

## Experience with Online Teaching with Open Source MOOC Technology

*Geoffrey Fox 10 June 2014*

**Abstract:** This memo describes experiences with online teaching in Spring Semester 2014. We discuss the technologies used and the approach to teaching/learning.

**Scenario:** We have taught from December 2013 to now, a course that was originally called X-Informatics but later renamed Big Data: Applications and Analytics. This course has four separate offerings so far

- 1) A MOOC with no credit and no tuition. This started first and continues with a total of 1637 enrolled students as of June 9 2014; about 6 new students enroll every day. This class is online only and does not offer instructor or AI grading. A largely meaningless “piece of paper” is available to those who complete class.
- 2) An Informatics Undergraduate class I400 offered spring 2014 semester with 55 residential students and regular tuition and credit. This had regular instructor/AI grading.
- 3) An Informatics Graduate class I590 offered spring 2014 semester with 22 residential students and regular tuition and credit; most students were Computer Science Masters students. This had regular instructor/AI grading.
- 4) A Data Science Certificate Graduate class I590 offered spring 2014 semester with 12 non-residential students and (discounted to ~instate value) tuition but regular credit. This class was online only. This had regular instructor/AI grading.

For offering 1) link is <https://bigdatacourse.appspot.com/preview>

For offering 2, 3, 4) link is <https://bigdatacoursespring2014.appspot.com/preview>

Most of the the web material is hosted on Dropbox/Box@IU/Google resources to reduce traffic and hence cost; Google App Engine for base site and YouTube for videos. The videos can be watched either through YouTube directly or embedded in the App Engine web page.

Google Course Builder is an app built for Google App Engine that offers free hosting upto a certain quota

The building blocks of these classes were

**A online course resource** built with Google Course Builder and enhancements CGL Mooc Builder <http://moocbuilder.soic.indiana.edu/> built by us and available as open source that allow convenient assembly of the different course components. These components include

5-15 minute video segments called **lessons** and containing curricula material (instructor desktop often containing PowerPoint slides).

Lessons are assembled into **units** totaling around 45 minutes – 2 hours and roughly equivalent to a traditional class.

Units linked into **sections** that together make up a coherent description of a major topic in course; for example “introduction” “Big Data and the Higgs Boson” and “Cloud Technology” are sections in these classes

The offerings 2), 3) and 4) share the same online site with 14 sections; 33 units and 220 lessons totaling 28.7 hours of video. The average lesson length was 7.8 minutes with 52 minute

average for units and sections averaging just over 2 hours with a maximum length of 5 hours 18 minutes. Offering 1) was similar but had earlier versions of material.

Each lesson had a video located on YouTube and an abstract (called lesson overview in figure 1 below). This interface show all lessons (13) for this unit and that each unit has its own abstract and slides available. There are also a list of follow-up resources associated with units and illustrated at bottom of figure 1. In the middle of figure 1, one sees the link to YouTube hosting of this lesson and 3 discussion links; one for each offering (2), 3) and 4). These are described later.

The screenshot displays a MOOC interface for a course titled "INFO 400 (Undergraduate), INFO 590 (Graduate & Data Science Certificate)". The user is logged in as "gcfexchange@gmail.com" and has options for "Logout" and "Profile". A navigation bar shows 29 units, with unit 13 selected. The main content area features a video player showing a slide titled "Big Data Use Cases" by Geoffrey Fox, Indiana University. The slide text includes: "Based on work of NIST Big Data Public Working Group (NBD-PWG) June-September 2013", "http://bigdatawg.nist.gov/", and "Leaders of activity: Wo Chang, NIST; Robert Marcus, ET-Strategies; Chaitanya Baru, UC San Diego". To the right of the video is a list of 13 lessons, with "13.1 Government Use Cases I" selected. Below the video are buttons for "Watch on YouTube", "Unit Slides", "Unit Overview", "Lesson Overview", and "Next Lesson". At the bottom, there are three discussion links: "INFO 400 Discussion", "INFO 590 Discussion", and "INFO 590 DSC Discussion". A "Resources" section at the bottom lists seven links (a-g) related to cloud standards, NIST working groups, and use cases.

Figure 1: A typical lesson (the first in unit 13) showing features discussed in text. Note links to all units across the top (29 of 33 units shown)

Note that we have a course – section – unit – lesson hierarchy (supported by Mooc Builder) with abstracts available at each level of hierarchy. The home page has overview information (shown in figure 2) plus a list of all sections and a syllabus with latter shown in figure 3.

Course
gcfexchange@gmail.com | [Logout](#) [Profile](#)

### 1. COURSE SUMMARY

This course INFO 400 and INFO 590 credit course investigates the use of clouds running data analytics collaboratively for processing Big Data to solve problems in Big Data Applications and Analytics. Case studies such as Netflix recommender systems, Genomic data, and more will be discussed.

### 2. FOR IU STUDENTS

1400 Undergraduate class for residents will have enhanced content, Python, conventional grading.

1590 Graduate class for residents: further enhanced content, Java and Python, conventional grading.

Data Science Graduate class for remote students: further enhanced content, Java and Python, IU course credit with normal grading.

### 3. OTHER INFO

Homework for your class will be posted in Oncourse. Grading will be also done conventionally using [IU Oncourse](#).

Course Material: [syllabus](#), [course files](#), [slides](#), or all-in-one zip file. Use the appropriate Google community forum for course discussions. [INFO-400](#), [INFO-590](#), and [INFO-DSC](#).

Please enroll only using a GMAIL account

#### Instructor

Professor Geoffrey Fox received a PhD in Theoretical Physics from Cambridge University and is now Professor of Informatics and Computing as well as Physics at Indiana University, where he is director of the Digital Science Center and Associate Dean for Research and Graduate Studies at the School of Informatics and Computing. He previously held positions at Caltech, Syracuse University and Florida State University.

He has published around 1,000 papers in Physics and Computer Science, supervised the PhD candidacies of 65 students, and received an h-index of 67 along with over 23000 citations. Professor Fox currently works in applying

#### IMPORTANT LINKS

- [Contact Us](#)
- [Team](#)
- [Tell us what you think](#)
- [FAQs](#)

#### OTHER MOOCs AT INDIANA UNIVERSITY

- [Big Data MOOC](#)
- [Cloud MOOC Repository](#)
- [Future Grid MOOCs](#)
- [Educational Assessment BOOC](#)
- [Information Visualization MOOC](#)

Powered by Open Source CGLMOOC Technology built on [Google Course Builder](#)

Figure 2: Course Home Page with Overview material

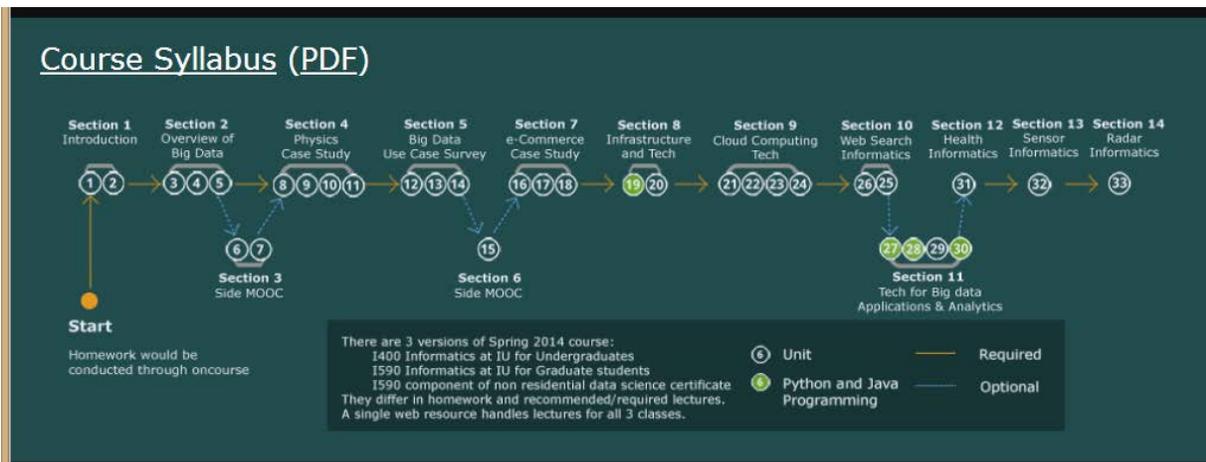


Figure 3: Course Home Page showing Syllabus

The screenshot displays a course navigation interface. At the top, there is a 'Course' button, a user email 'gcfexchange@gmail.com', and 'Logout' and 'Profile' buttons. The main content area shows a list of sections, each with a plus sign to expand it. Section 11 is expanded, showing a video abstract and a list of constituent units. The abstract text describes a K-means clustering example using Python and SciPy. The units listed are:

- Unit 27 - Technology for X-Informatics: K-means (Python & Java Track) - 42min
- Unit 28 - Technology for X-Informatics: MapReduce - 30min
- Unit 29 - Technology for X-Informatics: Kmeans and MapReduce Parallelism (Python & Java Track) - 17min
- Unit 30 - Technology for X-Informatics: PageRank (Python & Java Track) - 29min

Below the units, sections 12 through 14 are also visible, each with a plus sign to expand it.

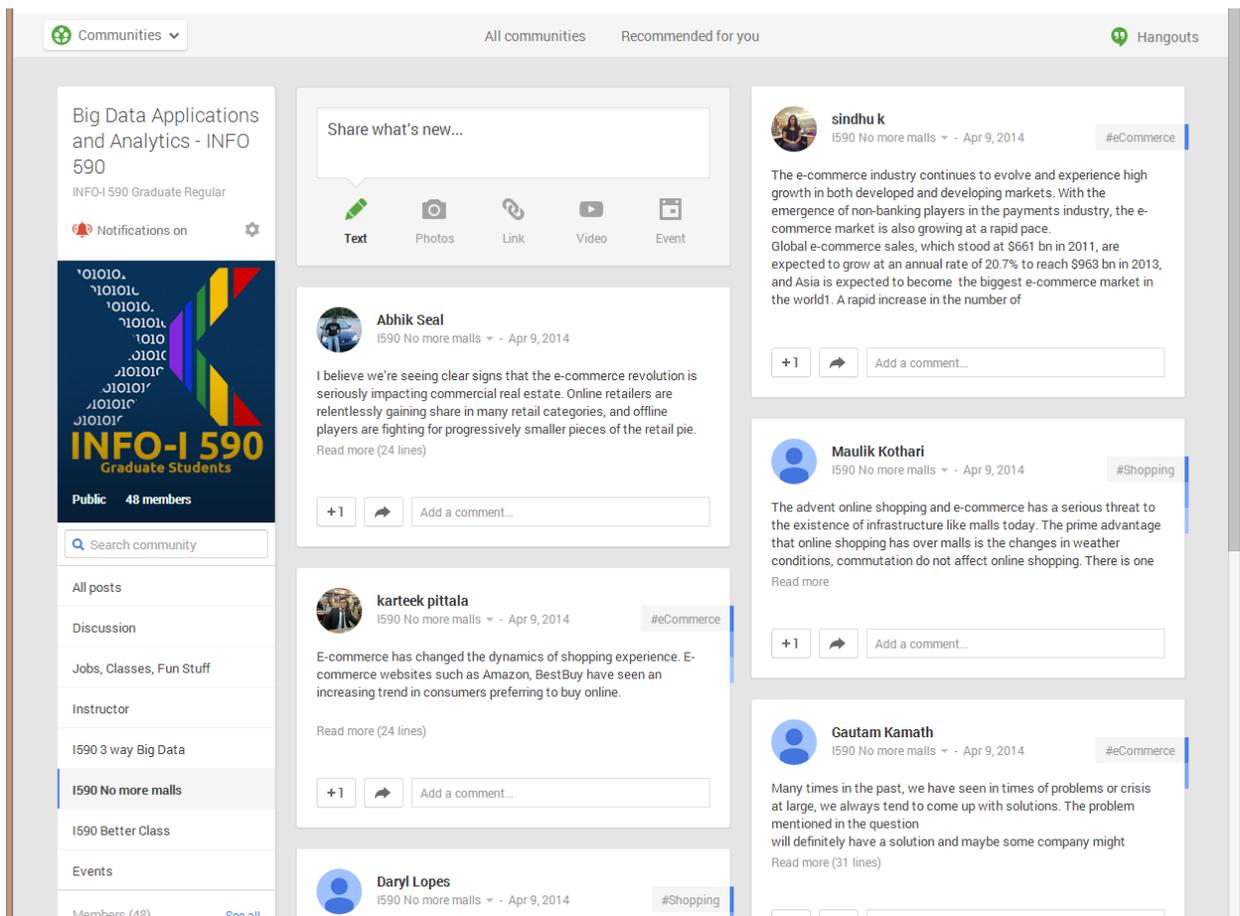
Figure 4: List of Sections with one (Section 11) expanded to show abstract and constituent units.

Figure 4 shows a partial list of sections showing how one can interactively browse the hierarchy. The next level would expose a unit as shown in figure 1.

**Homeworks.** These are online within Google Course Builder for the MOOC 1) with peer assessment. In the credit offerings 2-4), all graded material (homework and projects) is conducted traditionally through Indiana University Oncourse. Oncourse was additionally used to assign which videos should be watched each week and the discussion forum topics described later (these were just “special homeworks in Oncourse). In the non-residential class 4), the students were on a variable schedule (as typically working full time and many distractions; one

for example had faculty position interviews) and considerable latitude was given for video and homework completion dates.

**Discussion Forums:** each offering had a separate set of electronic discussion forums which were used for class announcements (replicating Oncourse) and for assigned discussions. Figure 5 illustrates an assigned discussion on the implications of the success of e-commerce for the future of “real malls”. The students were given “participation credit” for posting here and these were very well received. Our next offering will make greater use of these forums. Based on student feedback we will encourage even greater participation through students both posting and commenting. Note I personally do not like specialized (walled garden) forums and the class forums were set up using standard Google Community Groups with a familiar elegant interface. These community groups also link well to Google Hangouts described later.



*Figure 5: The community group for one of classes and one forum (“No more malls”) described in text*

As well as interesting topics, all class announcements were made in the “Instructor” forum repeating information posted at Oncourse. Of course no sensitive material such as returned homework was posted on this site.

**Hangouts:** For the purely online offering 4), we supplemented the asynchronous material described above with real-time interactive Google Hangout video sessions illustrated in figure 6. Given varied time zones and weekday demands on students, these were held at 1pm Eastern

on Sundays. Google Hangouts are conveniently scheduled from community page and offer interactive video and chat capabilities that were well received. Other technologies such as Skype are also possible. Hangouts are restricted to 10-15 people which was sufficient for this course. Not all of 12 students attended a given class. The Hangouts focused on general data science issues and the mechanics of the class.

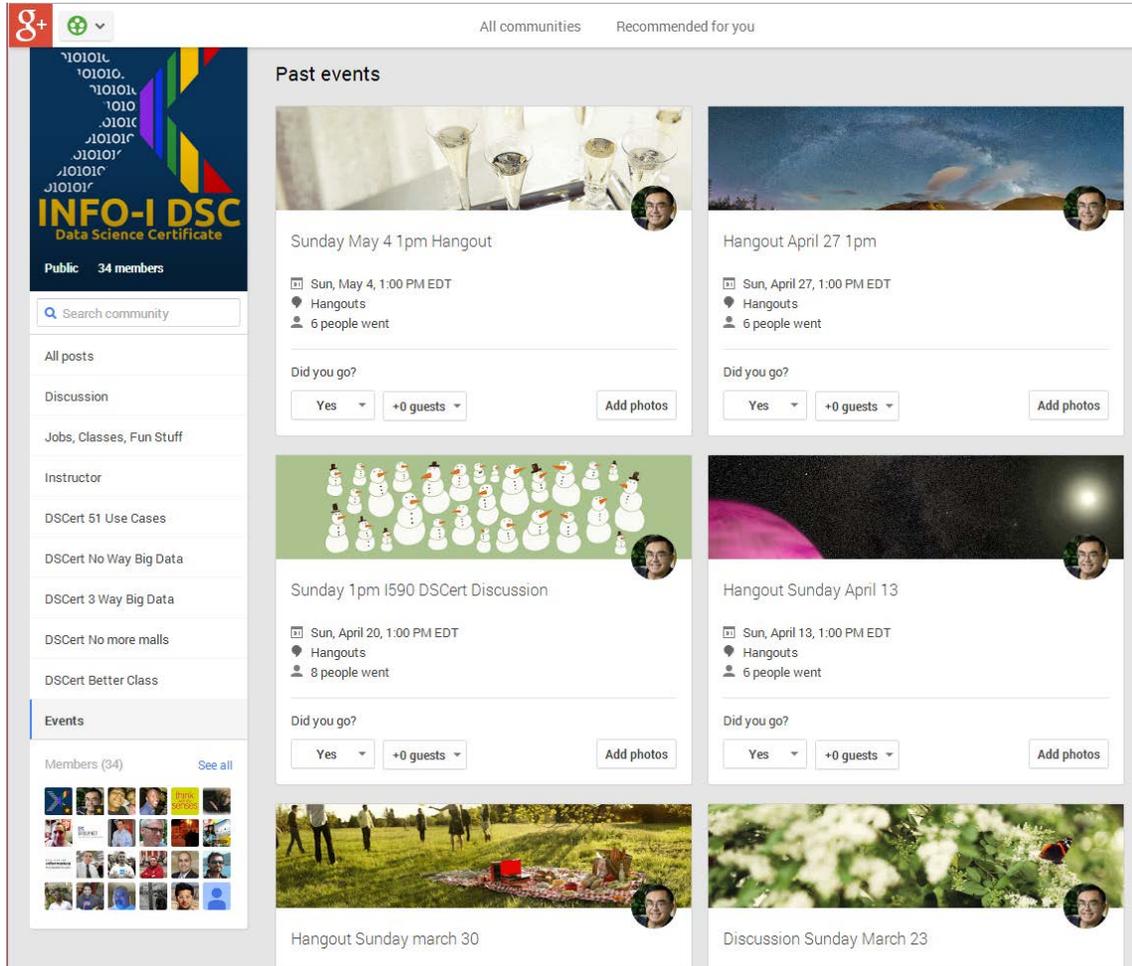


Figure 6: Community Events for Online Data Science Certificate Course 4).

**In class Sessions:** The residential sections 2) and 3) had regular in class sessions; one 90 minute session per class each week. This was originally two sessions but reduced to one partly because online videos turned these into “flipped classes” with less need for in class time and partly to accommodate more students (77 total graduate and undergraduate). These classes were devoted to discussions of course material, homework and largely the discussion forum topics. This part of course was not greatly liked by the students – especially the undergraduate section which voted in favor of a model with only the online components (including the discussion forums which they recommended expanding). In particular the 9.30am start time was viewed as too early and intrinsically unattractive.

**Workflow:** The above approach has no revolutionary ideas and is perhaps even prosaic. It did however get developed by several false starts and here we capture the actions that would be taken by a faculty copying the above strategy.

Let's assume we start with a traditional class with therefore an existing curriculum and content. Then we start by dividing material into sections and units. This should be easy as unit is roughly material for a traditional class.

Now we take a unit and prepare it for online recording. I did this in low-end style using a high quality web-cam (Logitech HD Pro Webcam C920, 1080p at ~\$73 from Amazon) on my laptop with attention to appropriate lighting and a black backdrop. I then took the slides for unit and inserted PowerPoint dividers separating this into lessons. This is convenient for recording as you know any video will be divided at these separators so you can take a deep breath at each divider. The lesson is recorded using Adobe Presenter. In my case this was either done as PowerPoint in full screen mode or an application (browser or IPython window perhaps). I recorded one or two units at each session. The final Presenter material is then edited – we used Camtasia – into lessons to be uploaded to YouTube. The next step is to write the necessary abstracts and assemble the online course using Moco Builder. Obviously the material needs to be checked. I note students found the video quality and organization of web site as quite good.

Operationally when the course is in session, the instructor needs to monitor Oncourse and the Class community group and set each week, the class homework, videos to watch and discussion topic.

### **Analytics**

Course Builder at time does not support good student analytics. The only measure available to me was the ability to look up a student's peer reviewed assignments. We developed some analytics capability as part of the MOOC builder that offers basic demographics data such as Age, Country, Education, and Enrollment Number per day. There are intrinsic difficulties due to multiple web sites and currently for example it's hard to link YouTube access to the Google App Engine access.

### **Summary**

We described an approach that can either offer an unsupervised Mooc, a purely online credit class or a flipped class. We showed how an extended Google Course Builder, the traditional IU Oncourse system and Google community groups could be integrated. The relatively new features: Course Builder and YouTube videos and Community groups were well received by students. For the instructor the weakest part of the system were the analytics available to monitor student use of online resources. Some Demographics analytics are available as part of Moco Builder that may be useful for gaining insights.

### **Acknowledgements**

Sidd Maini led the design and construction of CGL Moco Builder available at <http://mocobuilder.soic.indiana.edu/>. He supervised carefully temporary students who helped in the video editing and course assembly,