MOOC Data Analysis:
Open-Source Code for Open Education

Ted Polley
Research & Editorial Assistant
Cyberinfrastructure for Network Science Center
School of Informatics and Computing
Indiana University Bloomington

Abe Murray
Senior Product Manager
Google Research

Information Visualization (IV) MOOC

• Sept 2012: Indiana University explores MOOCs to reach more students as part of the IU Online initiative

• Oct 2012: Dr. Katy Börner is approached about translating her Information Visualization course into a MOOC format

• Nov 2012: Collaborative team assembled including Center of Innovative Teaching and Learning (CITL) and IU experts

• Jan 2013: Course launches
IV MOOC: Course Goals

- Scale Dr. Börner’s Information Visualization Course to anyone
- Survey state of the art information visualization technology
- Introduce data analysis algorithms to find patterns and trends in data
- Establish theoretical foundations, IV techniques across modes
- Provide theoretical foundation so students can navigate IVs in other settings
- Experience in creating unique and insightful visualizations
- Encourage collaboration: Allow students to collaborate in a highly collegial environment with each other and real world clients on actual information visualization projects

IV MOOC: Initial Planning Phase

- During the initial planning phase Google Course Builder was identified as a possible platform for this course
- After some research, we decided to go with Course Builder because:
  1. Open Source
  2. Highly flexible
  3. Scaled well to our needs (we were hoping for 2000 students in our first implementation of the course)
IV MOOC: Initial Planning Phase

- We decided to mirror Dr. Börner’s face-to-face Information Visualization course as much as possible
- Working with real world clients on actual visualization projects was an essential part of the course
- Students in Dr. Börner’s face-to-face class will also take the MOOC, resulting in a hybrid style MOOC

Information Visualization MOOC: Instructors

**Katy Börner - Theory**
Instructor, Professor in Department of Information and Library Science at School of Informatics and Computing

**David E. Polley - Hands-On**
CNS Staff, Research Assistant with MIS/MLS
Teaches and Tests Sci2 Tool

**Scott B. Weingart - Client Work**
Assistant Instructor, PhD student in Department of Information and Library Science at School of Informatics and Computing
IV MOOC: Course Schedule

Course Started on January 22, 2013

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

Midterm Exam

- Week 5: “With Whom:” Trees
- Week 6: “With Whom:” Networks
- Week 7: “Dynamic Visualizations and Deployment

Final Exam - March 10, 2013
Final Projects - March 18, 2013

IV MOOC: Unit Structure

Theory: Videos and Slides
Self-Assessment

Hands-On: Videos and Wiki
Homework
IV MOOC: Tools Taught

Science of Science (Sci2) Tool
- Open source
- Modular
- Supports analysis and visualization of:
  - Temporal data
  - Geospatial data
  - Topical data
  - Network data

Gephi
- Open source
- Interactive visualization platform
- Network
- Bridge from Sci2
IV MOOC: Tools Taught

Pajek

- Free program for Windows
- Visualization and analysis of large networks

IV MOOC: Client Projects

- 15 clients provided their own data for unique, real-world visualization projects
- Students formed groups of 4-5 to work on client projects
- 13 student projects for 9 different clients
IV MOOC: Client Project Results

- Final grades were based on Midterm (30%), Final (40%), and Client Project (30%).

- All students that received more than 80% of all available points get an official letter of accomplishment and badge.
IV MOOC: Development

Google Course Builder 1.0
- Implemented aesthetic changes to match CNS branding
- Sought to create a unique feel to the course interface

IV MOOC: Development

Drupal Forum
- Allowed students to communicate with other students and clients
- Facilitated group formation with student profiles
- Provided a way for students to submit work and get feedback from instructors
IV MOOC: Social Media

Twitter

Flickr

IV MOOC: Course Tour
IV MOOC: Analysis

- 1,780 students from 93 different countries
- 1,517 registered in time to successfully complete all the work and earn a badge
- 58 students earned a letter of accomplishment and a badge

Student Demographics

<table>
<thead>
<tr>
<th>Age</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>62+</td>
<td>2%</td>
</tr>
<tr>
<td>51-60</td>
<td>8%</td>
</tr>
<tr>
<td>41-50</td>
<td>17%</td>
</tr>
<tr>
<td>31-40</td>
<td>34%</td>
</tr>
<tr>
<td>21-30</td>
<td>37%</td>
</tr>
<tr>
<td>Under 20</td>
<td>1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ph.D.</td>
<td>15%</td>
</tr>
<tr>
<td>Graduate</td>
<td>56%</td>
</tr>
<tr>
<td>College</td>
<td>22%</td>
</tr>
<tr>
<td>High School</td>
<td>6%</td>
</tr>
</tbody>
</table>

Profession

- 40% Other
- 11% Government
- 11% Not for profit
- 11% Industry
- 31% Faculty
- 11% Student
IV MOOC: Analysis

IV MOOC: Future Iterations

- We will be running the IV MOOC again in January 2014
- This time the course will be taught by Scott Weingart
- Planned extensions include information visualization for the digital humanities and statistics
Acknowledgments

We would like to thank Miguel I. Lara and his colleagues at the Center for Innovative Teaching and Learning for instructional design support, Samuel Mills for designing the web pages, Robert P. Light and Thomas Smith for extending the GCB platform, and Mike Widner and Mike T. Gallant for adding the Forum. Support comes from CNS, CITL, DILS, SOIC, and Google.

The tool development work is supported in part by the Cyberinfrastructure for Network Science Center and the Department of Information and Library Science in the School of Informatics and Computing at Indiana University, the National Science Foundation under Grants No. SBE-0738111 and IIS-0613650, the US Department of Agriculture, the National Institutes of Health, and the James S. McDonnell Foundation.

Visualizations used in the course come from the Places & Spaces: Mapping Science exhibit, online at http://scimaps.org, and from the Atlas of Science: Visualizing What We Know, MIT Press (2010).

Open Discussion