Title of Project: 
Uncovering the Geography of Learner Engagement in MOOCs

Team Members:
College of Earth and Mineral Sciences, University Park
Anthony C. Robinson, Lead Faculty for Online Geospatial Education, John A. Dutton e-Education Institute
Sterling Quinn, PhD Candidate, Department of Geography

Designated Principal Investigator:
Anthony C. Robinson – arobinson@psu.edu – 814-574-0030
Lead Faculty for Online Geospatial Education
John A. Dutton e-Education Institute
University Park Campus

Date of Submission:
May 15, 2014

Amount Requested: $48,168

Project Time Period: (September 1, 2014 – June 30, 2015)
Uncovering the Geography of Learner Engagement in MOOCs

Abstract

This project will characterize the geographic patterns associated with learner engagement in Massive Open Online Courses (MOOCs). MOOCs generate datasets which capture a wide range of learner information, including their interactions with the course content, performance on assessments, and contact with other students through discussions. These data can be geolocated through the use of Internet Protocol (IP) address records, and new methods in development at Penn State provide the opportunity to extract the places learners talk about from their discussion posts. Contemporary research on MOOC learner engagement focuses on country-level geographic data gathered through survey or IP address data. Considering the interest in global learning communities that MOOCs create, there is surprisingly little known about how place matters beyond crude measures of the proportion of students coming from a given country. We will advance geographic knowledge about MOOC learners by investigating these core questions:

• What are the spatial patterns associated with learners in a MOOC?
• How do learner outcomes vary by space and time?
• How does learner engagement vary by space and time?
• Which places are relevant to learner engagement with course topics?

Proposed Innovation

One of the greatest value propositions associated with the investment in time and capital required to develop a MOOC is their ability to attract very large and globally-diverse audiences. Champions of MOOCs have lauded their potential for reducing barriers to access in higher education, especially among learning communities in less-developed countries [1]. At present, most of the published work that explores global dimensions of MOOCs focus on simple
measures of the raw proportion of students coming to a class from a given country [2]. While these lists tend to include hundreds of countries, there is frustratingly little detail of the geographic dimensions of MOOCs beyond these simple measures. Highly populated countries and cohorts with high broadband internet access tend to make up larger proportions of every class, but very little analysis has gone deeper to explore how cohorts within a MOOC may vary across spatial scales, or to uncover how student interactions in the course may vary across space and time as a course progresses from start to finish.

We see a clear linkage between our focus on geographic patterns in learner engagement and outcomes in MOOCs and our ability to develop improved models for MOOC instruction. A richer understanding of spatial patterns and dynamics in MOOC student populations would make it possible for us to develop place-specific targeted efforts to increase retention, for example.

**Research Aims**

Our project focuses on the following four broad research aims. For each of these broader aims, we provide two examples of specific, answerable questions for the initial project period.

- **What are the spatial patterns associated with learners in a MOOC?**
  - When we normalize by population or other factors, which places have the highest rates of interest, engagement, or disengagement with a course?
  - How does the geography of individual demographic groups compare to the aggregated whole?

- **How do learner outcomes vary by space and time?**
  - Where do certificate earning students live compared to those who drop out or who have interest in the content but not in earning a credential?
  - Does the geography of dropouts decline evenly, or does it have a spatial signature?

- **How does learner engagement vary by space and time?**

2
Where are the students who are the most actively engaged, and when are they active?

What are the temporal dynamics of student engagement, and how do they relate to course milestones?

Which places are relevant to learners when discussing course topics?

Which places do students reference when they talk about course concepts?

Do the places students talk about in the course correlate to their home locations?

**Project Methodology**

To answer our research questions we will analyze data gathered from two offerings of Maps and the Geospatial Revolution, a Penn State MOOC offered on Coursera. Combined, these courses have enrolled over 70,000 students. Maps MOOC datasets include discussion forum posts, clickstream logs of user interactions with the content, and pre/post-course survey data to capture demographics, motivation information, and other learner attributes.

Making sense of these datasets in relation to our research questions will involve the following key phases of work, categorized by each of our four research aims:

**Phase 1: Identify Spatial Patterns Associated with Learners in a MOOC**

2. Create a spatial database using PostgreSQL with PostGIS extension to link and index geospatial datasets from the MOOC.
3. Implement interactive OpenLayers web map to leverage the PostGIS database to select learner demographic categories and visualize a wide range of learner cohorts.

**Phase 2: How do learning outcomes vary by space and time?**

1. Develop queries to extract learning outcome cohorts (based on motivation and assessment data) from the PostGIS database.
2. Extend interactive web map to support spatial visualization of these outcome cohorts and compare with spatial patterns identified in Phase 1.

**Phase 3: How does learner engagement vary by space and time?**

1. Develop queries to extract time slices from demographic categories and learner outcome cohorts from the PostGIS database.

2. Extend interactive web map to support visualization of temporal dynamics in conjunction with demographic and learning outcome cohorts.

**Phase 4: Which places are relevant to learners when discussing course topics?**


2. Develop topic categories across discussion forum posts using a topic modeling tool such as the Stanford Topic Modeling Toolbox.

3. Integrate geolocated forum data (and topic model categories) into PostGIS database.

4. Extend interactive web map to support spatial visualization of topic data in relation to places that students discuss.

Throughout this effort, we will rely on the OpenLayers open source web mapping platform. Sterling Quinn, a collaborator on this project, is an expert on developing open source web mapping technologies using OpenLayers, and this platform coupled with a PostGIS database is a common design pattern in interactive web mapping today.

To evaluate the impact of our research beyond our analytical findings from the Maps MOOC, we propose to extend our project into 2015-16 and compare our results with sample data generated by other Penn State MOOCs. We would approach this task through the development of a generic framework with which data coming from any MOOC could be evaluated from multiple geographic and temporal angles. We would leverage contacts made through COIL to
connect to other MOOC instructors and demonstrate the potential value of collaboration with our team in this way.

In addition to publishing our findings, a key dissemination goal for our project is to provide our geolocated datasets as well as our mapping interfaces and any programming code we develop to support other ongoing MOOC research projects at Penn State.

To execute our project, we envision a partial overlap between each phase given their dependencies, with additional overlap early on to deal with inevitable data integration issues. Less time is required for Phase 4 given our preliminary progress on those methods as seen in [3].

Funding Rationale

This project requires funding to develop the initial geospatial datasets and visualization products necessary in order to support follow-on proposals to funding agencies. The effort to geolocate MOOC data requires an initial investment in time and technical expertise before useful results can emerge. Finally, this project provides value for the rest of the PSU MOOC research community through the development of shared geolocated datasets and analytical approaches to augment existing and future MOOC data research projects.

Future Research Potential

The anticipated results from this project will provide the catalyst to support a follow-on funding proposal to the National Science Foundation Geography and Spatial Science Program. We will also explore opportunities to submit a proposal that links to the Research on Learning in Formal and Informal Settings set of programs.
Qualification of Key Personnel

The research team is composed of Anthony C. Robinson and Sterling Quinn. Anthony Robinson is the Principal Investigator for this project, and serves as Lead Faculty for Online Geospatial Education in the John A. Dutton e-Education Institute in the College of Earth and Mineral Sciences. In this role, he directs the Master of GIS degree program as well as several certificate programs, which serve over 1000 adult learners annually with more than 30 online courses. In addition, Anthony is the Assistant Director for the GeoVISTA Research Center in the Department of Geography, where he leads geographic visualization interface design and evaluation efforts on externally funded research projects. He has taught online courses since 2008, and in 2013 developed one of Penn State’s first MOOCs on Coursera.

Sterling Quinn is a PhD Candidate in the Department of Geography. Sterling began his PhD studies at Penn State in Fall 2013, following an eight-year stint working as a senior software developer for Esri, the world’s largest geographic information systems company. Sterling earned his Master of GIS degree from Penn State online in 2009, and upon entering the PhD program in 2013 became the first student in Penn State Geography history to have made such a transition from our online programs. Sterling has taught in the MGIS program since graduating in 2009, and he has developed two courses that focus on geospatial software development, including our all-new GEOG 585: Open Web Mapping course. Sterling brings to this project a unique mixture of technical savvy required to execute on our project aims as well as deep experience with online learning, both from the perspective of a former student as well as an online educator.

Our team could easily expand to include faculty, students, or research staff from the College of Information Science and Technology to aid with our text processing objectives and to

Supplementary Material 1
provide data mining expertise. If Mr. Quinn is able to complete his Comprehensive Exams before Spring 2015 (which we anticipate), his tuition remission costs will drop considerably, and we will seek additional collaborators at that time from IST to fill these roles.

**Proposed Budget**

<table>
<thead>
<tr>
<th>COIL RIG Budget</th>
<th>Fall 2014</th>
<th>Spring 2015</th>
<th>Summer 2015</th>
<th>Totals</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncovering the Geography of Learner Engagement in MOOCs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 Month of PI Salary</td>
</tr>
<tr>
<td>Salaries (Category I)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 Semesters GA Stipend</td>
</tr>
<tr>
<td>Graduate Assistant (Category II)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fringe for PI Salary</td>
</tr>
<tr>
<td>Fringe (Category I)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fringe for GA</td>
</tr>
<tr>
<td>Fringe (Category II)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tuition Remission for GA</td>
</tr>
<tr>
<td>Tuition Remission</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Summer Wage for GA</td>
</tr>
<tr>
<td>Wage Payroll</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Semester Totals                |           |             |             |                 |                                    |

| Budget Request                 |           |             |             |                 |                                    |

**Budget Justification**

Dr. Robinson will devote one month of time (8.3% of a full year) to directing, implementing, and disseminating results from this project, split into equal parts across the 2014-2015 academic year. This funding will apply toward his available summer salary.

Two semesters of Graduate Assistant stipend are requested to support Sterling Quinn, at a Grade 12, half-time appointment level during Fall 2014 and Spring 2015. Tuition remission is also requested for both semesters, although we believe it is likely that Mr. Quinn will be post-
Comps in Spring 2015 and we can see significant tuition savings as a result, which will support the addition of further research personnel from IST to improve the impact of our project.

Fringe benefit charges are included for Dr. Robinson’s salary and Mr. Quinn’s graduate stipend, calculated using current Penn State rates for their respective appointment categories.

Finally, we request $XXXX for wage payroll to pay Mr. Quinn to complete the project during the Summer of 2015. This amount will support 20 hours of work per week during the Summer at a rate comparable to the normal Grade 12 GA stipend.

No travel, equipment, or consulting charges are requested. Dr. Robinson will support any related conference travel to present research results through in-kind funds provided by the online geospatial education programs that he directs. Computing, lab, and IT support infrastructure to carry out this project already exists through the GeoVISTA Center and we do not anticipate the need for direct support of those items.

References