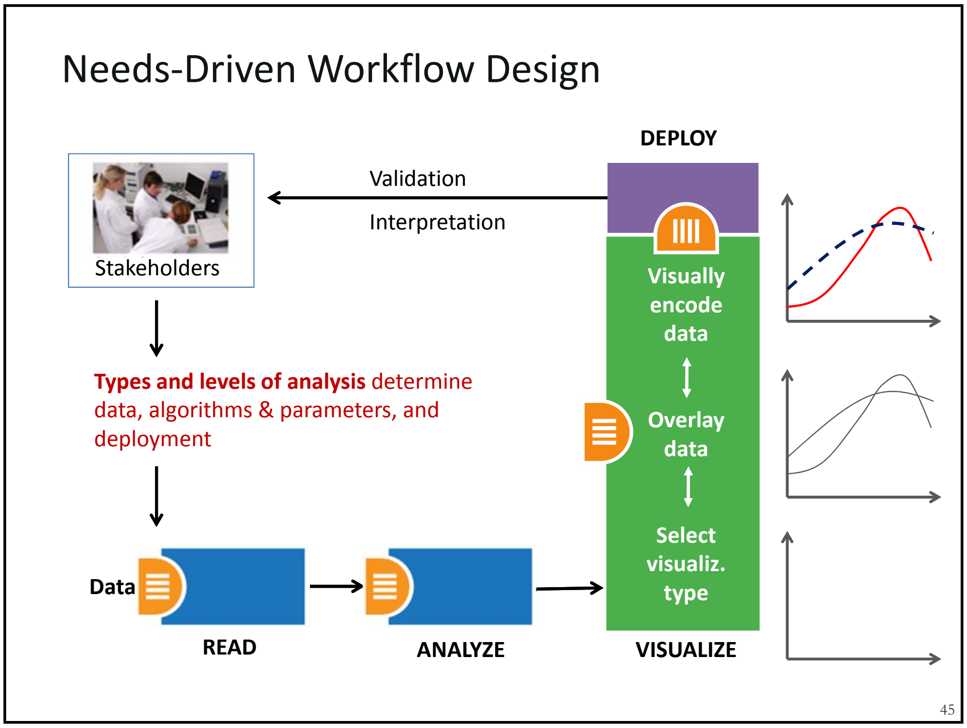
| | *Micro/Individual (1-100 records)* | *Meso/Local (101–10,000 records)* | *Macro/Global (10,000 < records)* |
|---|---|---|---|
| ***Statistical Analysis/Profiling*** | Individual person and their expertise profiles | Larger labs, centers, universities, research domains, or states | All of NSF … , all of scienc… |
| ***Temporal Analysis (When)*** | Funding portfolio of one individual | …ic bursts …f *PNAS* | 113 years of p… research |
| ***Geospatial Analysis (Where)*** | Career trajectory of one individual | … intellectual l… | *PNAS* p… |
| ***Topical Analysis (What)*** | … s. | Knowledge f… chemistry research | |
| ***Network Analysis (With Whom?)*** | NSF … ork of one … | C… k | NIH's … cy |

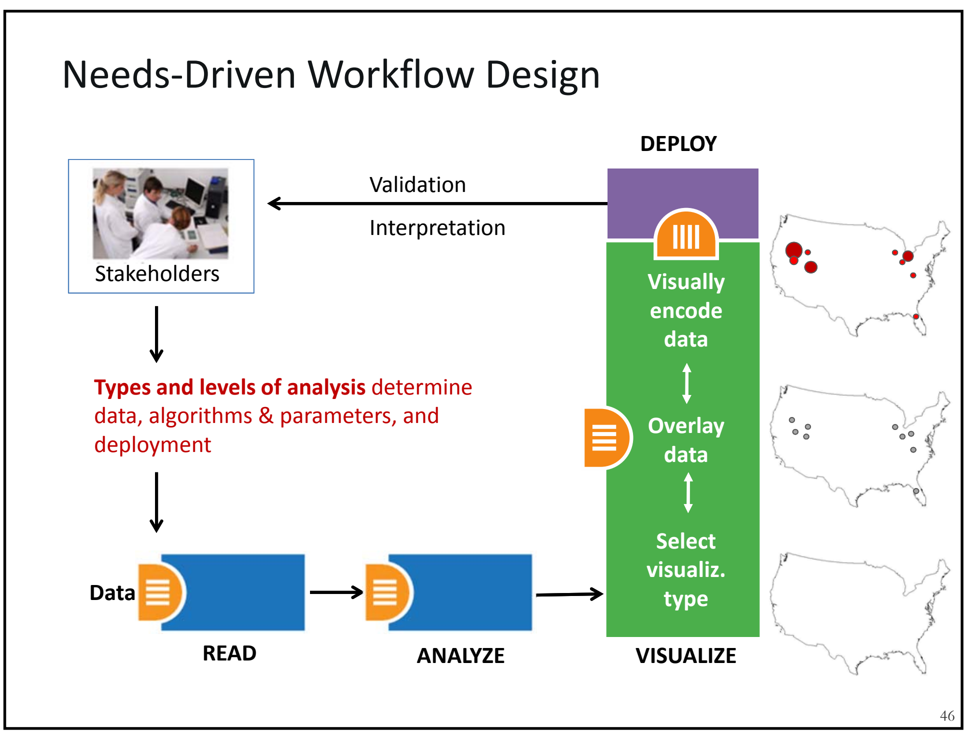





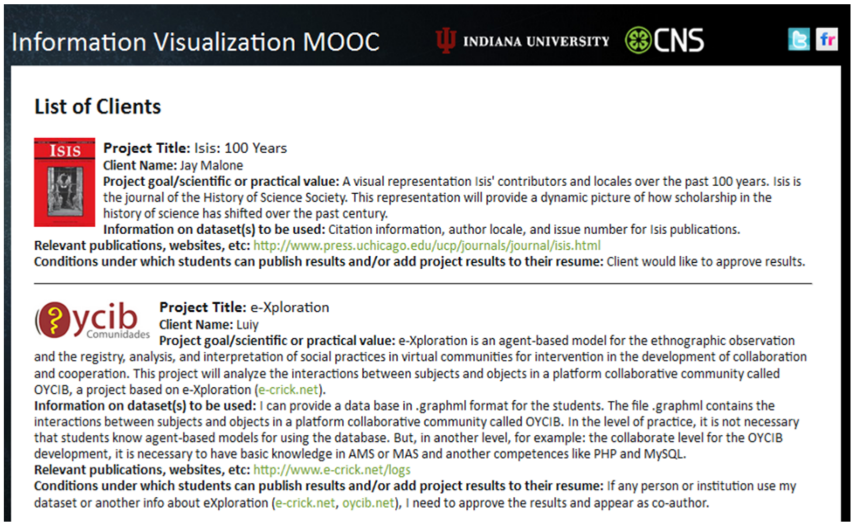







44


Needs-Driven Workflow Design
DEPLOY
Validation
Interpretation
Stakeholders
Types and levels of analysis determine data, algorithms & parameters, and deployment
Visually encode data
Overlay data
Select visualiz. type
Data
READ
ANALYZE
VISUALIZE
45


Needs-Driven Workflow Design
DEPLOY
Validation
Interpretation
Stakeholders
Types and levels of analysis determine data, algorithms & parameters, and deployment
Visually encode data
Overlay data
Select visualiz. type
Data
READ
ANALYZE
VISUALIZE
46

# Clients


Information Visualization MOOC INDIANA UNIVERSITY CNS

**List of Clients**

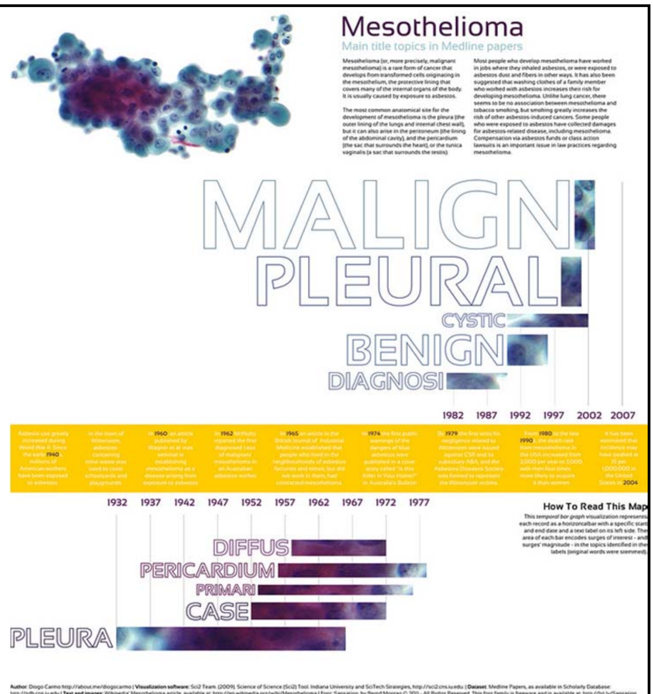
**Project Title:** Isis: 100 Years
**Client Name:** Jay Malone
**Project goal/scientific or practical value:** A visual representation Isis' contributors and locales over the past 100 years. Isis is the journal of the History of Science Society. This representation will provide a dynamic picture of how scholarship in the history of science has shifted over the past century.
**Information on dataset(s) to be used:** Citation information, author locale, and issue number for Isis publications.
**Relevant publications, websites, etc:** http://www.press.uchicago.edu/ucp/journals/journal/isis.html
**Conditions under which students can publish results and/or add project results to their resume:** Client would like to approve results.

**Project Title:** e-Xploration
**Client Name:** Luiy
**Project goal/scientific or practical value:** e-Xploration is an agent-based model for the ethnographic observation and the registry, analysis, and interpretation of social practices in virtual communities for intervention in the development of collaboration and cooperation. This project will analyze the interactions between subjects and objects in a platform collaborative community called OYCIB, a project based on e-Xploration (e-crick.net).
**Information on dataset(s) to be used:** I can provide a data base in .graphml format for the students. The file .graphml contains the interactions between subjects and objects in a platform collaborative community called OYCIB. In the level of practice, it is not necessary that students know agent-based models for using the database. But, in another level, for example: the collaborate level for the OYCIB development, it is necessary to have basic knowledge in AMS or MAS and another competences like PHP and MySQL.
**Relevant publications, websites, etc:** http://www.e-crick.net/logs
**Conditions under which students can publish results and/or add project results to their resume:** If any person or institution use my dataset or another info about eXploration (e-crick.net, oycib.net), I need to approve the results and appear as co-author.

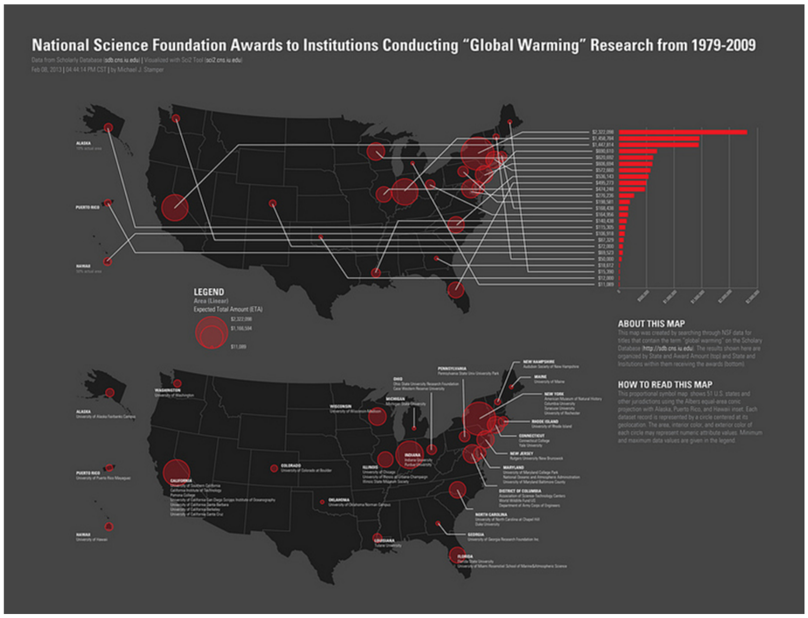
http://ivmooc.cns.iu.edu/clients.html

47

Diogo Carmo




48




mjstamper_ivmooc



# 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/

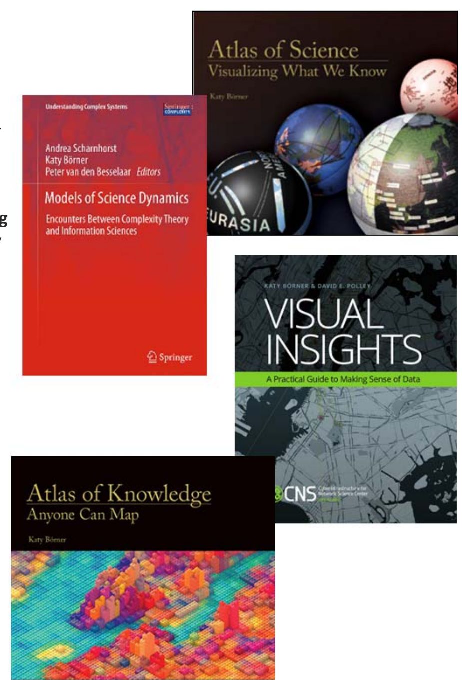
Börner, Katy (2010) **Atlas of Science: Visualizing What We Know**. The MIT Press. http://scimaps.org/atlas

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

Katy Börner, Michael Conlon, Jon Corson-Rikert, Cornell, Ying Ding (2012) **VIVO: A Semantic Approach to Scholarly Networking and Discovery**. Morgan & Claypool.

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
These slides will soon be at http://cns.iu.edu/docs/presentations

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

51