Dynamic dashboard for educators and students in FutureLearn MOOCs: experiences and insights

Lorenzo Vigentini\textsuperscript{1,2,3}, Andrew Clayphan\textsuperscript{1,3}, Mahsa Chitsaz\textsuperscript{1}

\textsuperscript{1}UNSW Sydney Australia, Office of Pro-Vice Chancellor (Education)  
\textsuperscript{2}UNSW Sydney Australia, School of Education  
\textsuperscript{3}School of Computer Science & Engineering

l.vigentini@unsw.edu.au, a.clayphan@unsw.edu.au, m.chitsaz@unsw.edu.au

ABSTRACT. One of the differentiating aspects of the FutureLearn platform, compared with other MOOC providers such as Coursera and EdX, is the approach to data sharing with partners. This is grounded on the release of a small set of relatively simple source files, which can be downloaded and used as required by end users (e.g. educators, researchers and so on). This approach has both advantages and disadvantages. The major advantage is the simplicity; the most important drawback is the lack of an ‘out-of-the-box’ set of analytical representations which the end-user can use and digest to obtain immediate insights regarding their online course. In this paper, we discuss these aspects in more detail and document the approach adopted at UNSW Sydney, to use the data as released, and how we produced a set of analytical dashboards for educators and students. The architecture underpinning the dashboards built is explained with a link to a GitHub repository with more detailed information.

Keywords. MOOCs; visualization dashboard; learning analytics, FutureLearn.

1 Introduction

MOOCs have been around for several years and there are now platform providers hosting courses with a variety of educational designs reaching millions of learners each year [29]. With the establishment of this way of learning online, the increased availability of MOOC data can offer the opportunity to provide insights to educators and developers into learners’ behaviours, and empower learners to understand their patterns of engagement and performance through learning analytics [30]. The former allows exploring learning design at scale and has the potential to inform pedagogy. The latter can improve the learning experience and develop crucial metacognitive skills essential for self-directed and lifelong learners. As mentioned in [37], the field of learning analytics...
has seen a push to move from descriptive analytics to analytics able to inform and direct practice [11,41,42]. This was also advocated by Gasevic and colleagues as a key area of further research in their review of research in MOOCs [16] and took centre stage in many presentations at LAK ‘17 and in other Workshops about MOOCs presented in this volume [40].

There appears to be two critical problems hindering the application of learning analytic methods to support and shape pedagogy in MOOCs: 1) constraints of the platforms (e.g. tools and course design) and 2) the availability of data – when it is needed, by different stakeholders to gain a better understanding of learning and take action.

Looking at the wealth of research in MOOCs, a lot is done ‘post-hoc’, or after the course is completed, when the respective platforms release data for exploration. Furthermore, data is often restricted to individual institutions, limited by contractual agreements with platform providers, and this makes it very hard to draw generalizable conclusions. Nevertheless, research has been carried out using Coursera data [4,26] focusing predominantly on the dashboard offered to partners’ institutions [12]. In addition, the relative openness of EdX, allowed different teams to develop extensions/plugins to access and use analytics [8,25,28,32]. Yet, to date, the focus has been primarily on educators and administrators rather than students.

2 The UNSW way of ‘moocing’

UNSW entered the MOOC space with the specific intent to learn at scale and take this back to mainstream education. This meant that the key driver for the development of different MOOCs was research-oriented and experimental in nature. This afforded the use of multiple MOOC platforms, which allowed for different learning designs and social engagement structures to be implemented across the deployed course offerings. This approach has been documented in various outlets [5,6,15,24]. This shows that learning design took center focus, and with it, the teachers whom were part of the MOOCs were able to push the boundaries of the technology – which has oft been criticised for its inability to deliver beyond the hype of connectivist approaches.

Experimenting with multiple platforms, using very different philosophies has been an essential aspect of the process. To date, UNSW Sydney has delivered ten distinct courses in Coursera; twenty in FutureLearn; a handful in SmartSparrow (mostly closed, on-campus courses focussed on Engineering and Medicine) and a dozen courses in OpenLearning (on-campus courses, namely in computer science).

Another important aspect has been the intentional integration (with other tools via LTI) and differentiation of the offerings, based on learners’ own preferences (elements of personalisation at scale via interactive activities and adaptive tutorials). Examples of this approach are demonstrated in the ‘Learning To Teach Online’ MOOC in Coursera [23], featuring a system to recommend pathways through the material based on learner’s preferences and the grouping mechanism in the ‘Entrepreneurship’ course in Coursera evolved from successful on-campus blended experiments [2]. In FutureLearn, adaptivity and personalisation featured in the integration of LTI activities using SmartSparrow (‘Through Engineers’ eyes’) [14,18], and activities
aimed to interact with participants in ‘Personalised Medicine’ and ‘Disability’ courses. All these, within the constraints and affordances of each MOOC platform, provide insights into the experimental approaches to course design taken.

Another element of experimentation was done with the actual development and funding of different courses. Multi-disciplinary teams have been used to support the academic leads (content experts). Driven by the idea of curriculum alignment [1] and the RASE model for representing elements of course design [6], the MOOC design process at UNSW focussed on ongoing conversations between educational developers and academic leads to match and fit technology and tools to their pedagogical intentions. Finally, experimentation was done with different levels of de-centralisation to support the planning, design and implementation of the various courses.

Figure 1 provides an overview of the familiar learning analytics cycle [7], which may be implemented to develop and evaluate a single course (in the middle), but the transition to mainstream is an essential component of the larger cycle situated at institutional level, in which the focus is not only the improvement of a single course, but a more systemic improvement across all educational provision from an institution.

Given the context, a natural extension of such course experimentation at scale has allowed us to explore certain areas with evidence-based approaches to, these are: 1) the evaluation of projects and programs related to learning and teaching [3,22]; 2) a desire to systematically implement approaches and systems aiming to enable the use of learning and teaching data to improve pedagogy and 3) bridging the gap between traditional academic development approaches and data-driven approaches with Learning Analytics and Educational data mining [38].

In this paper, we present the work we completed with the FutureLearn platform, and we make available our work efforts to aspiring and other course development teams.
3 Visual exploration as an aid to pedagogical design

As mentioned earlier, the use of dashboards to support sense-making from learning and teaching data is not new. There are several examples in which the use of visualisations provide a good starting point for discussing course design and student learning. Our specific focus was to offer tools for educators and educational designers, allowing the visual exploration of data to help understand the way in which learners engage with different elements of the course and provide valuable information to inform future course designs.

To achieve these goals, several streams of work were carried out, resulting in a systematic process based on aspects identified to be used starting points for development:

- Collate questions from diverse stakeholders (including educators, developers, learning designers and academic managers)
- Come up with the definitions of key terms (i.e. – What is engagement? What do we mean with learner experience?)
- Explore design variants (Coursera/FL dashboard examples)
- Reduce to smaller components/simplify
- Re-connect with academics and educational developers
- Obtain feedback from learners, by getting them to use the dashboards
- Refine and reiterate

At the conceptual level, defining what is intended with terms like engagement and the learning experience is essential to inform the process put in place to measure these constructs. Vigentini and Zhao [39] provide an overview of the evaluative tools used in MOOCs in their recent meta-analysis of several platforms.

At the practical level, previous research [19,31,36] identified five key areas which appear to direct the attention of educators and developers in MOOCs: 1) An overview of the course, 2) Who are the participants/learners, 3) How participants interact/engage with the material, 4) How participants interact in the forums, and 5) How participants perform in the course.

In the design of our dashboard around the FutureLearn Data, we took these areas into account as well as principles for dashboard design to showcase information to achieve one or more objectives [13]. It was important for us to keep the representations simple, avoid overwhelming users with information [5], thus we employed a minimalist design combined with appropriate structuring of information.

3.1 The purpose of dashboards

Intuitively, dashboards are only useful if there are actionable insights. In this sense, data visualizations have often been used as tools to help teachers gain insight about how students engage with the content and resources provided in learning environments, [33,34], but can also be offered directly to students to help them reflect on their approach to learning [9,17].
Three areas or domains of representation have proven popular: 1) Learner activity, 2) Learner engagement, 3) Learner experience. Learner activity focuses solely on a descriptive level of what students do in the learning environment. The analysis of student engagement moves further, and attempts to make sense not only of what students do, but also question the effectiveness of activity (usually correlating activity with outcomes, exploring in detail – assessment and the way in which students communicate and interact with each other as well as the platform, or digging into why students behave in the way they do, or attempting to model behaviour by enriching the logs with learning and teaching metadata – for example about the course or learning design). The third level of analysis is an abstraction of what the students’ experience is like, usually emerging from surveys or what students voice about their learning experience (before, during or after taking part in the course).

All of these provide valuable insights in the effectiveness of learning design leading to more effective evaluation and informing redesign [Vulic], but, if done consistently and in a timely manner, can also provide useful information to monitor progress, and allow for much quicker interventions while the course is running.

3.2 Why another dashboard?

Over the past year UNSW developed and delivered twenty MOOCs in the FutureLearn platform. This provided an exceptional opportunity to put into action the work done in this space, by using data as the course unfolded. As mentioned earlier, at this point in time, FutureLearn still does not offer a fully-fledged dashboard to partner institutions, but rather simple to understand accessible files, served up on a daily basis, of which a static report of activity can be generated.

The granularity of the information provided is thus enough to provide some insights, without overwhelming stakeholders. Given the near real-time availability of data, this provides a major opportunity to respond to students’ engagement as a course unfolds.

4 A practical approach to dashboard development

As it was the case for the design and development of MOOCs, the data and evaluation team in the PVC Education portfolio at UNSW Sydney took an experimental stance, creating minimum-viable products and moved from exploration and representation (i.e. supporting sense making of learner activity) to the characterization of patterns related to engagement with elements of the course. In the Coursera dashboard [36], the focus was on reporting what students did, and this was organised under functional headings (i.e. the use of videos, content, forums/discussion, assessment/learning activities and evaluative activities such as polls and surveys).

As will be presented in the following sections the approach tackled four distinct problems/phases: 1) cater for different systems, different formats, different philosophies; 2) determine appropriate processes for designing dashboards; 3) identify core dashboard ‘building blocks’; and 4) develop and test through a structured development process (Figure 2).
There were two fundamental drivers for the development:

1. Studying learning engagement focusing specifically on active participation, measured via proxies of attempts/completion of learning activities and the volume of discussion. Admittedly neither measures learning directly, but the characterisation of the patterns provide insights in the way that participants go about their learning.

2. Beginning to include more detail about the learning experience and what participants said about their motivations, quality and satisfaction when learning through MOOCs.

What we strived to achieve was to put analytics into practice and empower both students and educators with the information and knowledge derived from the data.

![Fig. 2. Learning and teaching dashboard development processes at UNSW Sydney](image)

### 4.1 FutureLearn MOOC structure and data

Each MOOC published in the FutureLearn website is presented in a hierarchical structure with weeks, activities and steps (Fig. X). Steps contain different types of material and can be recognized by the label next to the step title. Typically, in the FL courses designed at UNSW, eight different step types were used: article, discussion, video, exercise, quiz, test, audio and LTI activity.

At the time of writing, FutureLearn provides eight separate data sources prepared as comma separated values (CSV) files. Table 1 describes each file with detail about purpose. The datasets provided in FutureLearn have three main limitations: 1) granularity of user activity (currently limited to the time of the first/last access rather than a full
interaction log); 2) minimal contextual information (lack of metadata about the learning context, such as video interaction data), and 3) partial demographic information to understand learners (only about 10% of participants have chosen to share personal details in the platform). These limitations are the byproduct of FutureLearn’s choice to provide easy to access and stable datasets. Nevertheless, these sources provide an excellent starting point to demonstrate the use of analytics in action.

![Diagram of FutureLearn MOOC structure with Weeks, Activities and Steps](image)

**Fig. 3.** The Overview of a FutureLearn MOOC structure with Weeks, Activities and Steps

Although each file can be used in isolation to answer particular types of questions, in order to gain deeper insights, the data required transformations and integration.
Table 1. FutureLearn Datasets.

<table>
<thead>
<tr>
<th>File</th>
<th>The purpose of the file</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrolments</td>
<td>This file provides basic information regarding enrolled learners (and staff). It includes demographic information of learners derived from responses to FutureLearn’s ‘more-about-you’ survey, which captured gender, country, age range, highest education level, employment status, and employment area.</td>
</tr>
<tr>
<td>Step Activity</td>
<td>This file stores information regarding step activity from learners, e.g. the time when a step is first visited, and the last time a step is marked as completed.</td>
</tr>
<tr>
<td>Comments</td>
<td>Information about learners’ contributions to the discussion in each step is noted. The file includes the full-text of comments, timestamped according to when the comment was made. Likes associated with comments are also stored.</td>
</tr>
<tr>
<td>Question Response</td>
<td>This file holds information about the quiz activity of learners. It stores learners’ responses, whether correct or not, and when the quiz was answered.</td>
</tr>
<tr>
<td>Team Members</td>
<td>Information about organization staff members including their FutureLearn ID and full names are stored in this file.</td>
</tr>
<tr>
<td>Peer Review</td>
<td>This file provides information regarding peer review assignments including when the assignment was first viewed, when it was submitted and the number of reviews associated with the assignment.</td>
</tr>
<tr>
<td>Assignment</td>
<td></td>
</tr>
<tr>
<td>Peer Review</td>
<td>This file provides information about the reviewers on an assignment, including when the review was submitted, the reviewer’s ID and feedback for each of the assignment guidelines.</td>
</tr>
<tr>
<td>Reviews</td>
<td></td>
</tr>
<tr>
<td>Campaigns</td>
<td>Information about the referral used to advertise a course is stored in this file, following the number of enrolments and active learners for each referral.</td>
</tr>
</tbody>
</table>

4.2 The technology stack

Based on our previous experience in building Learning and Teaching dashboards with Coursera MOOC data where we choose Tableau cloud to present insights, [35,36], we quickly identified key differences in the process required to develop a FutureLearn dashboard. The choice of appropriate tools was driven by previous experience combined with a preference for wanting to experiment with different products and services to realize different perspectives.

In the case of Coursera data, despite being able to present information in a rich and interactive way, the solution adopted was fit for a scenario in which the data was available only at the end of the course. This meant that after the first preparation step was completed, a data analyst could build an entire dashboard based on the Tableau templates by simply adapting the dashboard to the new course by fitting the released dataset and thus authoring a dashboard to the cloud.

However, given that data from FutureLearn is available daily, one of the main challenges was to develop a sustainable, dynamic and near real-time dashboard. This meant the choice of an appropriate technology stack to make the update process regular and seamless. Additionally, in a similar way to the use of dashboard templates for Coursera, keeping the visuals up-to-date, meant a deciding on a series of re-usable building blocks...
– thus further allowed common questions to be answered.

In Coursera, the technology used focused on the preparation and pre-processing of the data, predominantly leveraging on Python to create a replicable processes and data products ready to be used in Tableau.

In FutureLearn, given the relative simplicity of the dataset provided, what we wanted to focus on was more of the analytical process and therefore shifted toward the use of R scripts in conjunction with a Shiny Server for publishing of outputs developed across the analysis areas chosen. Fig. 3 provides an overview of the technologies adopted in the two cases. Notably the front-end technologies did not change drastically between the two implementations.

<table>
<thead>
<tr>
<th>Coursera (sessional)</th>
<th>FutureLearn</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Csv, json, mysql</td>
<td>- Csv, R, sql</td>
</tr>
<tr>
<td>- Python, sql</td>
<td>- Html, bootstrap, Rstudio</td>
</tr>
<tr>
<td>- Html, angular js, tableau</td>
<td>server, shiny server, d3js</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coursera (on demand)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Docker, csv, postgres</td>
<td></td>
</tr>
<tr>
<td>- Python, sql</td>
<td></td>
</tr>
<tr>
<td>- Html, angular js, tableau</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 4.** A parallel between the technologies used at UNSW in the implementation of dashboards for data coming from Coursera and FutureLearn.

### 4.3 From post-hoc to near live representation

A simple architecture was developed around a two-stage model: 1) data extraction and pre-processing and 2) dashboard development. This allowed automation of the download and preparation of data, where more time could be spent on the customization of dashboard views to enhance the end-user experience (note: the dashboards were accessible via the web, via authentication – password / LTI) (see Fig. 5). As noted earlier, the R platform was used as the analytical engine, with a cloud-based Shiny server with an Apache server to serve the web content.

For the R dashboard creation process, Python was used to pull the data daily (a scheduled job), transformations were applied, and outputs stored in cloud-based MySQL database. In the data extraction and pre-processing phase, Python scripts were used to automatically login into the FutureLearn platform and download all available files for each course. Each time new CSV files were downloaded, they are stored, named as per the course title, with previous CSV files archived. After downloading all course files and loading the data into the database, the pre-processing routines (written in R) prepared the data, transforming the source files into views for specific purposes. This step was essential to make the web data requests scalable at run-time by querying
pre-computed information sets, keep the hosting server size minimal and also keeping costs down.

Fig. 5. Architecture for automating dashboard development for FutureLearn MOOCs.

4.4 Experimenting with student dashboards

Although there are several examples of the use of dashboards with educators, there is a lack of examples with learners. Corrin et al. [9,10] discussed the issues and difficulties of exposing data to students, but others like Liu and Pardo [21], have demonstrated how careful selection of specific data representations for students can lead to action. Here we present our current work-in-progress efforts which have focused around two key questions: (1) what data is useful to students? and (2) which user interfaces are likely most effective for the type of data being presented?

Based on previous research and continuing conversation and engagement with students, we propose a framework for student dashboard elements in four areas:

- **Learning Communities**: this includes an awareness of their social context, the nature/structure of interactions with others in this context and learning conversations;
- **Student Progress**: focuses on an awareness of the learning space in which a student learns, and provides a sense of the learning activities required to achieve learning goals. This also means that students can obtain markers (feedback) to motivate and promote self-regulation (i.e. articulate their goals in the context and know where they are compared to others);
- **Student Performance**: focuses on both formative and summative elements of the course providing an overview of strengths and weaknesses as well as clear references to other students’ performance. This feedback is essential to inform and provide evidence for action; and
- **Student Experience**: this is harder to measure and in most cases, comes from surveys about learners’ satisfaction with different aspects of the course.
Further, based on the previous work with educators’ dashboards we specifically focused on four elements of design: 1) Layout, 2) Chart types, 3) Chart features and 4) Other visual/functional features.

Table 2. Basic dashboards design building blocks.

<table>
<thead>
<tr>
<th>LAYOUT</th>
<th>CHART TYPES</th>
<th>CHART FEATURES</th>
<th>OTHER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tabs</td>
<td>Scatter plots</td>
<td>Titles</td>
<td>Tables</td>
</tr>
<tr>
<td>Navigation Bar (Located at the top)</td>
<td>Bar charts</td>
<td>Drilldown</td>
<td>Show/Hide Elements</td>
</tr>
<tr>
<td>Side menu Bar</td>
<td>Line/spline charts</td>
<td>Hover</td>
<td>Styled Value boxes</td>
</tr>
<tr>
<td>Radio Buttons</td>
<td>Pie charts</td>
<td>Legend</td>
<td>Textboxes</td>
</tr>
<tr>
<td>Accordions</td>
<td>Heat maps</td>
<td>Labels</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Choropleth maps</td>
<td>Symbols</td>
<td></td>
</tr>
</tbody>
</table>

5 The code repository overview

The repository of our work-in-progress efforts is shared in GitHub under a AGPL GNU licence. This is intended to foster collaboration and provide benefits for both end-users and others who may want to contribute to developments of the dashboard.

In the repository, there are two separate streams/folders: A dashboard for educators/administrators and a dashboard for students. These are now presented.

5.1 A set of dashboard building blocks

In this section, we provide an overview of the dashboard design and the visual building blocks used. At this point in time, the dashboard design is based on a minimalist html design template which dynamically loads the shiny dashboard applications to an iframe. An alternative to this is to use a native Shiny Dashboard ‘fluid page’, but this would require additional maintenance coding between individual apps. Displaying visuals in the iframe kept the design of the dashboard scaffold separate from the R scripts of each application, and promoted rapid iterative development.

A more extensive description of the various elements of the dashboard can be found in the Git repository [https://github.com/moocunsw/FL-dashboard]. The key questions driving the overall structure are in the table below.
Table 3. Key questions driving the visual representations and menu structure.

<table>
<thead>
<tr>
<th>Learners and engagement</th>
<th>Aspects of evaluation</th>
<th>Areas of research</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Who are the learners?</td>
<td>• Have design goals been achieved?</td>
<td>• Instructors-driven questions</td>
</tr>
<tr>
<td>• Why are they taking the MOOC? (precursors / motivation)</td>
<td>• Were learning outcomes achieved?</td>
<td>• pedagogy and design</td>
</tr>
<tr>
<td>• What are the learners’ explicit goals?</td>
<td>• Was the experience worthwhile?</td>
<td>• data mining and learning analytics</td>
</tr>
<tr>
<td>• How did they engage with instruction / content / activities?</td>
<td>• Overall evaluation of effectiveness</td>
<td></td>
</tr>
<tr>
<td>• How did they interact with others?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspects of educational design</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Course characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Intended learning outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Assessment and activities</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.2 A dashboard for students: what they may find useful

In this section, we provide a description of the dashboard currently used in the FutureLearn course “Enhancing Learning and Teaching in Higher Education”. The tool is a set of Shiny applications wrapped into a simple LTI application. The LTI connector allows users to authenticate to the application, leveraging on the learner ID to present a personalised data view of the data for that use. The following four Shiny apps are available in the GitHub repo:

- **myClass**: gives an overview of the participants’ geographical distribution, gender and ages.
- **myCommunity**: provides a basic representation of the learners’ interactions with others in a network, and a view of the learning discussions in the course (based on popularity (views and likes) as well as general sentiment).
- **myProgress**: displays a detailed account of the learners’ engagement with the content/activities of the course. The visual display also provides a ‘snapshot of the course dna’ [27] as it allows to see what types of activities are available.
- **myPerformance**: focuses on the quizzes, showing a detailed account of responses.

At the time of writing the dashboard is about to be deployed to a live FutureLearn course, so as to collect student feedback across the duration of the course, to inform both future designs and as a measure of success of the implementation.

6 Future work

The work presented here shows a clear trajectory of development which UNSW has taken to bring analytics to both educators and students in the MOOC space. The expe-
rience provided several opportunities to build internal capacity, but, by sharing the artefacts of this work, we hope to provide a solid starting point for others just getting started. By sharing this work, we also hope that is seen as an opportunity to begin to build a community of practice toward pushing the implementation and envelope of learning analytics practice. Of course, automating the development of an analytics dashboard for the FutureLearn platform for all would be a very desirable goal for both FutureLearn and the partner institutions.

The development process allowed us to explore alternative ways to implement a dashboard that other FutureLearn partner institutions may find useful, with several different visualizations explored, as well as consideration of related literature [31]. Furthermore, a direct comparison with a similar effort by Leon-Urrutia et al. [20] demonstrates the viability and effectiveness of the implementation.

Our implementation provides an opportunity to consider possible ways to use the tool with both educators and students. However, considering earlier discussion, we also show a cautious increase of sophistication in the student dashboard, compared with the educator dashboard. The increase of sophistication, however, is highly dependent to what FutureLearn will provide in the future. There is plenty of opportunity for more data to be exposed, for example: in-video behaviours, clickstream information, and user-session based information. All these are important to explore in more detail, particularly in trying to determine and judge on/off task behaviours, retention and resilience as well as relating survey data about goals and motivations to performance and achievement.

7 References

35. Lorenzo Vigentini and Andrew Clayphan. 2015. Pacing through MOOCs: course design or teaching effect? In Conference on Educational Data Mining (EDM).


