Data Life-Cycle Management's Massive Open Online Course on Research Data Management

Silas Krug
Information Sciences Dept.
Geneva School of Business
Administration
Carouge, Switzerland

Basma Makhlouf Shabou
Information Sciences Dept.
Geneva School of Business
Administration
Carouge, Switzerland
ORCID 0000-0003-0980-0517

Abstract—*In this paper, we will first have a quick look at what is a MOOC (Massive Open Online Course) and e-learning in general. Then we will present the five main modules which structure the course, who is responsible of which one, and what are the three optional specialized modules. We will then talk about the established partnerships for the project, with Swiss MOOC Service and DevPro. Then we will have a look at the content types which you will find in this course, and finally list the future project's milestones.*

I.   INTRODUCTION

In its mission to be the reference for Research Data Management (RDM) in Switzerland, DLCM (Data Life-Cycle Management) must provide accurate learning tools. The European Commission appeals to include e-learning in HEI's (Higher Education Institution) training strategies, which is why one of the DLCM's deliverables is to create MOOC (Massive Open Online Course) on RDM. This MOOC will have the following objectives :

- Provide accessible online materials to researchers,

- Support HEI units that assist researchers,

- Assist swiss and external researchers,

- Offer a complementary resource to OLOS users,

- Promote DLCM and OLOS services and expert network.

The MOOC belongs to family of e-learning methods. Several definitions have been given to the e-learning, such as the one from the Commission of the European communities in 2001, « *[the e-learning is the] use of new multimedia technologies and the Internet to improve the quality of learning by facilitating access to resources and services as well as remote exchanges and collaboration* », the one from the Walloon Telecommunications Agency in 2008, « *E-learning (e-learning) : online learning centred on the development of skills by the learner and structured through interactions with the tutor and peers* », and the one from the Joint Information Systems Committee (JISC) in 2017, « *learning facilitated and supported through the use of technology to support learning as part of a 'blended' approach (a combination of traditional and e-learning approaches), to learning that is delivered entirely online. Whatever the technology, however, learning is vital element* ».

There are several varieties of e-learning, notably depending on the teaching timing which can be simultaneous or not. The MOOC is one of these sorts and here is what the words composing the acronym imply :

- Massive : without pre-requisites, the number of enrolments is potentially higher than for universities or schools.

- Open : because registration is open to all, without any special conditions (registration at a university, linked to a level of study or professional status, etc.)

- Online : the whole course is online, courses, homework and exercises, but also registration and exams.

Also, two kind of MOOC can be distinguished :

- xMOOC is designed as a classic course with a teacher-defined progression.

- cMOOC is decentralized. The material is made available to each learner who chooses his or her own course.

DLCM's MOOC will be somewhere between both concepts, as the contents will be freely available at any time, but its five general modules will be involved in an evaluation process, with a defined time schedule, leading to a certification.

II.   DLCM's MOOC's general modules

The general modules are the following :

A. *GM1 – Research Data Governance*

This first module introduces RDM, giving its context in the institutions' governance. It clarifies RDM's role and impact on the institutions' policies and the work done inside.

B. *GM2 – Active Research Data Management*

The second module concentrates on the life-cycle of the research data in its active part, meaning its management while the actual research is conducted. It starts on several details regarding the capture, collection and creation of the data. It then explains how it is structured, described and as a final step, analyzed and visualized.

C. *GM3 – Research Data Sharing*

RDM's problematic is highly related to another one : the Open Science, and its philosophy. Open research data is at the crossroad of both problematics and has to consider ethical and legal aspects in order to be led correctly. The data sharing is also deeply related to practical considerations such as the interoperability factors ; these are also studied in this module.

D. *GM4 – Research Data Preservation*

This fourth module basically talks about the last step of the Active Research Data Management, but it requires to be studied separately. Its content will review basics of digital archiving, such as the Open Archival Information System standard (ISO 14721). This module presents the swiss official solution OLOS.

E. *GM5 – Data Management Plans*

Considering the fact that the usually most familiar aspect of RDM for researchers is the Data Management Plan (DMP), for the simple reason they are regularly appealed to produce one to obtain research funding, this specific module is mostly about practical considerations around DMP's creation and use in researchers daily activities. It is focused on DMP addressed to the Swiss National Science Foundation (SNSF), but it will be universally useful for any kind of DMP.

The lead of all these modules have been distributed between very competent personalities in each of the respectively concerned domain. Dr. Basma Makhlouf-Shabou, co-director of the Data Life-Cycle Management project and responsible of the Master of Information Studies at the Geneva School of Business Administration, has major knowledge in information sciences and knows in particular how to manage an institutional governance, since she was project manager of the *Information and organizational governance policy : from design to implementation* project, from 2014 to 2016. Considering these abilities she took the responsibility of the module 1 on research data governance. In the DLCM project, she is also in charge of the Coordination Desk management, e-mail address that receives requests of help from researchers. Current statistics show that the most usual demands are DMP corrections, which is why she also took the lead of the module 5. The lead of the second general module is shared by Dr. Basma Makhlouf-Shabou and Eliane Blumer, coordinator of the EPFL's Research Data Team. Dr. Basma Makhlouf-Shabou is teaching research methodology in her IS Master, and thus is very familiar with the active research data management. Eliane Blumer is also in regular contact with the EPFL's researchers and, in particular, have a deep active life cycle. The third module is taken in charge by Dr. René Schneider, responsible of Information Studies at the Geneva School of Business Management. As a specialist of Linked Open Data, the Open Science and research data sharing largely belongs to his field of competences. Finally, Dr. Pierre-Yves Burgi, director of the DLCM project, and deputy IS director at the University of Geneva is responsible of the fourth module, on Research Data Preservation. Belonging to the development team of OLOS, the DLCM's solution for research data preservation, he is naturally well indicated to lead this part of the course.

III.   DLCM's MOOC's specialized modules

There will be four optional modules which can be studied as complement, depending on students' interests. Three of them have exclusive contents, but the fourth will be a compilation of what has been told about life-cycle facilities in each general module. The three exclusive modules are the following :

A. SM1 – Research Data types and paradigms

In this module, different types of data, depending on the nature of the research they are resulting from (quantitative or qualitative) and on their file formats (text, picture, sound, mixed, etc.), are studied. The main goal of this specialized module is to make discover the difference of management induced by these data types.

B. SM2 – RDM fields of study

This module will review of RDM depending on different academical fields of study. Some fields can have important specificities such as high caution on data sensitivity in medical fields, massive quantity of data in nuclear research or few actual data creation in literary studies.

C. SM3 – Tools, technologies and services

This final specialized module will review existing solutions and their features, open source or not, whose use covers the entire data life-cycle or only a specific phase.

IV. CONTENTS' CREATION

This MOOC is created in close contact with Swiss MOOC Service. Like DLCM, Swiss MOOC is a P-5 project of swissuniversities, what makes the collaboration between the two projects obvious. Within the framework of the MOOC project, Swiss MOOC provides in particular four services which are:

- An introduction to the arcane of shooting MOOC videos for the DLCM team as well as for some external contributors.

- The availability of their studio, in the EPFL premises, for the shooting of the videos.

- The post-editing of the videos shot on location.

- And the addition of subtitles to the videos, customizable in four languages.

For practical considerations, and also because the HEG, in a sense thanks to the pandemic crisis, now possess its own video studio, some of the footage will be directly shot there. Some video editing will also not be delegated to Swiss MOOC but done internally at the HEG with the software Camtasia. Finally, the MOOC's hosting will be done in collaboration with Cyberlearn, responsible of the HES-SO continuing formation, which is the official partner for the DLCM trainings' organization.

The didactic contents of the different modules will take various forms. Naturally, the emphasis will be placed on videos, which will consist of interviews and monologues, shot in the Swiss MOOC or HEG film studios, or in iconic locations. Other contents will have a classic PowerPoint format created on the base of a generic template, with or without voice cover to present them, and finally a few simple texts will complete the set. A short quiz is available for each lesson to ensure the learners test their comprehension of the presented content.

Although if the general structure of the modules will vary, each one will offer for between one and two hours of study (depending on the student's level), a bibliography of all the used sources, a follow-up certificate, as well as overall summaries in English, German, French and Italian. An advanced certification with fully assessment by OLOS experts is also available, delivered in collaboration between OLOS, the HEG-Geneva and DevPro. With regard to this last aspect, DLCM's aims at producing deliverables that are accessible to the entire Swiss population, with people who doesn't necessarily speak the project's language. Therefore, all national languages will be represented.

V. CONCLUSION

Today, three further milestones are planned for the MOOC : modules' certification for the end of March 2022, the introduction of a community management before summer 2022, and the completion of the specialized modules for February 2022. The subscription to the course implies that it will be possible to obtain an official certification by succeeding to scheduled exams. While the MOOC will permit to obtain an attestation for free, it will have a cost to subscribe to the advanced certification.

Built in collaboration with very specialists of RDM and MOOC's creation, with an ambitious plan and various didactic formats and experts from all places, DLCM's MOOC join a set of other trainings in the domain of RDM. Indeed, DLCM also organize half-day interactive trainings in collaboration with DevPro, and provides a Coordination Desk (dlcm@hes-so.ch), where any question on RDM can be asked, and receive an answer in the three following days.

This MOOC will probably become a major element of the training in the domain of RDM at the swiss level. No doubt that it will also be a very original experiment and an opportunity to open RDM training to a broader public, very familiar with the most advanced way of studying.

REFERENCES

AGENCE WALLONNE DES TELECOMMUNICATIONS (AWT). (2008). Qu'est-ce que l'e-learning ? *Site de l'Agence du Numérique* http://www.awt.be/web/edu/index.aspx?page=edu,fr,gui,080,010

ALLEN, I. E., SEAMAN, J. (2008). *Staying the Course : Online Education in the United States, 2008*. Needham : Sloan Consortium, November. https://www.onlinelearningsurvey.com/reports/staying-the-course.pdf

BENRAOUANE, S. A. (2011). *Guide pratique du e-learning : Stratégie, pédagogie et conception avec le logiciel Moodle*. Paris : Dunod.

COMMISSION OF THE EUROPEAN COMMUNITIES. (2001). *Communication from the Commission to the Council and the European Parliament : The eLearning Action Plan – Designing tomorrow's education*. Bruxelle, 28 March. http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52001DC0172

PLOURDE, M. (2013). MOOC : every letter is negotiable. *Flickr*. 4 April 2013. https://www.flickr.com/photos/mathplourde/8620174342/in/photolist-e6NW3q-dSyZ94-e8JDvy-gTdSqd-gskoKF-eejpte

POMEROL, J-C., EPELBOIN, Y., THOURY, C. (2014). *Les MOOC : conception, usages et modèles économiques*. Paris : Dunod.

ROSCORLA, T. (2012). Massively Open Online Courses Are 'Here to Stay'. *Converge*. 18 July. http://www.centerdigitaled.com/policy/MOOCs-Here-to-Stay.html

ROYAL HOLLOWAY. (2017). Defining E-learning. Royal Holloway Website. 13 October. https://www.royalholloway.ac.uk/e-learning/aboutelearning/what-is-e-learning.aspx

THOT CURSUS. (2016). Plates-formes de e-learning et e-formation – 2016. *Thot Cursus – Formation et culture numérique*. 21 August. http://cursus.edu/institutions-formations-ressources/formation/13486/plates-formes-learning-formation-2016/#.WddDPmc6w8A