Upskilling IT Professionals: A MOOC for the Data Science and IoT domains

Vasileios Gkamas
University of Patras, Greece

Maria Rigou
University of Patras, Greece

Alan Bruce
Universal Learning Systems, Ireland

Teemu Patala
Universal Learning Systems, Ireland

ABSTRACT

Given the overall Data Science and IoT impact on the European economy – the impact of the European Data Market on the EU economy as a whole exceeded the threshold of € 400 Billion in 2019 – there is a constant demand for experts at both domains. Although many training programs exist at different levels of education (i.e. higher education, VET) a skills gap in the domains of Data Science and IoT is observed. We present a MOOC for Data Science and IoT co-developed by academics and enterprises, as part of the SEnDIng vocational education and training program. The MOOC has been developed based on a modular multidisciplinary curriculum and has been positively evaluated by the trainees.

CCS CONCEPTS

• **Applied computing** → Education; Learning management systems; • **Social and professional topics** → Professional topics; Computing profession; Computing occupations.

KEYWORDS

Data Science, Internet of Things, Massive Open Online Courses, Vocational Education and Training

ACM Reference Format:
Vasileios Gkamas, Maria Rigou, Alan Bruce, and Teemu Patala. 2020. Upskilling IT Professionals: A MOOC for the Data Science and IoT domains. In *24th Pan-Hellenic Conference on Informatics (PCI 2020), November 20–22, 2020, Athens, Greece.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3437120.3437340

1 INTRODUCTION AND BACKGROUND

According to the latest estimates, the number of data professionals in the EU27 plus the UK reached € 76 million in 2019 (3.6% of the total workforce), but there is still an imbalance between the demand and the supply of data skills in Europe. The data skills gap is forecast to continue as demand will continue to outpace supply [1]. In the IoT domain, the skills demand has also been extremely high during the last years, with the total installed base of IoT devices projected to amount to 75.44 billion worldwide by 2025 [2]. A survey conducted in 2017 showed that 76% of enterprises need IoT specialists with more advanced skills and 80% did not have the skills needed to keep their IoT infrastructure working as it is [3]. Moreover, according to another report published in 2017, 1/3 of a survey's respondents declared that there is a major skills gap in IoT readiness [4]. In the next years we can expect the expansion and impact of the IoT to be even more important, especially when we look its potential when IoT is used in combination with other technologies like AI, machine learning and Big Data, domains closely related to Data Science.

We present a MOOC for Data Science and IoT developed by the Erasmus+ project SEnDIng. The SEnDIng project aims to address the Data Science and IoT skills gap of IT professionals, by providing them with knowledge, skills and competences that meet the needs of Data Science and IoT industries, are transferable and recognized among EU countries. To achieve this goal SEnDIng designed two multi-disciplinary VET programs comprising online training and work-based learning along with a horizontal course on soft skills.

2 MOOC FOR DATA SCIENCE AND IOT

2.1 MOOC design

The MOOC has been developed to provide cost-effective and improved learning experiences beyond those available in classrooms. The following pedagogical principles [5] have been adopted for MOOC design and for the successful e-learning provision:

- Match to the curriculum: there must be clear objectives, relevance to content covered, appropriateness of students' activities.
- Inclusion: inclusive practices should be seen in terms of different types and range of achievement, physical disabilities, different social and ethnic groups and gender.
- Learner engagement: learners should be engaged and motivated, activities should have a worthwhile educational aim, not just to occupy the learners, be enjoyable, not to produce adverse emotional reactions, improving the learning atmosphere.
- Effective learning: promoting personalized learning, learner autonomy; encouraging metacognitive thinking and collaboration, providing authentic learning exhibiting multiple perspectives on a topic.
- Provision of formative and summative assessment: This is essential for the purposes of improving and grade learners.
- Coherence, consistency and transparency: objectives, content, activities, and assessment should match to each other. It should be clear to the user what to expect.

PCI 2020, November 20–22, 2020, Athens, Greece

ACM ISBN 978-1-4503-8897-9/20/11…$15.00
https://doi.org/10.1145/3437120.3437340

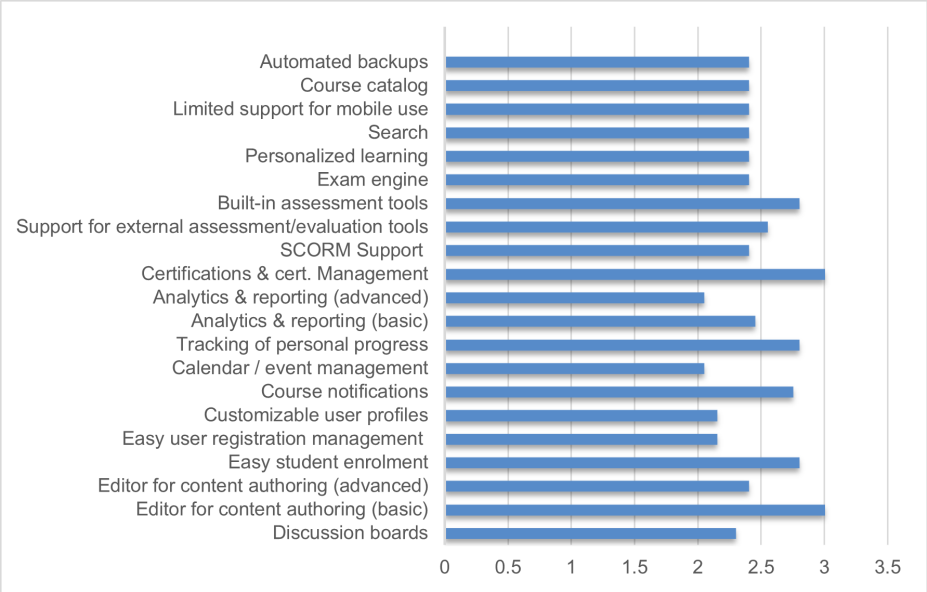

Figure 1: High scoring features for selecting the MOOC platform

- Ease of use: being open and accessible, intuitive and not requiring guidance on use, providing appropriate guidance to learners of teachers.

We conducted a survey among SEnDIng project's partners (i.e. HEIs, enterprises, associations of IT professionals, certification body) to define the target specifications of the MOOC. Each partner was asked to define the importance of the target features considering the training needs. The weight of each feature takes 4 values: 0 (useless), 1 (nice to have), 2 (useful) and 3 (must have). We gave categorized the importance of the target specifications into 4 categories based on the average score they got from the survey: (a) Low importance: average score 0.00 – 0.99, (b) Medium importance: average score 1.00 – 1.49, (c) High importance: average score 1.50 – 1.99 and (d) Very high importance: average score: 2.00 – 3.00. Figure 1 presents the features of the platform which have got an average score higher than 2, thus categorized as having very high importance.

2.2 Survey of MOOC solutions

Taking into account the aforementioned features, we surveyed 4 alternative platforms for MOOC development considering both open source and commercial solutions:

- Moodle (https://moodle.org) is an open source eLearning platform written in PHP and distributed under the GNU General Public License. It has been developed based on pedagogical principles and is used for blended learning, distance education, flipped classroom and other e-learning projects in schools, universities, workplaces and other sectors.
- Open edX (https://open.edx.org) is an eLearning technology that enables online campuses, instructor-led courses, degree programs, and self-paced courses using a single platform. It is available for desktop, iOS and Android versions, and provides access to course content and supporting infrastructure. It includes Open edX Studio a cutting-edge authoring tool which empowers learning and development through custom experiences powered by the latest in instructional design.
- Sakai LMS (https://www.sakailms.org) is an open source learning management system providing a wealth of powerful, flexible tools that enable great teaching, compelling learning, and dynamic collaboration. It includes a modern and easy to use User Interface, which thanks to its responsive design, enables instructors and students to achieve their academic goals.
- Cypher Learning NEO LMS (https://www.neolms.com) is a commercial eLearning platform. It is suitable for managing all classroom activities, whether it's creating classes, assessing students, facilitating collaboration, or tracking student achievement.

The survey was conducted based on a test scenario where the project partners have been asked to test the features of the 4 platforms and provide their feedback. The main results derived are summarized as follows.

Moodle is the most popular and widely spread open source learning management system, which in principle, has all the required features, at least through extensive customization. It is a stable platform, but some users find the functions a bit unnatural, and the look of the platform feels a bit outdated. Additionally, Moodle has chosen not to prioritized mobile-friendly development or APIs, which makes things difficult for some users. Open edX has better usability features and exploits APIs which foster easier use. Open edX covers most of the required features, at least through add-ons and plugins. It has an excellent customer base, positive references from the users and a large active community. Considering Sakai LMS, it has limited functionalities compared to Open edX and Moodle, and a rather limited community in Europe. On the other

Table 1: Data Science and IoT online courses

Data Science online courses	IoT online courses
Introduction to Data Science (DS-EM1)	Introduction to IoT (IoT-EM1)
Python for Data Science (DS-EM2)	IoT Devices (IoT-EM2).
Statistics for Data Science (DS-EM3)	IoT Communication Technologies (IoT-EM3).
Storing and Retrieving data (DS-EM4)	Architectural Design and Applications in IoT (IoT-EM4)
Applied Machine Learning (DS-EM5)	IoT Security and Privacy (IoT-EM5)
Data Visualization (DS-EM6)	IoT Business Value (IoT-EM6).

hand, Cypher Learning NEO LMS seems to be the strongest MOOC platform compared to the other alternatives by its visuality and clarity. However, it has a major drawback, since it is a commercial solution, and the associated cost prevents limited-budget projects from using it. Considering the open source solutions, Open edX is more oriented as a MOOC environment, compared to Moodle, which is more oriented as an LMS environment. For this reason and given that most of the training material developed will be available in the form of videos, Open edX is the MOOC platform of choice for hosting the SEnDIng online courses.

2.3 MOOC description

The MOOC is accessible at http://mooc.sending-project.eu and contains the Data Science and IoT courses listed in Table 1. Moreover, the curriculum includes an additional series of 5 courses (shared by both training programs) on Transversal Skills, namely Effective communication and presentation, Change management, Team working, Goal setting and Creative thinking. The macro-level design of the curriculum was based on a desktop research and a survey conducted among companies running Data Science and/or IoT projects [7], [8].

Each course, except for the 2 introductory courses, is divided into training units at three levels of proficiency: **Introductory** (covers the most important aspects of the course), **Core** (principles and methods of the course are covered in adequate depth, so that the learner can discuss matters with other stakeholders and acquire additional knowledge when necessary) and **Advanced** (advanced aspects of the course are covered in sufficient detail so that the learner can apply the knowledge and skills on the job).

Moreover, each course has an informational section which describes its content, the learning outcomes and assessment methods. The training material is cut down to small parts in either hypertext (enriched with images and code examples) or video (embedded at the MOOC). Each training unit ends with a self-assessment quiz and learners are offered the related training material as OERs for downloading. In addition, courses offer a set of projects learners can work on to test and enhance more their skills on the taught topic.

3 MOOC EVALUATION

The trainees have been asked to evaluate the MOOC at a scale ranging from 1 up to 5 (1: Strongly Disagree, 2: Disagree, 3: Neutral, 4: Agree and 5: Strongly Agree) based on the following 10 statements. 1: I have enjoyed the course, 2: This course was challenging, 3: The course meets my expectations, 4: The quality of the training material was high, 5: The content was well organized and easy to follow, 6: The course will be useful in my work, 7: The objectives of the course were clearly defined, 8: The time allocated for the course was reasonable, 9: The course enhanced my knowledge of the subject matter and 10: In this course, I have been challenged to learn more than I expected.

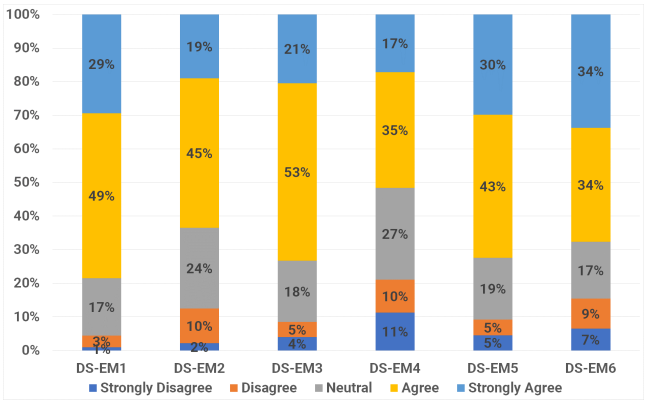

Figure 2: Evaluation of Data Science online courses

Figure 2 depicts the feedback collected for the Data Science online courses, while Figure 3 depicts the feedback collected for the IoT online courses. Considering the Data Science online courses, on average 68% of the trainees declared that they agree or strongly agree with the aforementioned statements. Among the 6 Data Science courses, the course "Python for Data Science" and "Statistics for Data Science" have collected the most positive feedback. Considering the IoT online courses, on average 80% of the trainees declared that they agree or strongly agree with the aforementioned statements. Among the 6 IoT courses, the courses "IoT Devices" and "IoT Business Value" have collected the most positive feedback.

4 LOOKING BACK AT THE BIG PICTURE: THE OVERALL TRAINING PROGRAMS

The MOOC supports the first of the three phases of SEnDIng training programs and its focus is to provide the new knowledge to trainees. Specifically, each SEnDIng VET program is delivered in three phases: (1) Online training (103 hours per program) through self-paced online learning; (2) Face-to-face training (20 hours) aiming to cultivate the transversal (or soft) skills of trainees; (3) Work-based learning (4 months), where trainees work on real-life case

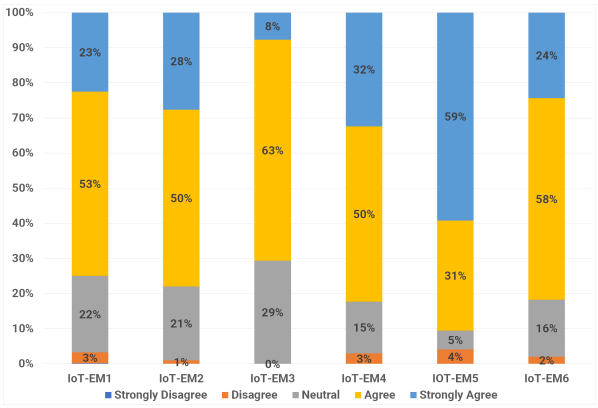

Figure 3: Evaluation of IoT online courses

studies in their workplace, so that they are transformed into productive IT professionals on Data Science and IoT [6]. To accomplish this, trainees implement well-defined and autonomous projects where they execute typical stepwise tasks in the respective domains, supported by in-company mentors [9].

5 OPPORTUNITIES AND CHALLENGES

COVID-19 lockdown has affected the initially planned project activities, as face-to-face transversal skills training was organized online, and part of the work-based learning and mentoring was done remotely. Drop out has been a challenge for the SEnDIng initiative, which is a very common problem to online learning and the COVID-19 pandemic has increased it. There were cases of trainees that started the online training, which was self-paced but at some point, before the start of the work-based-learning, were fired and had no company to 'host' them. It was decided that those of them that went through the transversal skills training, would be supported remotely (by the University of Patras and Cyprus teams) and provided with a series of projects to work on with technical mentoring, when necessary. Another way that the pandemic has affected the dropout rates was that after the long period employees stopped working or worked remotely, companies entered a phase where they have other priorities (than providing work-based training for their employees) and this in many cases resulted in employees not being able to allocate sufficient effort to work-based learning.

6 CONCLUSION AND FUTURE WORK

Our future plans include the support of personalized learning paths for acquiring skills on specific Data Science and IoT professional roles. According to the EDISON project [10], we can identify 5 professional profiles for data scientists (Data Analyst, Data Architect, DB Administrator, Machine Learning Engineer, Data Scientist). Moreover, SEnDIng proposed 6 scientific profiles for IoT professionals (IoT Product Manager, IoT Architect, IoT Software Developer, Data Scientist, IoT Cloud Engineer, and IoT Industrial Engineer) and provided a mapping between the training units provided in the respective training programs and these profiles for both domains [6]. The incorporation of this mapping in the MOOC in the form of learning paths that lead to specific professional profile certifications is a direction we already work on and consider important for project sustainability. This approach, along with setting up a mechanism of enriching the collection of hands-on projects and collaborating with certification bodies for designing certification schemes using the SEnDIng training content could keep the SEnDIng initiative active after the end of EC funding.

ACKNOWLEDGMENTS

This research was conducted with the support of the Erasmus+ programme of the European Union (contract no 591848-EPP-1-2017-1-EL-EPPKA2-SSA). The European Commission support for the production of this publication does not constitute endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. The authors would like to thank the SEnDIng project partners for their valuable input and feedback to the MOOC design.

REFERENCES

[1] European Commission. 2020. The European data market monitoring tool. Key facts & figures, first policy conclusions, data landscape and quantified stories: d2.9 final study report. Directorate-General for Communications Networks, Content and Technology, Directorate G - Data, Unit G1 — Data Policy and Innovation. doi: 10.2759/72084, ISBN 978-92-76-19505-4

[2] Statista I.H.S. 2018. Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025 (in billions). https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/

[3] Inmarsat Research Programme. 2017. The Future of IoT in Enterprise - 2017: Investigating the views of 500 enterprise IoT decision-makers. https://research.inmarsat.com/2017/

[4] CMO COUNCIL. 2017. The Impact of Connectedness on Competitiveness: Creating Business Value from Operational Insight Driven by the Industrial IoT. https://cmocouncil.org/thought-leadership/reports/the-impact-of-connectedness-on-competitiveness

[5] John Anderson and Robert Mccormick. 2005.Ten Pedagogic Principles for E-learning. Observatory of The Technologies and Education.

[6] Vasileios Gkamas, Michael Paraskevas and Maria Rigou. 2020. The SEnDIng project Data Science and Internet of Things professionals' training. *Information, Intelligence, Systems and Applications, 1(1),* 89-93.

[7] Vasileios Gkamas, Maria Rigou, Michalis Paraskevas, Thomas Zarouchas, Isidoros Perikos, Vaos Vassiliou, Ivaylo Gueorguiev and Pavel Varbanov. 2019. Bridging the skills gap in the Data Science and Internet of Things domains: A Vocational Education and Training Curriculum. *https:// wcol2019.ie*, 312.

[8] Vasileios Gkamas, Maria Rigou, Isidoros Perikos, Ivaylo Gueorguiev, Pavel Varbanov and Christina Todorova. 2019. Learning outcomes design for Data Science and Internet of Things training programs. In *10th International Conference on Information, Intelligence, Systems and Applications (IISA)* pp. 1-2. IEEE.

[9] Konstantinos Ampatzidis, Dimitrios Oikonomou, Paris Kitsos and Maria Rigou. 2019. A Smart Home Energy Management System Based on Internet-of-Things. In 2019 *Panhellenic Conference on Electronics & Telecommunications (PACET)* pp. 1-4. IEEE.

[10] Andrea Manieri, Steve Brewer, Ruben Riestra, Yuri Demchenko, Matthias Hemmje, Tomasz Wiktorski, Tiziana Ferrari and Jeremy Frey. 2015. Data Science Professional uncovered: How the EDISON Project will contribute to a widely accepted profile for Data Scientists. In: IEEE 7th International Conference on Cloud Computing Technology and Science, pp. 588-593, IEEE.