Analysis of MOOC Features in a Regional Platform: Design and Delivery of Courses

Annamaria De Santis(✉), Katia Sannicandro, Claudia Bellini, and Tommaso Minerva

University of Modena and Reggio Emilia, Reggio Emilia, Italy
annamaria.desantis@unimore.it

Abstract. At least two motivations make relevant research on regional MOOC providers: (1) they are Learning at scale environments where a large amount of data is produced, the high number of learners can show different approaches to learning, the whole population can be analysed, teaching and instructional design methods can be compared; (2) regional platforms reach the interests and needs of population groups that have only sometimes registered in international platforms.

Therefore, we conducted a study at a large scale on EduOpen, an Italian MOOCs platform to which 28 institutions joined. The research aims to describe the current situation on EduOpen MOOCs, identify a latent model for the design and delivery of courses, and detect points of interest to enhance the platform's procedure. We used courses as statistical units and identified three groups of variables: basic features, design features, and delivery features. We used multiple correspondence analysis (MCA) with descriptive statistics to answer the research questions. MCA is a method for data reduction for qualitative variables in which the categories assumed by the variables play a central role in defining a smaller number of dimensions.

We found two dimensions that define the structure that underlies the design process – COURSE DENSITY and CONTENT ATTAINABILITY, and the delivery process – DIFFUSION and PARTICIPATION. Some focal points that EduOpen members can consider improving strategies in the design of the courses are related to the organization of activities and content, the scheduling of activities and whole courses, the levels of interaction in the courses, and the definition of qualified professional figures for design and tutoring.

Keywords: Learning@Scale · Multiple Correspondence Analysis · EduOpen

1 Introduction

Numerous regional and global providers offer MOOCs (Massive Open Online Courses) all around the world. They differ in the number of registered learners, number and kinds of courses, language, business models, levels of openness, and so on [1].

At least two considerations have to be made on this globally diffused practice.

© The Author(s) 2023
G. Fulantelli et al. (Eds.): HELMeTO 2022, CCIS 1779, pp. 30–43, 2023.
https://doi.org/10.1007/978-3-031-29800-4_3

1. *MOOCs platforms are an example of Learning at scale (L@S) environments where teachers deliver the sequence of learning activities. L@S environments offer opportunities and challenges in data-driven research, teaching and instructional design methods.*

In global and national contexts, Learning at scale (L@S) environments host courses and resources realized by a few experts/tutors for a huge number of learners.

In these environments, the focus goes indeed to *learners*, that are *at scale* for their numerosity and the diversity of experiences, goals, opinions among them. At the same time, we can assign the attribute "at scale" also to *research and data*, *adaptation* (meant as mass personalization), *time and space* (to learn anywhere and anytime), *pedagogy* (as new opportunities to enhance educational experiences in the digital age) [2].

In particular, the research on these environments used data to answer educational questions. Even if it perhaps did not conduct the hoped disruption in education [3], it can provide pieces of evidence in studies on educational policies, students' behaviours, and social phenomena. In fact, interdisciplinary studies about course design conducted with experimental approaches on more learning platforms contribute to using educational data to offer new directions in learning science.

The collected data in L@S environments refer to the whole population and not only to a sample; the big dimensions of platforms allow comparisons among instructional design and teaching methods with explanatory, predictive, and normative purposes. Interpretation of the meaning of variables and results, generalization of results to a wide population, and confidence in statistical tests affect the validity of the studies and make us risk of obtaining ambiguous results [4].

Two elements are fundamental for better efficacy of the studies, according to [3]: (1) researchers have to propose studies that test theory and practices in learning, introducing different kinds of instructional design, and use granularity of data to plan instructional designs functional to reach better learning outcomes; (2) they have to design and choose measurement tools and indicators to conduct the studies effectively.

2. *Global and regional platforms reach different goals in the diffusion of open education.*

Recent multi-platforms studies [5–7] underline that different learners register into global and regional non-English platforms. Regional platforms can receive and reply to the interests and training needs of a limited population (if compared to large providers) that, in some cases, could look at national platforms as the only way to participate in the training because of the language, the participation of well-known national universities, or the cultural background of the lecturers, teaching methods and course design [6].

For example, comparing Arab students registered on EdX and Edraak – a Jordanian platform – in [5], authors show that younger, female, and less educated students have registered in Edraak. They also confirm, for example, that the completion rate for similar courses on the two platforms can vary (three times higher in the Edraak), probably for the higher level of participation or, it cannot be excluded, for the levels of difficulty of the courses. These elements of diversity could signal that regional platforms, better than global ones, can use MOOCs as instruments for sharing and democratizing knowledge

and training thanks to the role that language and culture can play in the collaboration and achievement of educational goals.

Starting from these two considerations, we propose an analysis on EduOpen, one of the small European regional Moodle-based platforms [1] for MOOCs delivery. Since 2016, about 400 courses were published on EduOpen (approximately one-third of all those produced in the Italian academic context), attended by more than 120 thousand who acquired about 90,000 attendance certificates.

This platform is part of the Italian MOOCs ecosystem that, according to Italian studies on open education [8], seems to be dynamic. In fact, about one thousand MOOCs was produced by 28 universities until 2021. Besides, Italian scenario in MOOC production appears more open for content licenses and accessibility mechanisms than similar countries, and fragmented, probably because of the lack of a national policy.

Nevertheless, EduOpen is the only network of universities and entities (28 at the moment) that offers MOOCs in a unique environment in Italy. Also for this reason, we defined this research as a *macro-level analytics* according to [9, 10]. Our analysis focuses on the strict relationship between research on MOOCs and Learning Analytics [11] and tries to identify a latent model for creating and delivering MOOCs in a cross-institutional context, EduOpen. Even if the analysis we present here is only realized on one platform, we considered it *macro* for the high number of universities/institutions involved and for the purpose of looking at the whole community and impacting the design and delivery processes of MOOCs. We didn't consider individual students' performances (*micro analytics*), groups of courses or individual institutions (*meso analytics*).

Applying a multivariate technique for data analysis, we try to reply to the following research questions:

- Q1. What dimensions can we identify in the design and delivery features of EduOpen MOOCs?
- Q2. Considering relevant events related to the platform updates in November 2018 and the Covid19 pandemic, did the course features change during the years?

The answers to Q1 e Q2 allow us to describe the current situation on EduOpen MOOCs, identify the latent elements in the course production and publication, and detect points of interest to enhance the platform's use.

The following sections describe the dataset and method used for the analysis (Sect. 2), the results (Sect. 3), and the conclusions (Sect. 4).

2 Method

2.1 Data

The research dataset comprised the total number of MOOCs published on EduOpen from 2016 to September 2022, which is equal to 418 courses.

We excluded 77 courses. In particular:

– 15 courses that stayed open for only a few weeks at the portal launch (2016) and didn't use all the rules defined by the coordination team and become the standard over the years;

– 12 courses published after June 2022;
– 39 courses called Capstone and used only for assessment in Pathway (groups of MOOCs that have a common educational goal);
– 10 courses called Courseware that are repositories of well-done teaching materials and not courses at all.

So, after the selection, the dataset consisted of 341 MOOCs.

2.2 Variables

We identified 13 variables retrieved from the design plans of the MOOCs (e.g., Course Categories, Starting Year, Training Hours, Language) and from the platform dashboard filled by logs and users' participation (e.g., number of Learners and Completion, number of Posts). Some of the variables considered, such as Number of Learners, were metric; we created ranges using quartiles to transform them into non-metric variables and apply multiple correspondence analysis (the method uses only qualitative variables). The variables in this procedure are Training hours, Activities/Hours, Duration, Number of Learners, and Completion Rate.

We divided variables into three groups: the first one collected Course categories and Starting year that represent the BASIC FEATURES of the courses; the second group detected the DESIGN FEATURES of the courses and comprised eight variables; the third one collected five variables related to DELIVERY FEATURES. The variables Duration and Editions were included in both the second and third groups because we believe that these elements influence the two phases of course realization. To clarify with an example: instructors can propose more than one edition of a course to take a tutored mode (Editions so can be considered a design feature) or because of good results in the participation in the first edition (Editions as a delivery feature).

A complete explanation of the variables is in Table 1. The first column contains the number of categories for each group of variables; the last one includes the percentage of courses belonging to each category.

2.3 Analysis Method

We used multiple correspondence analysis (MCA) with descriptive statistics to reply to the research questions.

MCA is a method for data reduction for non-metric variables in which the categories assumed by the variables play a central role in defining a smaller number of dimensions [12, 13]. This method has exploratory and descriptive purposes and aims to show the latent structure of a dataset through values and graphical representations (perceptual map) where the categories are plotted.

The distance of categories gives information on the relation among them. The proximity of the points in the perceptual map makes the similarity between categories visible. Points closer to the origin represent categories with frequencies more similar to the mean; points further away indicate modality values that deviate from the expected values. Quadrants and half-planes in the biplot can be considered in interpreting the distribution of points.

In this method, we can produce symmetric and asymmetric biplots that differ in how the distance among categories is calculated. In the former, the ones we used in our analysis, we can only compare categories of the same variable and make general comments on categories belonging to different variables.

The distance calculated between the frequencies of the categories leads to the calculation of an index called *inertia.* It represents the variance of the dimensions in which the variables are summarized; it also represents the distance of the categories from the axes in the visualization.

Other indexes are:

- `ctr` that denotes how much each category contributes to the dimension and so to the variability of the dataset.
- `cos2` that assumes values between 0 and 1 and is the representation quality level, i.e. how well the dimension represents the category.

We decided to use this method because we needed:

- to identify and describe a small number of elements (dimension) that can summarize the different features of a great number of courses;
- to include non-metric variables in the analysis that are particularly relevant in the design and delivery of the courses;
- to create visualizations that a large community of stakeholders could simply understand.

We used `R/RStudio` as analysis software and the packages: `factoextra` and `FactoMineR` to perform MCA [14].

3 Results and Discussion

About half of EduOpen MOOCs regard themes on Social Sciences (44.9%), mainly in sectors of Education, Economy, and Law. Arts and humanities (23.8%) is the following category. 87.4% of courses are delivered in Italian. The courses can be divided into two similar groups if we consider the delivery mode: tutored/self-paced. In Table 1, besides, we can observe that number of Posts/Participant is very low. These data suggest that interaction and collaboration processes are not beware in the design and delivery of the courses. 29.9% of courses are part of a pathway, and 57.7% are courses with multiple editions (the first or the following editions). Around 38.4% of courses have no closure data, so they can be considered “always open”.

To identify dimensions in the design and delivery features of EduOpen MOOCs starting from course characteristics (Q1), we performed two MCA. We used the variables related to Basic and Design features in the first one; in the second one, we included variables of Basic and Delivery features.

As can be seen, the inertia value is around 20% in both analyses. Even if this value is quite low and more information could be retrieved by adding more dimensions, we decided to propose a solution with two dimensions to obtain a graphical visualization of the latent structure of data.

Table 1. Variables table.

Group	Variable	Categories	%
Basic features (13)	Course categories	Arts and humanities Computer & Data Science Health and Pharmacology Sciences Social Sciences Technology, Design and Engineering	23.8 10.3 4.7 12.3 44.9 4.1
	Starting Year	2016 2017 2018 2019 2020 2021 2022	14.1 16.1 22.6 20.5 11.4 7.9 7.3
Design features (21)	Training Hours	0–10 h 11–14 h 15–18 h 19–100 h	29.3 25.5 21.7 23.5
	Language	Italian English	87.4 12.6
	Mode	Tutored Selfpaced	50.2 49.9
	Pathway	Yes No	29.9 70.1
	Activities/Hour (ratio of the number of activities and the number of training hours)	0–2.9 3.0–3.7 3.8–5.1 5.2–22.8	27.0 23.8 24.3 24.9
	Duration (number of days the course was available)	0–100 days 101–150 days 151–257 days 258–1193 days Open (No closure data)	15.5 15.9 15.5 14.7 38.4
	Edition	Yes (Multiple edition) No (Single editions)	57.5 42.5
Delivery features (19)	Number of Learners	< 230 230–472 473–1142 > 1142	24.9 25.2 24.9 24.9

(*continued*)

Table 1. (*continued*)

Group	Variable	Categories	%
	Completion Rate	0–21% 22–29% 30–40% 41–72%	25.8 27.0 23.8 23.5
	Posts/Participant (ratio of the number of posts and the number of participant that completed course)	0.0 0.1 ≥ 0.2	44.0 31.4 24.6
	+ *Duration* and *Edition*		

Figure 1 is the biplot resulting from the first analysis. In the graph:

- Hard Science categories are in the top right sector, Social Sciences in the bottom right quadrant, Art and Humanities in the left. So, we can observe a clear division among categories;
- Self-pace (SP) and Tutored (T) are in the opposite sector near the origin. They didn't affect the dimensions related to the design;
- early Starting years (2016/2017/2018), long Duration, and high Training hours categories are on the right side of the figure;
- categories related to More editions and Pathway are in the bottom sectors;
- higher ratios Activities/Training hours are in the top sectors.

Dimension 1 (*inertia* = 11.2%) can represent COURSE DENSITY**,** which refers to the quantity and organization of teaching materials in the MOOC. Going from the left to the right, we can observe a change in the different types of workload: on the left, we have top categories of Activities/Hour and bottom categories for Hours; the contrary is on the right where, besides, we find also the category of courses that belong to pathways. This difference can be related to two periods: on the right between 2016–2018 and the left between 2019–2022.

Dimension 2 (*inertia* = 10.1%) can be seen as CONTENT ATTAINABILITY, that is, the possibility of reaching the content and the knowledge in training. On the bottom, we have "more attainable" categories related to the presence of more than one edition for a course and the organization of MOOCs in pathways where themes are set out more broadly; on the top, we have categories of Hard sciences and English that is more difficult for Italian learners that are the main participants on the platform. Besides, the categories related to the realization of one edition for a course or long-duration delivery that define courses that could be less updated or less findable in the platform among recent or featured courses are in the same sector.

According to `ctr` values, more involved variables are: Duration, Training hours, Pathway, Activities/Hour, Editions.

The sum of the inertia of the two dimensions is 21.3%. This means that these two dimensions explain the variance of the same percentage of the observations.

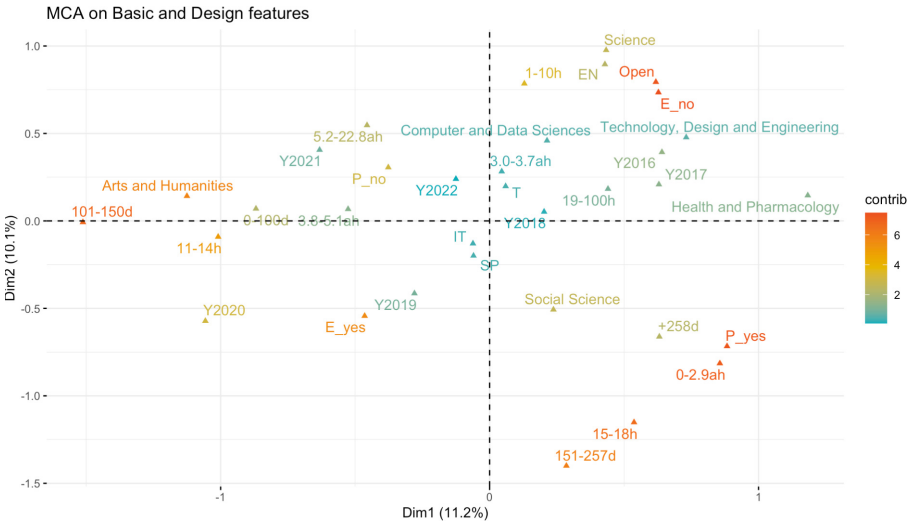

Fig. 1. Symmetric biplot of variables related to Basic and Design features of EduOpen MOOCs. Gradients show the contribution of each variable to the two dimensions (`ctr`).

In the second biplot (Fig. 2):

- Hard Science categories are distributed in the left sectors and more distant from each other than in the previous figure;
- the early Starting years (2016/2017/2018) are on the left side of the figure;
- the number of Learners enrolled in the course increases going from right to left;
- the number of Completion increases from the bottom to the top.

Dimension 1 (*inertia* = 11.5%) can represent the different kinds of DIFFUSION of MOOCs seen in the two half-planes of the biplot. On the left side, we find categories of Starting years between 2016 and 2018, with a Duration of a long period, no other editions, a high number of Participants, and Hard Sciences. On the right half-plane, we find categories related to more recent years (Starting year), Course categories in Social Sciences and Arts and humanities, more editions, and a lower number of Participants.

Dimension 2 (*inertia* = 7.8%) represents the level of PARTICIPATION in the courses. Going from the bottom to the top, the Completion rate, and the number of Posts per Participant increase.

According to `ctr` values, the variables involved are: Duration, Learners, and Editions.

The sum of the inertia of the two dimensions is 19.3%. This means that these two dimensions explain the variance of the same percentage of the observations.

The analysis allows us to identify at least two features (dimensions) of EduOpen MOOCs that characterize the design and delivery processes. The design dimensions that classify the courses relate to the kinds of organization of activities and the accessibility of contents. In course delivery, the dimensions to analyse MOOCs are the different diffusion and the level of participation in the activities.

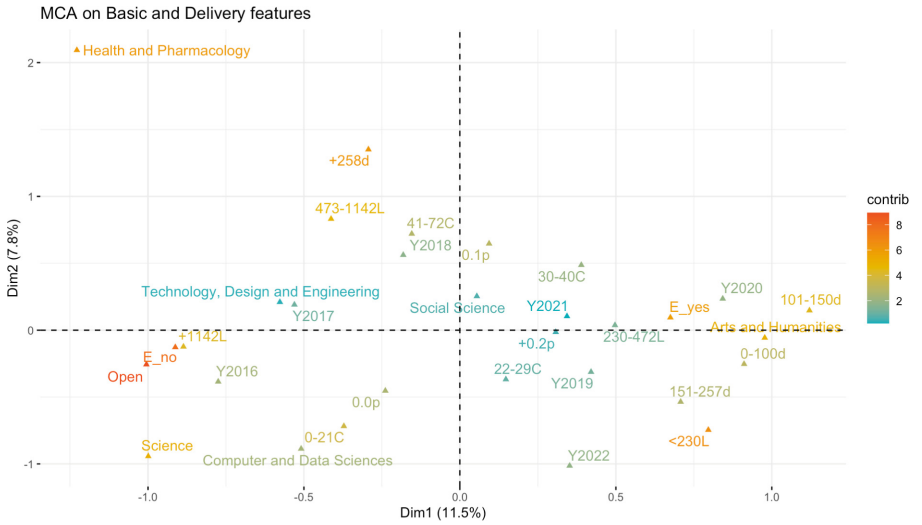

Fig. 2. Symmetric biplot of variables related to Basic and Delivery features of EduOpen MOOCs. Gradients show the contribution of each variable to the two dimensions (`ctr`).

In both cases, variable related to time in its various meanings Duration, Training hours, and Activities/Hour influence the two dimensions. We are used to dealing with synchronous, asynchronous, and self-regulation concepts, especially in blended courses [15]. MOOCs force us to talk about time factors as training hours and course scheduling that can change how students engage with contents [16, 17].

We added another more general observation in our discussion. Figure 3 shows the biplots for the categories of the variable named Course categories. In the graphs, the points represent the statistical units, and the colours in the legend define the categories. Our attention goes to the distance among categories. In the design biplot, courses related to scientific and technical disciplines are very close. Social sciences courses are in a different sector, and Arts and Humanities MOOCs are very distant from all the others on the left side. The same categories treated with the categories of the delivery variables take over another position. Health and Pharmacology courses are apart from all the others, and Science and Data Science ones remain near. Technology courses are nearer to Social science ones, and Arts and Humanities courses are now distant from the others but not as in the design biplot.

Similar effects can also be obtained by plotting the other common variables in the two analyses.

These two graphs seem to show an idea familiar to every teacher that proposes courses with the same structure or twice the same course: courses with similar design features could have very different levels of participation or diffusion when delivered.

As to Q2, MCA results show differences in MOOCs design and delivery over time in EduOpen. In Figs. 1 and 2, 2016-2017-2018 appeared in a different half-planes than the following four-year period. The splitting coincided with the period when the platform update was implemented. There has therefore been a variation in how courses

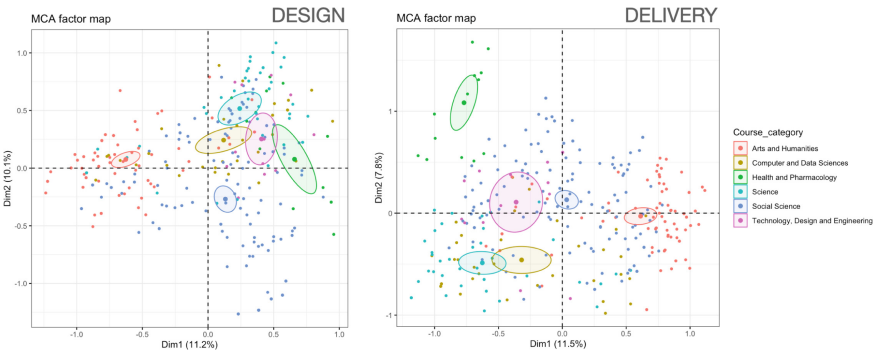

Fig. 3. Biplots for the categories of variable Course categories with Design (left) and Delivery (right) features. Points represent the statistical units and the colours in the legend define the categories.

were produced and delivered that could be related to the changes in the same platform, considering the influence of technical features on courses development. However, this break must be further investigated because it might also derive from more general issues related to the organisation of the network.

Besides, Fig. 4 shows two histograms: on the left, we can see the distribution of new learners registered from 2016 till June 2022; on the right, the distribution of courses by Starting year. The MOOCs in dark grey have only one edition; we find in light grey courses that were the following editions of a course.

In the first graph, 2020 represents an exceptional year for the number of registered users (about 30,000). We can hypothesize that EduOpen was used during the lockdown and Covid19 pandemic to satisfy training needs that universities and other entities had to meet or probably that training on EduOpen had been chosen as an activity to fill the empty moments due to the obligation of staying at home.

In the second graph, we can see that 2018 and 2019 were the years when more MOOCs were produced and published by the network institutions. However, in the same years, we observe an increase in courses with more than one edition. The distance between courses in light and dark grey was higher in 2020 and 2021, years of the pandemic, during which institutions were committed to the great challenge of online education and fewer (human and material) resources could be assigned to open education activities.

In reply to the second research question, we can say that platform updates in November 2018 didn't affect learners' participation but were part of an intense period of development of the platforms and seems to distinguish between two periods in the design and delivery of courses from the previous figures (Figs. 1 and 2). The effects of Covid19 pandemic/lockdown carried many new learners and a decrease in the universities' commitment to open education. The growth of learners in 2020 is consistent with the international scenario, which defines this year as the second year of MOOCs after 2012 [18]. At the same time, the opposite decrease in MOOCs production is a manifestation of the regional dimensions of EduOpen that didn't allow greater investment in production.

The number of MOOCs produced until June 2022 lets us think that the relatively stable health situation can bring a new engagement in MOOCs production.

4 Conclusion

Our analysis started from the consideration that studies on regional MOOC platforms as L@S environments can add information on the design and delivery of courses with respect to the more numerous research conducted at the international level. Using MCA, we analysed the courses published on EduOpen from 2016 to June 2022 to identify a latent structure in the design and delivery features of the courses that allows us to describe the actual status of the platform and identify interventions to better the courses.

Replying to Q1:

- we found two dimensions that define the structure that underlies the design process - COURSE DENSITY and CONTENT ATTAINABILITY, and the delivery process - DIFFUSION and PARTICIPATION. A limitation in the MCA is the values of *inertia* around 20%, which suggests that more dimensions have to be analysed to obtain more complete results;
- courses designed according to similar characteristics may have different delivery;
- variables related to time (Duration, Training hours, and Activities/Hour) influence the two dimensions identified in the design and delivery process.

Replying to Q2:

- in MCA results, the distinction on the half-planes among the years in Figs. 1 and 2 lets us think that course features are related to specific periods. The year of transition coincides with the technical updating of the platform. More investigation into this is necessary.
- Covid19 lockdown affected the number of new learners registered and accentuated a process already begun in 2019 of reduced commitment to the production of new courses by network institutions.

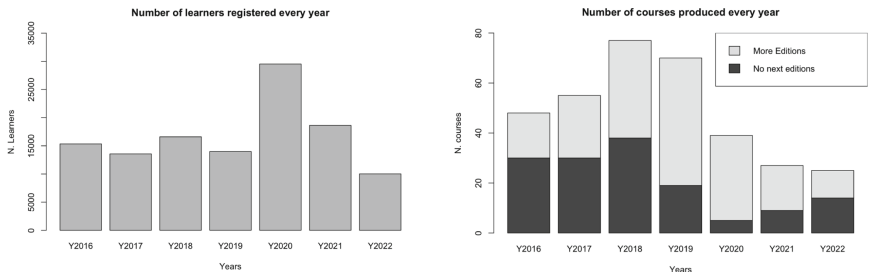

Fig. 4. Number of learners registered in the platform every year (left). Number of courses published every year; the light grey bars represent the courses that have one or more editions (right). Please, note that 2022 courses are considered till June.

Our study is set in the international scenario of MOOCs research where the production of papers is articulated around some thematic groups as institutional approach, pedagogical approach, evaluation, analytics, participation, and educational resources [19].

A previous study [20] analysed the design characteristics of more than 300 MOOCs in Portuguese using multiple correspondence analysis. Our research confirmed the use of this method for design analysis by applying it to MOOCs of a single platform. In our study, we were also concerned with analysing delivery issues along with design.

Some focal points that members of EduOpen (and not only) can consider to improve strategies in the design and delivery of the courses are related to:

- the awareness of the influence that the organization of activities and content in a MOOC can assume. In the first MCA on design, the contents are at the centre of the discussion for number, duration, amount of knowledge, and coherence with the whole course to which they refer. Future research can regard the relation between the features of contents and activities and the behaviours of learners in the courses;
- the attention to time-related issues that arose in the analyses with regard to the scheduling and duration of activities and entire MOOCs, course availability and estimated number of training hours, and changes over time;
- the reflection on levels of interaction of the courses that in this analysis confirm our previous results [21]. In fact, in the biplot in Fig. 1, the distance between the categories of course modes, tutored and self-paced, appears to be unremarkable. Participation in forum discussions, as noted in Table 1, is very low;
- the deployment of qualified professional figures for design and tutoring that are indispensable for enhancing and analysing the design and delivery mechanisms of the courses described here.

We can imagine a future for EduOpen to work in guidance, disciplinary and soft skills, recognition of credits, and mobility among universities. It has to start with attention to course design and scheduling, participatory mechanisms by learners, and the involvement of professional figures to design courses and support students.

Comparing these results with those from other (regional or global) platforms could be useful to generalize them or to characterize EduOpen with specific features associated with participants' cultural factors and design methods.

Authors' Contribution. The contribution represents the result of a joint work of the authors who collaborated in all the phases of the research work. According to CRedit system, Annamaria De Santis: Project Administration, Formal Analysis, Investigation, Writing – original draft; Katia Sannicandro and Claudia Bellini: Conceptualization; Tommaso Minerva: Supervision.

References

1. Perifanou, M., Economides, A.A.: The landscape of MOOC platforms worldwide. Int. Rev. Res. Open Distrib. Learn. **23**(3), 104–133 (2022)
2. Roll, I., Russell, D.M., Gašević, D.: Learning at scale. Int. J. Artif. Intell. Educ. **28**(4), 471–477 (2018)
3. Reich, J.: Learning analytics and learning at scale. In: Lang, C., Siemens, G., Wise, A.F., Gašević, D., Merceron, A. (eds.) Handbook of Learning Analytics, 2nd edn. SOLAR, Vancouver, Canada (2022)
4. van der Sluis, F., Van der Zee, T., Ginn, J.: Learning about learning at scale: methodological challenges and recommendations. In: Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale, pp. 131–140 (2017)
5. Ruipérez-Valiente, J.A., Halawa, S., Slama, R., Reich, J.: Using multi-platform learning analytics to compare regional and global MOOC learning in the Arab world. Comput. Educ. **146**, 103776 (2020)
6. Ruipérez-Valiente, J.A., et al.: Macro MOOC learning analytics: exploring trends across global and regional providers. In: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge, pp. 518–523 (2020)
7. Ruipérez-Valiente, J.A., et al.: Large scale analytics of global and regional MOOC providers: differences in learners' demographics, preferences, and perceptions. Comput. Educ. **180**, 104426 (2022)
8. Goglio, V., Nascimbeni, F.: MOOCS in Italy: an open and fragmented landscape. Ital. J. Educ. Technol. **30**, 82–96 (2021)
9. Shum, S.B.: Learning Analytics. Policy Brief, UNESCO Institute for Information Technologies in Education (2012)
10. Drachsler, H., Kalz, M.: The MOOC and learning analytics innovation cycle (MOLAC): a reflective summary of ongoing research and its challenges. J. Comput. Assist. Learn. **32**(3), 281–290 (2016)
11. Zhu, M., Sari, A.R., Lee, M.M.: Trends and issues in MOOC learning analytics empirical research: a systematic literature review (2011–2021). Educ. Inf. Technol. **27**(7), 10135–10160 (2022). https://doi.org/10.1007/s10639-022-11031-6
12. Bartholomew, D.J., Steele, F., Moustaki, I., Galbraith, J.I.: Analysis of Multivariate Social Science Data, 2nd edn. Taylor & Francis Group, CRC Press, Boca Raton (FL) (2008)
13. Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L.: Multivariate Data Analysis, 7th edn. Pearson, Edinburgh Gate, Harlow, Essex, GB (2014)
14. Kassambara, A.: Practical Guide to Principal Component Methods in R. STHDA Statistical Tools for High-Throughput Data Analysis (2017)
15. Norberg, A., Dziuban, C.D., Moskal, P.D.: A time-based blended learning model. On the Horizon **19**(3), 207–216 (2011)
16. Alemán de la Garza, L.Y., Sancho-Vinuesa, T., Gómez Zermeño, M.G.: Indicators of pedagogical quality for the design of a Massive Open Online Course for teacher training. International Journal of Educational Technology in Higher Education **12**(1), 104–118 (2015)
17. De Santis, A., Fazlagic, B., Sannicandro, K., Folloni, V., Tedeschi, C., Minerva, T.: Dalle linee guida di progettazione alla checklist di validazione: i Mooc di EduOpen. In: Rui, M. (ed.) Design the future, Extended abstracts della multiconferenza EMEMITALIA2016, Modena, 7–9 settembre 2016, pp. 264–275. Genova University Press, Genova (Italy) (2017)
18. Shah, D.: The Second Year of The MOOC: A Review of MOOC Stats and Trends in 2020 (2020). https://www.classcentral.com/report/the-second-year-of-the-mooc/. last accessed 2022/10/22

19. Despujol, I., Castañeda, L., Marín, V.I., Turró, C.: What do we want to know about MOOCs? Results from a machine learning approach to a systematic literature mapping review. Int. J. Educ. Technol. High. Educ. **19**(1), 1–22 (2022)
20. Lemos-de-Carvalho-Junior, G., Raposo-Rivas, M., Cebrian-de-la-Serna, M., Sarmiento-Campos, J.A.: Analysis of the pedagogical perspective of the MOOCs available in Portuguese. Revista Española de Pedagogía **75**(266), 101–119 (2018)
21. Sannicandro, K., De Santis, A., Bellini, C., Minerva, T.: Analysis of completion and dropout rates in EduOpen MOOCs. Ital. J. Educ. Res., 27–42 (2019)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.