Collecting data for feeding the online dimension of university rankings: a feasibility test

Raccolta dati per valutare la dimensione online dei sistemi di ranking delle università: un test di fattibilità

Albert Sangrà Morer*, Montse Guitert, Nati Cabrera-Lanzo, Maria Taulats, Lidia Toda and Anna Carrillo

Open University of Catalonia (UOC), Barcelona, Spain,
asangra@uoc.edu*, mguitert@uoc.edu, ncabrera@uoc.edu, mtaulats@uoc.edu, ltoda@uoc.edu,
acarrilloar@uoc.edu

*corresponding author

ABSTRACT Online universities have been always concerned about the quality of online education. However, the current rankings systems do not consider their singularities, and thus current rating criteria and indicators result in downgrading of their position in the rankings. CODUR (Creating an Online Dimension for University Rankings) is a European project that developed a set of criteria and indicators for evaluating the online dimension of university rankings. Criteria and indicators were created from a participatory research approach. A Toolbox to support data collection (and management) from the online universities to be ranked was also developed. This paper presents the testing process of this Toolbox, carried out with four international universities, located in different continents. These institutions have tested the feasibility of gathering the expected data to feed the indicators. The Toolbox test has been shown useful to forecast the availability of the data to feed the online dimension indicators.

KEYWORDS Teaching; Quality; Online Universities; University Rankings; Evaluation.

SOMMARIO Le Università online hanno sempre prestato attenzione alla qualità dell'insegnamento online. Tuttavia, i sistemi di ranking esistenti non tengono conto delle loro peculiarità, e pertanto gli attuali criteri e indicatori di valutazione penalizzano la loro posizione nei rankings stessi. CODUR (Creating an Online Dimension for University Rankings) è un progetto Europeo che ha sviluppato un insieme di criteri e indicatori per la valutazione della dimensione online dei ranking delle università. I criteri e gli indicatori sono stati individuati tramite un approccio partecipato. Il progetto ha inoltre sviluppato un toolbox per facilitare la raccolta (e archiviazione) dei dati da parte delle università online da valutare. Questo articolo presenta la valutazione di questo toolbox, condotta con quattro Università internazionali situate in altrettanti continenti. Queste istituzioni hanno valutato la fattibilità della raccolta dei dati richiesti per valutare gli indicatori. La valutazione del toolbox si è rivelata utile per predire la disponibilità dei dati necessari alla valutazione degli indicatori della dimensione online.

1. INTRODUCTION

Education and many other sectors are challenged by important societal and technological changes. Universities belong to the education system, so they are challenged as well. The transition from an Industrial Society to the Information and Knowledge Society has highlighted two important drivers: one is technology, and the other one is the Knowledge-Based Economy. Both of them exert pressures, changes and opportunities for education systems globally (Bates & Sangrà, 2011).

Reactions are quite diverse, but in general universities have been introducing technology to different levels, and shifting their teaching and learning models for several reasons: economic, social, or pedagogical.

According to Bates and Sangrà (2011) there are five main reasons universities are embarked on this transition:

- Enhancing the quality of teaching and learning
- Accommodating to the learning style of Millennials
- Increasing access to learning opportunities and flexibility for students
- Developing skills and competencies needed in the 21st Century
- Improving the cost-effectiveness of the system.

In other terms, by using technology as a driver for facilitating change, universities try to face the iron triangle (Daniel, 1999), generating wider access (Cannell, Macintyre, & Hewitt, 2015), reducing the costs (Deming, Goldin, Katz, & Yuchtman, 2015) and maintaining or enhancing the quality of their programming (McKnight et al., 2016).

There are several ways to integrate technology at universities but the emergence of the Learning Management Systems (LMS), in the 90's allowed many universities to start looking at online education models which were often considered as future scenarios. Today, online and blended teaching and learning models are no longer a promise, but a reality. Reports from the Sloan Consortium and the Babson College show that online education has been continuously growing during the last years (Seaman, Allen, & Seaman, 2018) in the US, as well as in other countries like Canada (Donovan et al., 2018) and Australia (Norton, Cherastidtham, & Mackey, 2018).

Although the increase has been very important, scepticism from a number of teachers, including even those who have experienced teaching online, persists (Straumsheim, 2014). It might be because most of them have been teaching face-to-face for years, and they feel much more comfortable and safe with that modality, and they may also show a weak conviction and meagre support for on online education.

This is one of the reasons, although it is not the only one, why quality of online education has been, and still is put under suspicion.

2. STATE OF THE ART

2.1. *Quality of online education*

The concept of quality has evolved from approaches that sought to develop control mechanisms of the final product (quality control), through its conception as a process which correctness has to be assured (quality assurance), to understand quality from a systemic point of view, which has to be properly managed (quality management). Currently, some voices are claiming to move towards the concept of quality enhancement

(Elassy, 2015), which considers quality has been reached and what it is going to be pursuit is its improvement. In parallel to these quality assurance approaches that are more focused on the evaluation, there has been a push for approaches seeking for continuous improvement as a process of internal understanding and of mirroring to other institutions and experiences or benchmarking. Benchmarking practices are usually intended to promote a commitment to quality rather than its control. Evaluation culture is evolving, and institutions are looking for assessment systems that could be better aligned with its mission and its special features.

In the field of online education, benchmarking initiatives had a great boom in the first decade of this century. A number of projects and initiatives produced benchmarking tools, like the Sloan Foundation (Five Pillars)[1], Quality Matters[2], EFQUEL (Unique or Open ECBCheck)[3], SEEQUEL[4], BENVIC[5] or E-xcellence (EADTU)[6], amidst others, which are still being used by most of the online or distance education universities.

2.2. *Rankings as a tool for market competition*

Because of some of the reasons stated at the beginning, the higher education arena is also becoming an even more competitive market (Pucciarelli & Kaplan, 2016). There is a growing willingness of governments to promote first-class universities that can compete globally (Song, 2018) and an accompanying request for increased accountability.

It is in this context that rankings become an important tool. They attempt to combine the identification of the achievement of a certain level of quality with the aim of establishing the valid parameters of competition between universities (Bilanow, 2010). It is no longer enough to be good and to do things well. One should be the best and make things much better than others.

Rankings emerge strongly in higher education have emerged as important tools for students, governments and industry, especially since the advent in 2003 of the Academic Ranking of World Universities[7], known as the Shanghai ranking. The Centre for World-Class Universities of Shanghai Jiao Tong University, supports the maintenance of this ranking. Overcrowding, marketization and globalization of higher education have made rankings popular tools and widely promoted by media organizations and (high ranking) institutions themselves.

In such a competitive environment, quality assurance procedures, benchmarking, accountability systems and rankings currently coexist as complementary mechanisms to support the prestige of universities (Berbegal-Mirabent & Ribeiro-Soriano, 2015). Rankings are being used as a means of becoming more visible, reputable, and "marketable". Reputation and research are the elements that currently most contribute to positioning in most of the rankings, but earning experience and student satisfaction are also considered (Locke, 2011).

There also are critiques of rankings. On one hand, creation of rankings is a very difficult task, since each university is different, has its particular mission and its focus, different resources, and is located in a dif-

---

[1] https://sloan.org/

[2] https://www.qualitymatters.org/

[3] European Foundation for Quality in e-Learning, https://www.eurashe.eu/about/partners/efquel/

[4] Sustainable Environment for the Evaluation of Quality in ELearning,
http://menon.org.gr/mind2innovate.org/wp-content/uploads/2012/11/SEEQUEL-eLearners-user-guide1.pdf

[5] Benchmarking of Virtual Campuses, http://www.benvic.odl.org/

[6] European Association of Distance Teaching Universities, https://eadtu.eu/

[7] http://www.shanghairanking.com/

ferent educational system The report accompanying the creation of the U-Multirank[8] ranking system has identified some of the weaknesses of the current rankings: ill-defined target audiences; homogenization that ignores institutional diversity; narrowness of focus, which sometimes is just on research; volatile methodologies; or the fact that there are few winners and many losers (Van Vught & Ziegele, 2011).

Rankings are based on various theoretical conceptual frameworks and indicators and thus they are often not comparable. Related to that, the most important critique is that institutions tend to change their mission focus to suit the rankings, deeply perverting what they should guarantee. Difficulties on gathering data are also a concern (Hazelkorn, 2008). However, it is generally accepted that rankings are socially ingrained and likely to stay.

3. METHOD

3.1. *Context of the study*

This lack of a ranking dimension for online education discriminates against and in some cases harms those universities with extensive online programs. The importance of this omission is also argued by current European Commission documents such those from the High Level Group on the Modernisation of Higher Education (2013; 2014). Online education, beyond Massive Open Online Courses (MOOCs) or Open Educational Resources (OERs), needs to be more visible and evaluated to improve its reputation and efficacy of the current higher education system.

In this context, under the European Commission Erasmus+ Call Strategic Alliance Partnerships in Higher Education, the CODUR (Creating an Online Dimension for University Rankings) project operated between October 2016 and September 2018. The main partners were the Open University (UK), the Consiglio Nazionale delle Ricerche-Istituto per le Tecnologie Didattiche (Italy), and the Universitat Oberta de Catalunya (Spain). The project aimed to generate:

   a. A set of quality criteria and indicators for the measurement of world-wide online education dimension, and

   b. The guidelines for integrating these online education metrics in the current U-Multirank ranking, the European funded university ranking which intends to be a European contribution to the development of a multidimensional ranking.

The identification of criteria and indicators of quality in online education started with an analysis of major issues and developments in online education (Giardina, Guitert, & Sangrà, 2017) and a parallel study on the existing quality assurance tools for online education (Brasher, Holmes, & Whitelock, 2017). The U-Multirank ranking was chosen due its multidimensional approach, which fits well with the purposes of the project.

The project took a participatory approach to the definition of the criteria and indicators, i.e. the design phase was not something that happened within the project boundaries, but, on the contrary, it involved additional stakeholders and informants and the broader Higher Education (HE) community. This was done under the assumption that considering the perspectives of the relevant stakeholders (both individual people and institutions) is crucial for the adoption of the criteria and indicators. The description of the criteria and indicators put forward by the project is out of the scope of this paper. For more information, see Pozzi et al. (in press).

3.2. *Toolbox development*

Once the indicators were defined in the project, the next step was to design and develop a Toolbox.

---

[8] https://www.umultirank.org/

The resulting CODUR Toolbox is a tool intended to support data collection (and management) from the online universities to be ranked. Starting from the results obtained in the previous phases of the CODUR project, we proceeded to design the Toolbox, by defining its main functions and capabilities at a goal level (Alexander & Beus-Dukic, 2009). In particular, the requirements' definition of the CODUR Toolbox revealed that the data to be collected could be classified in two different dimensions:

descriptive data that can be harvested from existing data sources (typically, institutional sources internal to the Higher Education Institution);

performance metrics that will be used as measures for comparing Online Universities and, in the final analysis, generating rankings.

Data collection (and management) is the core function of the CODUR Toolbox. Following the CIPO-model (i.e. Context-Input-Process-Output) (Scheerens, 1990), data belong to a context (in the CODUR project, the specific environment of an Online University) and represent a process (and its influences) at three different levels: input, process, and output. is important to note that, for the purposes of this work, the outputs of the process are those that must be measured, in terms of performance, to generate rankings. In practice, procedural aspects have more to do with accreditation of an Online University, and therefore deal with quality (in terms of assessment and assurance) by providing information and judgments (not ranking).

In the CODUR project, we are aware of the need to keep accreditation (which evaluates the process) separated from ranking, where a synthetic score could be derived based on objective data related to the outcome. Therefore, the CODUR Toolbox begins from this theoretical assumption, and its operative tools are designed to feed ranking system.

Generally speaking, data collection tools can be both qualitative and/or quantitative. The CODUR Toolbox is based on surveys and rating scales as data collection tools, devised to elicit information about qualitative and quantitative attributes of Online Universities.

In reference to the CODUR Toolbox, the data source is any existing person, group, or organization within the Online Institution that may provide information on whether and to what extent the intended output (or outcome) occurred. In this sense, data source for the CODUR Toolbox are the following: students, teachers, institution (as a whole or at the departmental level) and external reviewers (i.e. a panel of peer reviewers). Data types in the CODUR Toolbox include: alphanumeric strings such as numbers, percentages, composite scales, dichotomous scales (Yes/No), and Likert scales.

For the CODUR project, we are mostly interested in measuring and comparing data collected in reference to the outcomes. The measurement of outcomes is defined in terms of performance measurement and is obtained by means of synthetic and objective scores (performance metrics) for comparing Online Universities and generating rankings.

Thus, operational variables referring to specific empirically measurable features of Online Universities on which evidence can be collected have been identified, defined, and validated in terms of weighted indicators through a previous Delphi Study (Pozzi et al., 2019).

Performance metrics, next have to be scaled so that, before applying the weightings, they have the same relative importance. This is necessary because some indicators (e.g., internal budget devoted to disciplinary research per Full Time Employed (FTE) academic staff) are measured on a different order of magnitude than others (e.g., percentage of courses that propose personalized paths to reach the learning objectives). The former category of indicators risks to outweighing the ranking, unless corrective scaling is applied before weighting indicators. One way to perform such a scaling is transforming each measure in the corresponding percentile rank in the distribution of evaluated Higher Education Institutions (HEIs). This kind of transformation has the single drawback that it makes measures sample-dependent, i.e. measures can only be interpreted relative to each other, and not as absolute measures. However, since the goal of the Toolbox

is ranking, rather than evaluation, it is possible to adopt this approach without losing goal-relevant information. Therefore, the collected data can be exploited to generate performance comparisons, or rankings, and to sort Online Universities into "ranking order".

The Toolbox looks like a spreadsheet, structurally organized in 8 worksheets (i.e. "tabs") which allow the user to easily switch from one criterion to another (see Fig. 1). Each tab corresponds to one of the eight criteria previously identified: Teaching & Learning, Student Support, Teacher Support, Reputation-Impact, Research, Organization, Sustainability, and Technological Infrastructure (Pozzi et al., 2019). For each criterion, the relevant indicators are reported, as shown in Fig. 1. For each indicator, its weight, data source and data type are specified.

Criterion	Observable indicator	Weight	Data source	Data type	Data input
Quality of teaching & learning	Student satisfaction of the overall learning experience	16,00%	Student survey	Likert scale	
	Student satisfaction regarding adequacy of the adopted pedagogical approaches to the learning objectives	16,00%	Student survey	Likert scale	
	Institutional support for learning design (in terms of tools, formats, etc.)	15,00%	Institution	Yes/No	
	Percentage of courses that propose personalized paths to reach the learning objectives (for example offering different materials/activities depending on culture, learning style, background, etc.)	11,00%	Institution	Percentage	
	Student satisfaction regarding learning materials	15,00%	Student survey	Likert scale	
	Percentage of courses/examinations that make use of diverse forms of assessment (quantitative and qualitative approaches, human-based and technology-based tools, etc.)	14,00%	Institution	Percentage	
	Student and teacher satisfaction regarding performance reports	11,00%	Student and teacher survey	Likert scale	

Figure 1. The CODUR Toolbox. "Quality of Teaching & Learning" tab.

3.3. *Aims of this study*

Given that we knew by personal experiences, that collecting some kind of data can be particularly challenging and time consuming for universities and the risk is that - if it is too difficult to gather the data - the resulting information will not be provided, we proceeded with a testing phase for the Toolbox, so to determine what the challenges are of gathering data to quantify the indicators proposed by the project.

Thus, this study aims two main purposes:

To investigate whether and to what extent the necessary data from the universities can be achievable.

To investigate whether and to what extent the institutions already have these data or if they can get them in a reasonable time.

3.4. *Toolbox testing*

An instrument for testing the Toolbox was developed by the project. The test involved four universities from North-America, Australia, Africa and Europe. Online questionnaires were used to this end. This activity was undertaken between October 2017 and September 2018. This allowed testing the Toolbox with some word-wide universities from different continents, so to limit cultural biases on data gathering interpretation. The main aim of the process of Toolbox testing is to check whether and to what extent the data envisaged by the CODUR set of criteria and indicators can actually be provided by online Higher Education Institutions. There are a number of issues related to data collection that are unresolved and may be challenging in data collection and management, in the context of Online Universities. These include:

Estimated time required for data collection. How long would it take to collect the data for a HEI? If the estimated time turns out to be too long, online Universities might fail to provide the data in a timely manner.

Scheduling. Does the data need to be collected at a certain time or period of the year? Online Universities have internal and external constraints as far as time scheduling is concerned. Some data can be available at certain time of the year (for example, the number of new enrolments in one year might be available only after the enrolment phase is officially closed, etc.) and this might impose limitations on data collection schedule.

Data availability. Has this data already been collected / is this data (already) available? Most universities collect data about their activities for internal formative assessment purposes. If these data can be made public, they can be made readily available by the HEI with little or no additional cost.

Publicity/Privacy Issue. Can this data be made public / open? Online Universities might have constraints in terms of data openness and this needs to be taken into consideration to estimate if a datum will be available or not.

In order to understand whether and to what extent these issues might become an actual barrier that prevents (some of the) data collection, the CODUR Toolbox was integrated with a number of additional columns containing specific questions, each one addressing one of the above mentioned issues, namely:

The estimated time required for data collection, in the form of the following question: "How long would it take to collect the data?"

The scheduling for data collection, in the form of the following question: "Does the data need to be collected at a certain time or period of the year?"

The data availability, in the form of the following question: "Has this data already been collected / is this data (already) available?"

The publicity of the data, in the form of the following question: "Can this data be made public / open?"

In this field respondents should provide "Data example" that are aimed in this testing phase not so much at measuring the actual situation, but rather at providing an example of the kinds of data we could possibly expect.

[1] Criterion	[2] Observable indicator	[3] Weight	[4] Data source	[5] Data type	[6] How long would it take to collect this data?	[7] Does this data need to be collected at a certain time of the year?	[8] Has this data already been collected / is this data already available?	[9] Can this data be made public/open?	[10] Data example
Quality of teaching & learning	Student satisfaction of the overall learning experience	16,00%	Student survey	Likert scale					
	Student satisfaction regarding adequacy of the adopted pedagogical approaches to the learning objectives	16,00%	Student survey	Likert scale					
	Institutional support for learning design (in terms of tools, formats, etc.)	15,00%	Institution	Yes/No					
	Percentage of courses that propose personalized paths to reach the learning objectives (for example offering different materials/activities depending on culture, learning style, background, etc.)	11,00%	Institution	Percentage					
	Student satisfaction regarding learning materials	15,00%	Student survey	Likert scale					
	Percentage of courses/examinations that make use of diverse forms of assessment (quantitative and qualitative approaches, human-based and technology-based tools, etc.)	14,00%	Institution	Percentage					
	Student and teacher satisfaction regarding performance reports	11,00%	Student and teacher survey	Likert scale					

Figure 2. The CODUR Toolbox. (Quality of Teaching & Learning tab) integrated with additional questions aimed to test it.

The additional questions (green columns in Figure 2) were derived from the survey logical tree elaborated to this purpose, as shown in Figure 3.
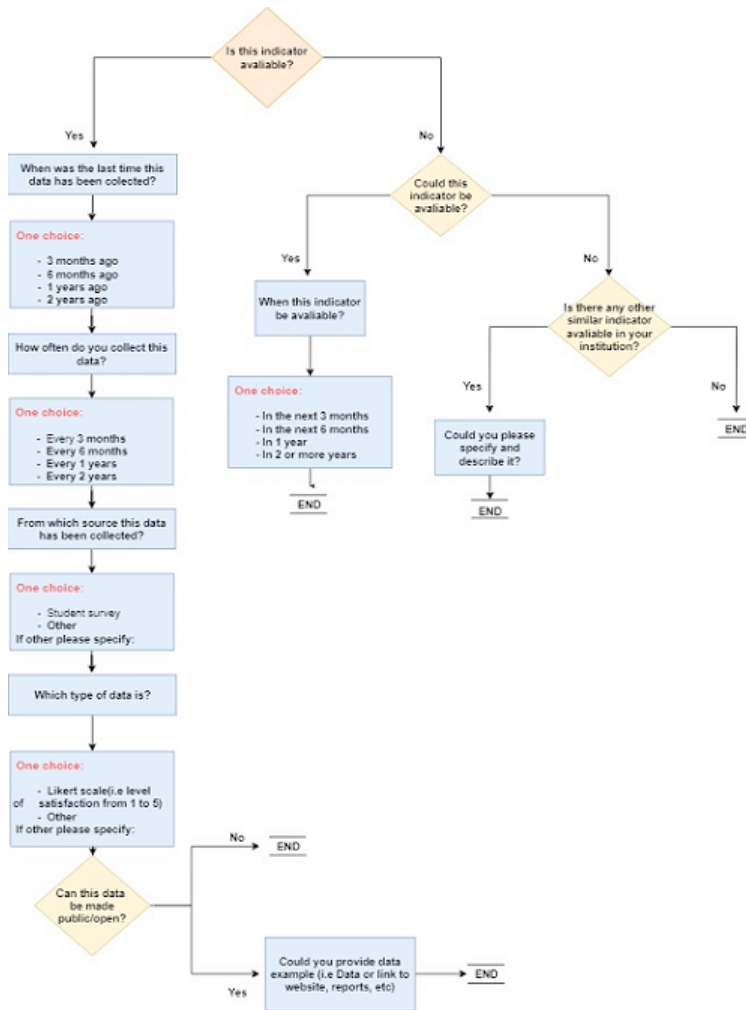
Figure 3. Survey Logical Tree.

Feedback collected in this process has been useful to validate the Toolbox. Furthermore, feedback collected regarding the last question helps to collect, with reference to the proposed model, useful data for:

Designing a basic data repository with filtering tools;

Implementing some preliminary data representation and visualization (such as league tables, charts or graphics).

To support the tester institutions in the test, a compilation guide has also been made available.

4. RESULTS

Our study focused primarily on the institution's availability of the data to feed the indicators. However, items as the estimated time required for data collection and its scheduling, the publicity of the data, and the data sources and information systems were also reviewed.

Regarding the estimated time for collecting data, it depends on the kind of indicator and its situation in the data fathering process at each institution. According to our results, it might take a range between few days

ad several months. Those institutions that do not have the data currently available consider they could have them in 6-12 months. The scheduling of the data collection ranges between 3-6 months (students' satisfaction surveys on the particular courses of programmes) and 1-2 years (satisfaction on the overall institution surveys).

As per the publicity of the data, most of the gathered data can be publicly open if they are aggregated. However, some cultural differences that have been identified show that some of these data, even aggregated, can be considered very sensitive and some restrictions applied.

Data sources and information systems are a key component for the purpose of gathering the required data. On one hand, data sources are usually found in the institutions. They have a lot of data, although sometimes they do not have the appropriate systems to manage them accordingly. Another important data source is also from the institutions: the students or faculty surveys. Experts' panels are usually another source for these data.

As said, to know the availability of data is the main aim of our study. It is important not only to know if the data is available, but if they could be made available in a reasonable period of time.

Table 1 shows this availability by universities, considering A) North-American university; B) Australian university; C) European university; and D) African University.

Criteria/Indicator	Institution/ Availability of data			
	A	B	C	D
QUALITY OF TEACHING AND LEARNING				
I1. Student satisfaction of the overall learning experience	Yes	Yes	Yes	No
I2. Student satisfaction regarding adequacy of the adopted pedagogical approach to the learning objectives	Yes	Yes	Yes	No
I3. Institutional support for the learning design	No	No	No	Yes
I4. Percentage of courses that propose personalized paths to reach the learning objectives	No	No	No	No
I5. Student satisfaction regarding learning materials	Yes	Yes	Yes	Yes
I6. Percentage of courses/examinations that make use of diverse forms of assessment	No	No	No	Yes
I7. Student and Teacher Satisfaction Regarding Performance Reports	No	No	Yes	No
QUALITY OF STUDENT SUPPORT				
I1. Student satisfaction regarding interactions with teachers/tutors	Yes	Yes	Yes	Yes
I2. Student satisfaction with technology support	Yes	No	Yes	Yes
QUALITY OF TEACHER SUPPORT				
I1. Teacher/Tutor satisfaction with technology support	Yes	Yes	Yes	No
I2. Number of hours of training devoted to teaching staff concerning online learning per year	No	No	No	Yes
I3. Teacher/Tutor satisfaction of training opportunities	Yes	Yes	Yes	No
I4. Teacher/tutor satisfaction with feedback on their courses derived from students' surveys	No	No	Yes	Yes
REPUTATION / IMPACT				
I1. Percentage of credits given in service-learning activities, in relation to the total number of credits	No	No	No	No
I2. Number of clicks/likes/shares/comments/followers/impressions on academic social network	Yes	No	No	No
I3. Percentage of post-graduated actively engages after graduation	Yes	Yes	Yes	No
I4. Percentage of former students employed in job sectors matching their degree	Yes	No	Yes	No
I5a. Joint/dual degree programmes	Yes	No	Yes	No
I5b. Inclusion of study periods abroad	No	No	No	No
I5c. Percentage of international (degree and exchange) students	Yes	Yes	Yes	No
I5d. Percentage of international academic staff	No	No	Yes	No
I6. Number of student internships	No	No	No	No
I7. Number of student mobility	No	Yes	Yes	No
I8. The proportion of external research revenues -apart from government or local authority core/recurrent grants- that come from regional sources	Yes	No	Yes	No
QUALITY OF RESEARCH				
I1. Internal budget devoted to research on online learning and teaching per full time equivalent (FTE) academic staff	No	No	No	Yes
I2. Percentage of full time equivalent (FTE) staff involved in research on online learning and teaching	No	No	No	No
I3. Yearly average n. of publications on online teaching & learning per full time equivalent (FTE) academic staff	No	Yes	No	Yes
I4. Yearly average number of publications with authors from other countries per full time equivalent (FTE) academic staff	No	No	Yes	No
I5. Internal budget devoted to disciplinary research per full time equivalent (FTE) academic staff	No	No	No	No
I6. External research income concerning disciplinary projects per full time equivalent (FTE) academic staff	No	No	No	No
I7. Yearly average n. of publications per full time equivalent (FTE) academic staff	No	Yes	Yes	Yes
QUALITY OF ORGANIZATION				
I1. Percentage of student complains or appeals solved/closed	Yes	No	Yes	No
I2. Number of full-time equivalents (FTEs) employed for non-instructional, non-technical support services (providing assistance for admission, financial issues, registration, enrolment, etc.) weighted by student satisfaction for the service	No	No	No	No
I3. Student satisfaction for room, laboratory and library facilities	Yes	Yes	Yes	Yes
I4. Student satisfaction for organization	Yes	No	Yes	Yes
SUSTAINABILITY OF THE INSTITUTION				
I1. Availability of an institutional strategic plan for online learning	Yes	Yes	Yes	Yes
I2. Percentage of curriculum changes resulting from an assessment of student learning (either formal or informal) within a fiscal year	No	No	Yes	Yes
I3. Percentage of total institutional expenditure dedicated to online programmes	Yes	No	Yes	No
QUALITY OF THE TECHNOLOGICAL INFRAESTRUCTURE				
I1. Student satisfaction with the overall learning platform	Yes	Yes	Yes	Yes
I2. Measure of compliance with the accessibility guidelines WCAG 2.0	No	No	No	No
I3. Measure of interoperability	Yes	Yes	No	No

Table 1. Current availability of data to feed the CODUR criteria and indicators.

Many of the indicators were already available and used by the participant universities or could be easily introduced in the near future (63.5% in the next two years). However, three out of the four universities accumulates the 80% of the available indicators, what means that one of the universities (D) is far away from the objective of having the indicators ready to be compared within the CODUR online dimension set of indicators.

Availability of the indicators is quite evenly distributed between the different criteria of the online dimension. While "Quality of student support" is the criterion that gets more availability of data (87,5%), "Quality of research" (25%) is the one with less data availability.

"Sustainability of the institution" also gets a good average availability (67%), as we can say similarly about "Quality of teaching support" and "Quality of organization" (56,2%). In a lower range of availability we find "Quality of the technological infrastructure" (50%), "Quality of teaching and learning" (46,4%) and "Reputation/Impact" (36,4%).

Any of the universities are able to currently feed eight of the indicators (I4, Quality of teaching and learning; I1, I5b, I6, Reputation/Impact; I2, I5, I6, Quality of research; and I2, Quality of organization).

5. DISCUSSION

The instrument for the testing of the CODUR project Toolbox was piloted with a subset of online and blended higher education institutions. The aim of this process was to document if online higher education institutions would be able to easily provide the information required by each of the CODUR indicators.

5.1. *Quality of teaching and learning*

The analysis of the indicators referring to the quality of teaching and learning reveals an incomplete but promising scenario. Universities periodically collected data on student satisfaction of the overall learning experience, student satisfaction regarding adequacy of the adopted pedagogical approach to the learning objectives, and student satisfaction regarding learning materials. Since these indicators are central for the quality of the institution, they are collected quarterly. These central indicators could be complemented with another two indicators identified by experts during the previous phases of the CODUR project: the institutional support for the learning design and student and teacher satisfaction regarding performance reports. The data for these indicators was either already available for universities or could be obtained in the next two or more years. Two more indicators were deemed less useful by institutions: the 'percentage of courses that propose personalized paths to reach the learning objectives' and the 'percentage of courses/ examinations that make use of diverse forms of assessment'. Data regarding the quality of teaching and learning was mostly available to the universities through student satisfaction surveys and were collected annually or biannually.

5.2. *Quality of student support*

Another relevant criterion for the assessment of online education is the quality of the student support offered by the online or blended higher education institution. The participant higher education institutions reported collecting data on student satisfaction regarding interactions with teachers/tutors. This data was collected quarterly, using student satisfaction surveys that include Likert scale questions on the topic. Half of these institutions were also taking into account student satisfaction with technology support. The other half, however, considered this indicator too difficult to manage. Data on this indicator was also collected through student satisfaction surveys.

5.3. *Quality of teacher support*

Quality of teacher support was assessed with a plurality of indicators. Most of the participating institutions gathered data on the teachers and tutors' satisfaction with technology support and with their satisfaction with training opportunities. In contrast, data on the 'number of hours of training devoted to teaching staff concerning online learning per year' and 'teacher/tutor satisfaction with feedback on their courses derived from students' surveys' were considered more difficult to manage. Some universities did gather data on these indicators or could make use of them in the future and other considered that they were just too hard to handle. The difficulty with the indicator 'teacher/tutor satisfaction with feedback on their courses derived from students' surveys' is that it requires two steps: surveying students and then surveying teachers about it. Universities proved to be more comfortable with data that could be obtained in one single step. As online learning becomes more pervasive and quality assessment improves, more universities may be able to consider this indicator. Data on the quality of teacher support is currently gathered annually or biannually with teacher surveys and/or staff climate surveys.

5.4. *Reputation/Impact*

A further criterion to be considered is the reputation and impact of online and blended institutions (Siemens, Gašević, & Dawson, 2015). The analysis of this criterion provides encouraging results. Most of the participant higher education institution already gathered data on this criterion and used five of the eight suggested indicators. Two of these indicators refer to the success of students after graduation. More concretely, almost all institutions counted the percentage of post-graduated actively engaged after graduation and on the percentage of former students employed in job sectors matching their degree. Another frequently used indicator refers to the availability of external research revenues, besides government funds, that come from regional sources. All institutions had data on this topic or were hoping to introduce it in the near future. Likewise, most institutions could report on the percentage of international (degree and exchange) students at their institution and on student mobility rate. Data on impact and reputation was mostly acquired quarterly or annually from institutional sources.

Some indicators on reputation and impact were more problematic. First, most of the participant universities considered the indicator 'Percentage of credits given in service-learning activities[9], in relation to the total number of credits' too hard to manage. This is due to the fact that there is not a common strategy for online institutions to offer service-learning activities and these can take place in a variety of ways, for instance through online simulation or through practices at the student workplace. Furthermore, some national agencies on higher education consider that institutions that demand students to complete service-learning activities offline cannot be considered completely online institutions regardless on how the rest of the instruction takes place.

Second, the indicator 'number of clicks/likes/shares/comments/followers/impressions on academic social networks', beyond the number of viewers of the institution's homepage, was also considered very difficult to manage. Advances in data mining and big data analysis may encourage universities to collect more information on this criterion (Macfadyen, Dawson, Pardo, & Gašević, 2014).

Three further problematic indicators refer to joint/dual degree programmes, the inclusion of study periods abroad and the percentage of international academic staff[10]. The participant institutions provided a wide

---

[9] "*Course-based, credit -bearing educational experience in which students (a) participate in an organized service activity that meets identified community needs and (b) reflect on the service activity in such a way as to gain further understanding of course content, a broader appreciation of the discipline, and an enhanced sense of civic responsibility.*" (Bringle & Hatcher, 1995, p. 112)

[10] Academic staff that participates, in any extent, in any sort of academic activities.

range of considerations regarding these indicators, from specific information on how they effectively gather this information to their conviction of the impossibility of their institutions reporting on the indicators. Mixed reports were also frequent.

This diversity of responses and the common lack of data may be due to the fact that it is not a priority for online and blended higher education institution to determine the location and provenance of their students and faculty as most of the academic activity takes place online (Altbach, 2016).

Finally, all institutions reported their interest in incorporating an indicator regarding the number of student internships.

5.5. *Quality of research*

The criterion that offers more room for the improvement of universities is that of the quality of research. Indeed, data in this area was very scarce and few universities indicated having data for some specific indicators. Nevertheless, the interest for incorporating indicators in this area was general. Universities suggested different moments for gathering these data, ranging from three months to more than two years. These differences respond solely to the characteristics of each university. The only indicator that was considered problematic out of the seven suggested by the CODUR team was 'the yearly average number of publications with authors from other countries per full time equivalent academic staff'. Again, this suggests that knowing the provenance or location of their staff may not be as relevant for online providers as it had been for their face-to-face counterparts. Although it could seem surprising that only one of these research indicators became critical, it has to be said that online universities are not used to measure research indicators. As they are traditionally considered as "distance teaching universities", indicators regarding research never were an issue. However, research is a key criterion for quality in online education, as it moves online universities to a similar and comparative level with conventional universities.

5.6. Quality of organization

The landscape regarding indicators on the quality of the organization is more varied. Most institutions reported having indicators on the percentage of student complains or appeals solved/closed and on student satisfaction for remote access to room, laboratory and library facilities. Data for these indicators was collected annually or biannually using institutional sources and student surveys. The participating universities provided contrasting responses regarding another indicator in this area: the student satisfaction of the organization'. Half of the organizations counted this data while the other half did not and could not even consider incorporating the indicator in the future. None of the institutions collected data regarding the number of full-time equivalents employed for non-instructional, non-technical support services (providing assistance for admission, financial issues, registration, enrolment, etc.) weighted by student. Indeed, even the definition of "full time equivalent" student is measured differently in different jurisdictions. However, most of the universities expressed their interest in incorporating this indicator in the future.

5.7. *Sustainability of the institution*

Data on the sustainability of the institution and on the quality of the technological infrastructure is easier to report. In fact, these criteria refer to the everyday functioning of the institution and most indicators were already available at the participating institutions.

Regarding the sustainability of the institution, most participant universities have some data on the availability of an institutional strategic plan for online learning and on the percentage of total institutional expenditure dedicated to online programmes. In contrast, universities presented very different responses regarding the indicator 'percentage of curriculum changes resulting from an assessment of student learning (either

formal or informal) within a fiscal year'. The variety of responses may be due to the fact that the gathering of data for this indicator requires two steps instead of one and thus, as mentioned above, discourages universities from collecting that data.

5.8. *Quality of the technological infrastructure*

Finally, all participant universities reported to be able to measure the quality of their technological infrastructure through their students' satisfaction with the overall learning platform. They also all expressed the interest in including an indicator regarding the compliance of their institutions with the Web Content Accessibility Guidelines 2.0 (Caldwell, Cooper, Guarino Reid, & Vanderheiden, 2008). Some universities had as well a measure of interoperability of their technological equipment with other platforms (Learning Tools Interoperability, Sharable Content Object Reference Model) gathered through subjective evaluations or automated performance checks.

6. CONCLUSIONS

Overall, the test of the CODUR Toolbox revealed that it is a useful instrument for online and blended higher education institutions to gather and manage data concerning their performance and outcomes. A slight difficulty was detected in determining the origin and current location of professors, researchers, and students since online everyday teaching and learning allows for an internationalization of the faculty and student body. Thus, practice reveals that perhaps this element may not be as essential to evaluate the quality of online and blended higher education institutions as first considered by experts.

The Toolbox was shown as useful to identify the availability forecast of the data to feed online dimension indicators. Although our results indicate that most of the indicators can be gathered in some of the universities, the results show that data do not fully match with the needs of the CODUR indicators yet. There are differences especially based in the period of the year data is collected or in the very nature of the indicator that can make it incomparable. However, most universities indicated that though this data is not currently available, they could gather most of them. The increasing use of data and the general concern regarding learning analytics should make this move easier.

The fact that universities usually collect data based on National requirements made difficult to compare some of the data retrieved. Other limitations of this study are related to the fact that surveys were answered by people in charge of statistics or quality units. This let the researchers to get those real data that were currently available. However, asking also to decision makers would have given a very interesting strategic approach to the future availability of these data.

Further research focused on identifying those specific indicators that could be equivalent to that described at the different CODUR criteria would be relevant to improve the current state of the online dimension and its more practical and useful application to most of the universities, both online or conventional. In addition, a more detailed analysis of each of the indicators could lead to determine more proper teaching and learning indicators, which would contribute to increase evidence about the quality of online education.

Online universities can benefit from the CODUR toolbox as a tool to identify those indicators to which they currently do not gather or assess and to develop strategies for their retrieval.

7. REFERENCES

Alexander, I. F., & Beus-Dukic, L. (2009). *Discovering requirements: how to specify products and services.* Chichester, UK: John Wiley & Sons.

Altbach, P. G. (2016). *Global perspectives on higher education.* Baltimore, MD, US: John Hopkins University Press.

Bates, T. A. W., & Sangrà, A. (2011). *Managing technology in higher education. Strategies for transforming teaching and learning.* San Francisco, CA, US: Jossey-Bass.

Berbegal-Mirabent, J., & Ribeiro-Soriano, E. (2015). Behind league tables and ranking systems: a critical perspective of how university quality is measured. *Journal of Service Theory and Practice, 25*(3), 242-266. doi: 10.1108/JSTP-04-2013-0059

Bilanow, J. (2010). *National university rankings and the evolution of global rankings*. Warsaw, PL: Perspektywy Foundation.

Brasher, A., Holmes, W., & Whitelock, D. (2017). Comparing online distance universities. In L. Gómez-Chova, A. López-Martínez, & I. Candel-Torres (Eds.), *Proceedings EDULEARN17,* IATED, (pp. 5411–5415). doi: 10.21125/edulearn.2017

Bringle, R. G., & Hatcher, J. A. (1995). A service-learning curriculum for faculty. *Michigan Journal of Community Service Learning*, *2*, 112-22. Retrieved from http://hdl.handle.net/1805/4591

Caldwell, B., Cooper, M., Guarino Reid, L., & Vanderheiden, G. (2008). *Web Content Accessibility Guidelines (WCAG) 2.0. W3C Recommendation 11 December 2008*. Retrieved from https://www.w3.org/TR/WCAG20/

Cannell, P., Macintyre, R., & Hewitt, L. (2015). Widening access and OER: developing new pràctic. *Widening Participation and Lifelong Learning, 17*(1), 64-72. doi: 10.5456/WPLL.17.1.64

Daniel, J. (1999). *Mega-universities and knowledge media*. London, UK: Kogan Page.

Deming, D. J., Goldin, C., Katz, L. F., & Yuchtman, N. (2015). Can online learning bend the Higher Education cost curve?. *American Economic Review*, *105*(5), 496-501. doi: 10.1257/aer.p20151024

Donovan, T., Bates, T., Seaman, J., Mayer, D., Martel, E., Paul, R., ... Poulin, R. (2018). *Tracking online and distance education in Canadian Universities and colleges: 2018*. Toronto, ON, CA: Canadian Digital Learning Research Association. Retrieved from https://eduq.info/xmlui/handle/11515/37136

Elassy, N. (2015). The concepts of quality, quality assurance and quality enhancement. *Quality Assurance in Education, 23*(3), 250-261. doi: 10.1108/QAE-11-2012-0046 https://doi.org/10.1108/QAE-11-2012 0046

Giardina, F., Guitert, M., & Sangrà, A. (2017). *The State of the Art of Online Education. Report*. CODUR Deliverable IO1.A1. Retrieved from: http://edulab.uoc.edu/wp-content/uploads/2018/06/CODUR-deliverable-IO1.A1_Stat-of-the-Art.pdf

Hazelkorn, E. (2008). Learning to live with league tables and ranking: the experience of institutional leaders. *Higher Education Policy, 21*(2), 193-215. doi: 10.1057/hep.2008.1

High Level Group on the Modernisation of Higher Education (2013). *Improving the quality of teaching and learning in Europe's higher education institutions*. Luxembourg, LU: Publications Office of the European Union.

High Level Group on the Modernisation of Higher Education (2014). *New models of teaching and*

learning in higher education. Luxembourg, LU: Publications Office of the European Union.

Locke, W. (2011). The institutionalization of rankings: Managing status anxiety in an increasingly marketized environment. In J. C. Shin, R. K. Toutkoushian, & U. Teichler (Eds.), *University rankings. Theoretical basis, methodology and impacts on global higher education* (pp. 201-228). Dordrecht, NL: Springer.

Macfadyen, L. P., Dawson, S., Pardo, A., & Gaševic, D. (2014). Embracing Big Data in complex educational systems: The learning analytics imperative and the policy challenge. *Research & Practice in Assessment, 9,* 17-28. Retrieved from https://www.learntechlib.org/p/157942/

McKnight, K., O'Malley, K., Ruzic, R., Horsley, M. K., Franey, J. J., & Bassett, K. (2016). Teaching in a digital age: How educators use technology to improve student learning. *Journal of Research on Technology in Education, 48*(3), 194-211. doi: 10.1080/15391523.2016.1175856

Norton, A., Cherastidtham, I., & Mackey, W. (2018). *Mapping Australian Higher Education 2018*. Carlton, AU: Grattan Institute. Retrieved from https://grattan.edu.au/wp-content/uploads/2018/09/907-Mapping-Australian-higher-education-2018.pdf

Pozzi, F., Manganello, F., Passarelli, M., Persico, D., Brasher, A., Holmes, W., ... Sangrà, A. (in press). Ranking meets distance education: defining relevant criteria and indicators for online universities. *The International Review of Research in Open and Distributed Learning*.

Pucciarelli, F., & Kaplan, A. (2016). Competition and strategy in higher education: Managing complexity and uncertainty. *Business Horizons, 59*, 311-320. doi: 10.1016/j.bushor.2016.01.003

Seaman, J.E.; Allen, I.E. & Seaman, J. (2018). *Grade increase: Tracking distance education in the United States*. Babson Park, MA, US: Babson Survey Research Group.

Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement, 1*(1), 61-80. doi: 10.1080/0924345900010106

Siemens, G., Gašević, D. & Dawson, S. (2015). *Preparing for the Digital University: A review of the history and current state of distance, blended, and online learning*. Arlington, AU: Link Research Lab. Retrieved March 2, 2019 from http://linkresearchlab.org/PreparingDigitalUniversity.pdf

Song, J. (2018). Creating world-class universities in China: strategies and impacts at a renowned research university. *Higher Education, 75*(4), 729-742. doi: 10.1007/s10734-017-0167-4

Straumsheim, C. (2014). Online ed skepticism and self-sufficiency: survey of faculty views on technology. *Inside Higher Education*. Retrieved from https://www.insidehighered.com/news/survey/online-ed-skepticism-and-self-sufficiencysurvey-faculty-views-technology

Van Vught, F., & Ziegele, F. (2011). *U-Multirank: Design and testing the feasibility of a multidimensional global university ranking: A final report*. Brussels, BE: Consortium for Higher Education and Research Performance Assessment. Retrieved from https://ris.utwente.nl/ws/portalfiles/portal/5145063/Vught11umulti.pdf