A Survey on Data-Driven Evaluation of Competencies and Capabilities Across Multimedia Environments

Sofia Strukova*, José A. Ruipérez-Valiente, Félix Gómez Mármol

Department of Information and Communication Engineering, University of Murcia, Murcia (Spain)

Received 22 July 2021 | Accepted 7 February 2022 | Published 3 October 2022

Abstract

The rapid evolution of technology directly impacts the skills and jobs needed in the next decade. Users can, intentionally or unintentionally, develop different skills by creating, interacting with, and consuming the content from online environments and portals where informal learning can emerge. These environments generate large amounts of data; therefore, big data can have a significant impact on education. Moreover, the educational landscape has been shifting from a focus on contents to a focus on competencies and capabilities that will prepare our society for an unknown future during the 21st century. Therefore, the main goal of this literature survey is to examine diverse technology-mediated environments that can generate rich data sets through the users' interaction and where data can be used to explicitly or implicitly perform a data-driven evaluation of different competencies and capabilities. We thoroughly and comprehensively surveyed the state of the art to identify and analyse digital environments, the data they are producing and the capabilities they can measure and/or develop. Our survey revealed four key multimedia environments that include sites for *content sharing & consumption*, *video games*, *online learning* and *social networks* that fulfilled our goal. Moreover, different methods were used to measure a large array of diverse capabilities such as *expertise*, *language proficiency* and *soft skills*. Our results prove the potential of the data from diverse digital environments to support the development of lifelong and lifewide 21st-century capabilities for the future society.

Keywords

Artificial Intelligence, Competencies, Computational Social Science, Data Mining, Multimedia Environments.

DOI: 10.9781/ijimai.2022.10.004

I. Introduction

Historically, educational research has been one of the most widely explored areas to improve teaching, learning and assessment [1]. Within this context, formal education has been understood as learning that happens within regular classrooms. However, a large body of research has explored the more informal learning constantly taking place across everyday activities that fall outside of a curriculum. Furthermore, the advent of the World Wide Web, followed by the spread of Internet usage, has changed how informal learning emerges in depth [2]. In this way, it also has led to a huge growth of online learning including not only Massive Open Online Courses (MOOCs) like Coursera[1] or edX[2] and language learning sites like Duolingo[3] or Babbel[4] , but also online portals where informal learning can emerge, such as websites that follow a Question-and-Answer format (Q&A), numerous online games, photo and video sharing platforms or social networking sites, amongst many others [3]. This kind of the platform can attract a diverse spectrum of users, especially those that are typically less interested in traditional learning approaches [4]. By creating, interacting with and consuming the content from these environments, users can, intentionally or unintentionally, develop different skills.

All the previously mentioned environments on the web, in turn, generate large amounts of data stemming from the interactions carried out by their users in such contexts. Consequently, it is important to highlight that big data can have a significant impact on education since it offers unprecedented opportunities to support learners and advance research in the learning sciences [5]. In addition, big data practices can help discover new and useful insights about the service of education providers, its students, competitors in private sectors and, most importantly, to gain its value for better educational outcomes [6]. According to Kizilcec et al. [7], research with heterogeneous samples of learners can provide a more inclusive science of learning that moves beyond tailoring to averages. This could help us to understand a better range of issues at the core of learning and technology research [8]. These topics are connected with the term *datafication*, which allows analyses of information across large data sets in more sophisticated ways [9]. According to Mayer-Schoenberger et al., datafication refers to the transformation of social actions into quantified data that permits real-time tracking and predictive analysis [10]. This is considered as a revolutionary research opportunity to investigate human behaviour [11] and an innovative way to inspect the behaviour of learners. Hence, within the context that we are exploring, the data generated by users in these environments hold the potential to infer valuable information about users' competencies and capabilities.

1 https://coursera.org/

2 https://edx.org/

3 https://duolingo.com/

4 http://babbel.com/

* Corresponding author.

E-mail address: strukovas@um.es

From an educational point of view, the landscape has been shifting from a focus on contents to a focus on competencies and capabilities [12], [13]. The rationale is that rapid technological evolution is having a direct impact on the skills and jobs needed in the next decade [14], [15]. Therefore, it is vital to raise a new generation that can self-regulate flexibly and rapidly acquire new skills and knowledge, as the world is changing in terms of economics, technologies, social and cultural life [16]. Accordingly, we need to focus on enhancing competencies and capabilities that will prepare our society for an unknown future during the 21st century. In this respect, interest in deploying modern techniques focused on both formal and informal learning has been rapidly growing in recent years. This has led to the fact that understanding learners and their contexts has become one of the most promising educational research topics during the past decade. For example, Redecker and colleagues [17] hypothesised that, by 2025, schools will have started integrating external learning resources and practical learning opportunities to address and implement students' individual needs and preferences. Therefore, one of the most exciting goals that researchers set in this field is to evaluate competencies and capabilities that the user potentially acquires by interacting with specific digital environments. This evaluation can be used to provide personalised feedback and to understand better how these skills develop through the interaction with these environments [18]. However, the challenge is how to perform such evaluations since there is no systematic way of doing it. Moreover, all previous studies have focused on developing and measuring competencies based on one specific platform while no one, to the best of our knowledge, studied the complete picture of the multiple existing online portals that can be used for this purpose. Hence, our survey aims to broaden the current knowledge on this issue by reviewing the research body that has already performed data-driven evaluations of competencies across different multimedia environments. As a result, this work could be the basis for developing a framework that can generalise well to different digital platforms and capabilities, and this can help to assess the capabilities of the users and to bring about a change in the educational system and training environments. Accordingly, we will try to answer the following overarching research question (RQ): Is there evidence of the existence of diverse technology-mediated environments whose data can be used to evaluate competencies and capabilities that the users can gain? For this purpose, we formulated the following specific RQs:

RQ1.What types of multimedia environments generate rich data sets that can be used for capabilities measurement?

RQ2.How are the data for the analysis accessed?

RQ3.What types of data are found across the environments?

RQ4.What methods or techniques are applied to infer competencies and capabilities?

RQ5.What competencies and capabilities are measured and/or developed across the environments?

RQ6.Are the findings across the publications validated?

RQ7.What are the main limitations or challenges faced by the authors of the studies?

The remainder of this paper is structured as follows. In Section II, we focus on the background of our study and related ones. Then, in Section III, we present the description of the RQs that drive our survey, followed by the detailed representation of the research methodology, including the description of the search process, inclusion and exclusion criteria, data collection and coding process. The research findings are outlined in Section IV, while in Section V we discuss our key findings, extend the article with a discussion beyond the results and present the implications of our research and the limitations of the selected approach. Finally, our conclusions are presented and future research directions suggested in Section VI.

II. Background

A. Competencies and Capabilities

Although both the terms 'competencies' and 'capabilities' are used to describe human abilities and are very closely related to each other [19], there are significant conceptual differences between them.

Generally speaking, the word competency can be defined as "something we already have or might be aiming to gain" [20]. However, the term's original definition may provide a deeper understanding of it. Initially, in 1982, a competency was defined by Boyatzis as an underlying characteristic of an individual that is causally related to effective or superior performance in a job [21]. Competencies are generally divided into functional competencies, used in daily activities, and integrative competencies, used to integrate and develop new components [22].

On the other hand, Lozano et al. [23] stated that competencies are externally demand-orientated as they are intended to provide the individual with the appropriate skills to solve external problems. In contrast, capabilities are not primarily externally demand-orientated, as they are guided by the exercise of individual freedom to choose and develop the desired lifestyle, and therefore the values individuals consider to be desirable and appropriate. Accordingly, the authors of [24] explained that an individual's set of competencies reflects their capability (i.e. what they can do) while a job competency may be a motive, trait, skill, aspect of one's self-image or social role or a body of knowledge. Consequently, capability approaches focus on developing the potential to achieve or acquire competencies, even if they are not present at a particular point in time, through certain personal qualities and attributes of individuals as well as ambition and effort [25]. At the same time, Nagarajan et al. [19] stated that capability integrates knowledge skills and personal qualities used effectively and appropriately in response to varied, familiar and unfamiliar circumstances. Finally, another important and related term is capacity. Morgan [26] stated that capacity is an emergent combination of attributes that enables a human system to create developmental value.

To sum up, competence is the quality or state of being functionally adequate or having sufficient knowledge, strength and skill, while capability is a feature, faculty or process that can be developed or improved [27]. Therefore, we can conclude that capabilities are made up of competencies that go beyond existing knowledge and experience. In this way, in our work, we refer to both, competencies and capabilities, when we use the term "capabilities". Conversely, by using the term "capacity", we refer to the overall ability of a person to achieve a goal or complete a task.

B. Related Work

We did not find any survey or literature review aiming to analyse data-driven evaluations of competencies and capabilities across several contexts at the same time. However, we did find studies that tackled one specific context or concept separately.

We found some surveys focused on detecting users online that have a high competency in specific skills, which is known in the literature as expert finding. Across these surveys, we found that one of the most established objectives within this context is detecting potential experts online. The authors of [28] aimed to review the existing literature to identify all the significant studies that addressed the task of expert finding in online communities and corporations. Accordingly to their goal, they summarised the existing graph-based and machine learning-based expert finding methods used. Similarly, Husain et al.

[29] conducted a systematic, state-of-the-art literature review of 96 articles on expert finding systems and expertise seeking. This study aimed to explore the domains that use the expertise retrieval systems, the expertise sources, the methods and the data sets used for expert finding systems. However, the most representative study related to expert finding discussed here is [30]. This comprehensive state-of-the-art survey reviewed 265 articles and proposed a framework that defines the descriptive attributes of Community Question Answering (CQA) approaches. The authors also introduced a classification of all approaches concerning problems they aimed to solve. From these studies, we can see that expert-finding is a common practice across those portals that could show users' competencies in particular topics.

Furthermore, there is a considerable amount of literature exploring the benefits of video games for developing and measuring competencies and capabilities. For example, Connolly et al. [31] examined 129 papers on computer and serious games. The game of every article was correlated with its genre, primary purpose as well as learning and behavioural outcomes. In their review, the authors found positive impacts on playing digital games on users with respect to learning, skill enhancement and engagement. The next review was conducted by Boyle et al. [32], who focused on 143 studies providing higher-quality evidence of the positive outcomes of games. While the previous work provided a useful framework for organising the research in this area, this article illustrated the increasing interest in the positive impacts and outcomes of games. Similarly, the authors in [33] investigated quality empirical studies associated with the application of games-based learning in primary education, mainly focusing on study design, the game genre, delivery platform, subject and curricular areas, as well as learning outcomes and impacts. We can see that in the field of games, the research focusing on exploring their benefits in terms of developing competencies and capabilities is fairly widespread. Moreover, we found related work performing a systematic review of publications about the potential of MOOCs in higher education aiming to investigate cognitive and behavioural learning outcomes, including obtained or mastered knowledge and intellectual skills [34].

The above review reveals that all previous studies explored how to develop and measure competencies based on one specific type of online environment. However, no single study, to the best of our knowledge, has provided a full picture of the multiple existing multimedia environments that can be used for this purpose. With this objective in mind, we surveyed the existing literature to study data-driven evaluations of competencies and capabilities across various digital environments.

III. Methodology

First of all, we stated the RQs and, accordingly, several steps were taken to select the publications for our survey, including (1) a literature search for identifying relevant articles, (2) the inclusion and exclusion criteria, (3) a full paper review and coding process, and (4) synthesis and analysis. The entire methodology process is represented in Fig. 1.

A. Search

Our literature search entailed an ill-defined area arising from many research communities. This led to several challenges regarding setting a clear search methodology to conduct the survey. The main challenge was that we could not directly retrieve all the literature on this topic by simply using a set of keywords and searches. Therefore, we were not able to apply an utterly systematic search as proposed in the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement [35]. On the other hand, while conducting our non-systematic literature review, we mainly relied on a knowledgeable selection of relevant current, high-quality articles [36]. Although we had to adapt the search process to the context of the target research

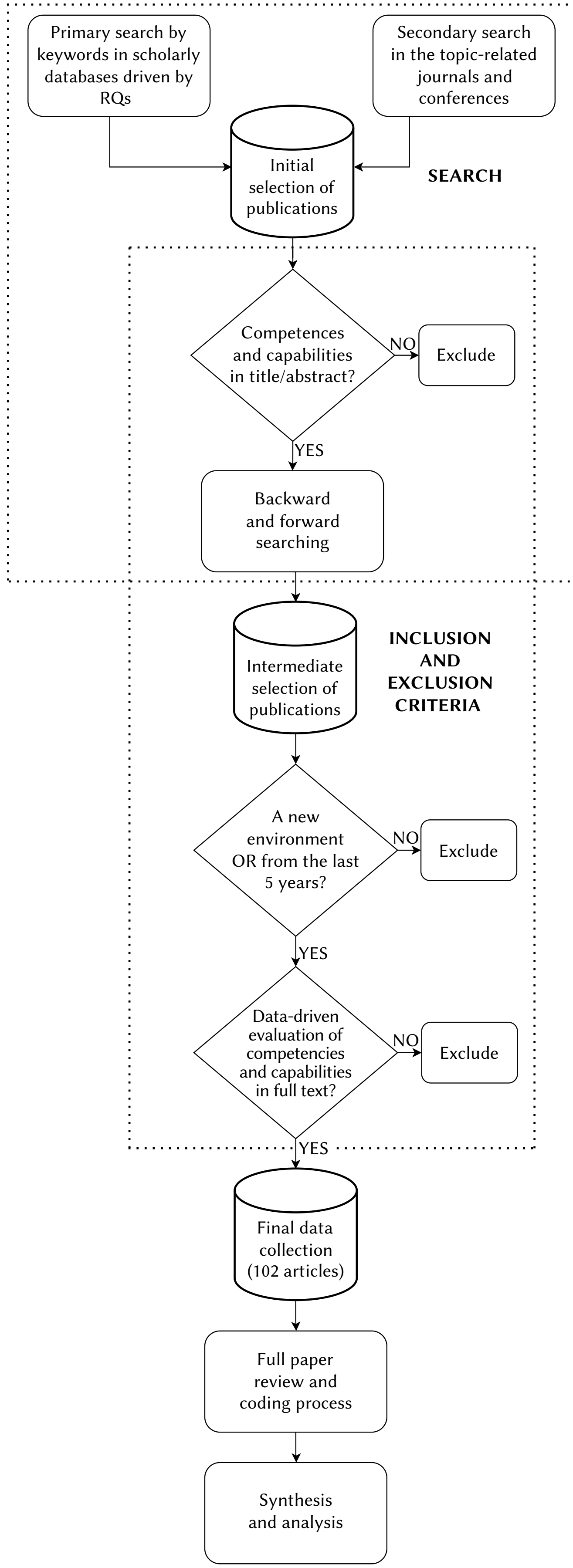


Fig. 1. Overview of the methodology used to conduct the survey.

area and, accordingly, change it, the rest of the methodology followed the PRISMA guidance and exemplars.

Initially, we identified keywords and formulated search strings according to our goal of exploring diverse multimedia environments in which data can be used to perform a data-driven evaluation of competencies and capabilities of different nature. We started the literature search with the following approach:

1. We looked for related papers using keyword search on various indexing platforms such as Scopus[5] and Google Scholar[6] . The examples of the keywords are as follows: "expert identification," "information retrieval," "game-based assessment," "forum," "question and answer portal," "online learning," "MOOC", "competence," "capability," "skill," "engagement," "soft skills," "language proficiency," "correct on/at (the) first attempt".
2. We explored the publications of the corresponding top conferences and journals related in the fields.

Firstly, we performed a fast initial paper screening to discard those that did not fit the survey by reading their title and abstract. More specifically, a paper was not included if there was no mention of the potential measurement and/or development of competencies and capabilities in its title. This probably meant that its primary focus was unrelated to our interest. In case the title, together with the abstract, did not provide sufficient information to justify its exclusion, the paper was not excluded. Next, we conducted the so-called 'snowballing technique' [37] comprising two methods: backward searching and forward searching. Backward searching included the review of the references of the publications fitting our survey while forward searching consisted in the examination of the papers that cited a known and relevant publication. We applied these methods to the publications that already proved to belong within our survey. Finally, the newly selected papers went through the same screening criteria as initially described by reading their title and abstract.

B. Inclusion and Exclusion Criteria

After performing the search and identifying the papers fitting our survey, we formulated several mandatory criteria for the final decision of whether to include the paper in the survey or not. Creating a valid set of inclusion and exclusion criteria required considerable trial and error pilot testing [38]. Accordingly, a paper was excluded if none of the following conditions was met (in other words, if at least one condition was met, the paper was included):

- The paper was published in the last five years.
- The paper represented a new environment or features that would deepen the results of the RQs.
- The paper was peer-reviewed.
- Reported outcomes in the paper included all the data needed for answering RQs outlined in Section I and were presented appropriately and consistently.
- The data from one paper did not overlap with the data from another paper.

[5] https://scopus.com/

[6] http://scholar.google.com/

We used, as a quality proxy, the type of publication and the venue, including full papers in conferences, book chapters and articles published in reputed journals. After that, we performed a more careful review of those papers originally selected. The abstracts were read again together with the methods and results sections. Through this process, we ensured that the selected publications measured concrete competencies and capabilities and performed their data-driven evaluation and thus correctly fitted the survey.

C. Data Collection

The final data collection included a total of 102 articles out of hundreds of initially selected publications. These articles are presented in ascending chronological order in Appendix B.

Most of the publications have a selection of keywords that define the research topic. We collected the keywords of all the papers selected for the survey, and we counted how many publications used the same keywords. The most frequent keywords are presented in Fig. 2.

The total sum of keywords is 415, while there are 329 different keywords. The average keyword was found 1.6 times, and its variance is 1.45. Based on the keywords (see Fig. 2) which define the publications' main topics, we did observe a high variability in the type of keywords used. Therefore, we see that the research topic covers many areas.

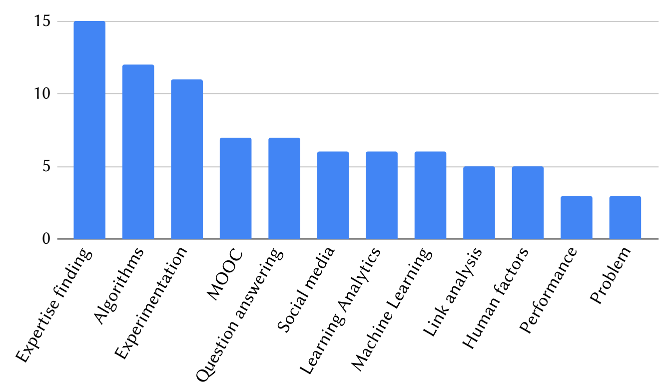


Fig. 2. Distribution of keywords across the articles selected for our survey.

D. Full Paper Review and Coding Process

In the coding stage, we identified variables that we consider as the most valuable to address the RQs stated in Section I. Based on the aim of our survey and the fact that we did not know in advance what data we could find, we followed an inductive coding scheme. This means that codes were created based on the qualitative data themselves. In Table I, we outline the variable coding scheme that we followed in our survey, indicating each code with its possible labels, which represent qualitative data. In the event of a paper deploying several experiments or using multiple methods [39]–[41], these were coded with several variables such as *"Network analysis, NLP"*. It is also important to mention that in the case of multiple methods, we coded only those used to infer users' competencies or capabilities rather than to report results. The full results of the coding process per paper are presented in Appendix B.

TABLE I. The Variable Coding Scheme

Environment (RQ1)	Data access (RQ2)	Data type (RQ3)	Methods (RQ4)	Capabilities (RQ5)	Validation (RQ6)
Content sharing & consumption	Open data set	Textual	Machine Learning	Expertise	Yes
Games	Public domain	Biometric	Statistics	Language proficiency	No
Online Learning	API	Clickstream	Network Analysis	Soft skills	
Social Network	Direct access	Audiovisual	Experiment/control	Performance	
			Natural Language Processing	Cognitive skills	

At the same time, from the coding process, three groups of publications emerged depending on their degree of relationship with the main goal of our survey:

1. The **strong relationship** group contains those papers that exactly matched the topic of the survey. This means that they proposed a method for identifying competencies or capabilities based on the data gathered by the user's interaction with a particular digital environment. We assigned 56 papers to this group.
2. The **weak relationship** group is similar to the first one; however, instead of measuring capabilities, the work belonging to this group explored other data-driven behavioural characteristics of users related to capabilities such as engagement or influence, to name some examples. Although this specific behaviour was not within the primary focus of our survey, these publications might be useful for analysing the applied methods and the selected environments. For example, several studies have explored a user's influence on others [42]–[46] or predicted dropout rate [47], [48]. We assigned 31 papers to this group.
3. The **high potential** group comprises those papers that had goals less related to our survey objectives but still described thought-provoking multimedia environments with large amounts of user data that could be used to perform a data-driven evaluation of capabilities. This group includes mainly those studies which did not provide evidence of using data generated through the user's interaction with the environment; instead, these studies measured other characteristics of users not related to their competencies or capabilities (identified in RQ5). We discuss these studies in more detail in Section B. We assigned 15 papers to this group.

IV. Results of the Synthesis and Analysis

In this section, we highlight our findings discussing each RQ in detail. We start with a description of all the types of environments within the scope of our survey. Next, we describe the data access, followed by the data type and the applied methods that emerged after analysing the articles. Finally, we discuss particular competencies and capabilities measured and/or developed across the selected environments. We also present the analysis of the validation of the results across the selected studies, followed by the main limitations mentioned by the authors of the publications.

A. What Types of Multimedia Environments Generate Rich Data Sets That can be Used for Capabilities Measurement? (RQ1)

Four groups of environments were identified during the review process described in the methodology section, namely, **content sharing & consumption** (41 articles), **video games** (13 articles), **online learning** (31 articles) and **social networks** (16 articles). They can all generate rich data sets through the users' interaction, which can be used to explicitly or implicitly perform a data-driven evaluation of competencies and capabilities. Next, we describe in more detail each one of them.

1. Content Sharing & Consumption

In this group of environments, we have all the platforms where users either intend to share content or consume it. Their categorisation is as follows:

- **Q&A portals and forums**. They include Quora[7] [49], Reddit[8] [50], StackOverflow[9] [51], [52], and Yahoo! Answers[10] [53], among others. These environments are quickly becoming rich sources of knowledge on many topics not well served by general web search engines [54]. They can be defined as online discussion sites where users can post messages stating what they are interested in or replying to others. On these platforms, users can ask questions, give answers and also provide their assessments about the quality of questions or answers through votes and choosing favourites [55].
- **Photo and video sharing online platforms**. Hosting services for video are called Online Video Platforms (OVPs); they allow users to upload, play, store and transfer video content online. These include, among others, YouTube[11] [56] and Vimeo[12] . On the other hand, users of Instagram[13] [57] can share both photos and videos, and Pinterest[14] allows to share only images so that users create, organise and share content by creating visual bookmarks called pins [58], so it is possible to characterise the volume and coherence of users' pinning activity in a given category [59].
- **GitHub**[15]. This platform [60] does not match any category above, but it still represents an appropriate and important portal corresponding to the stated goals. This service allows its users to host open-source projects and work collaboratively on them.

It is also worth mentioning that some papers discussed several *content sharing & consumption* platforms in their research scope. For example, the authors of [61] presented a method to create a detailed technology skill profile of a candidate. This was based on his code repository contributions through annotating user contributions on GitHub code repositories with technology tags found on StackOverflow.

2. Video Games

Even though the majority of *video games* are set out to entertain, nowadays there are many games created for other purposes, and their value as a learning tool has been widely accepted [62]. The explored *video games* that we found are categorised according to their main purpose into:

- **Entertainment games**. These are the commercial games that are played for entertainment purposes. For example, [63] described a racing game and explored the creation of its players' engagement profiles. The games of this type have an easily understandable user interface, their main goal is to entertain, and they are designed to provide an immersive experience for the player through gameplay, narrative and challenges [64]. Moreover, this kind of game can help develop various competencies and capabilities (e.g., multitasking and problem-solving skills) without users even noticing their involvement in the educational process.
- **Educational games**. This type of game is designed for educational purposes, and numerous teachers have already tried to teach foreign languages, history or programming with these games [65]. The use of educational games attempting to digitise and optimise the learning process has increased significantly [66]. This is due to the initial investigations suggesting that it can be rewarding at many levels, including academic achievement, concentration and classroom dynamics in general. Many experiments were conducted seeking to create *video games*, which could be beneficial for the learning process [67]. After the first studies succeeded, an interest in educational *video games* attracted the attention of many researchers. Nowadays, it is an up-and-coming field opening new ways of conducting educational processes [68].

[7] https://quora.com/

[8] https://reddit.com/

[9] https://stackoverflow.com/

[10] https://answers.yahoo.com/

[11] https://youtube.com/

[12] https://vimeo.com/

[13] https://instagram.com/

[14] https://pinterest.com/

[15] https://github.com/

According to the stated RQs, one example we were interested in is [69]. This article described the Virtual Age game, which is effective for learning about evolution, as its authors claimed. For this purpose, the scientific concept of evolution was implemented, concretised and gamified. The game's main idea is to create some extinct creatures, keep them alive and transmit them from one era to another. The results proved that Virtual Age is useful for learning about biological evolution.

- **Serious games**. These games are designed for purposes other than, or in addition to, pure entertainment. In other words, they can train specific skills and improve learning performance through real-world problem-solving [70]. The authors of [71] consider serious games as adaptive systems, as they continually adjust their responses to the learners' actions for preserving favourable conditions for playing and learning. The difference with educational games is that serious games are not directly connected with educational goals but can aim to change the users' behaviour, attitude, health or other features. Still, the educational purpose is not excluded.

  In this paper, we consider serious games as a subtype of educational ones. For example, a serious game described by Kang et al. [70] provides an authentic learning context for space science and astronomy. Users have to find a suitable home in the solar system to relocate a group of six distressed aliens because their home planets have been destroyed. The authors provided an analytical approach to understanding in depth students' sequential patterns of behaviour. At the same time, they showed that problem-solving strategies were different between low- and high-performing students.

3. Online Learning

Several types of online learning environments emerged after analysing the surveyed articles:

- **MOOCs (Massive Open Online Courses)**. They are defined by Daradoumis et al. [72] as one of the most versatile ways to offer access to quality education, especially for those residing in far or disadvantaged areas. Taken as a whole, MOOCs can provide not only traditional learning materials but also make students' forums or social media discussions available. There are many studies examining data from various MOOCs although, in our work, we mostly focus on the most well-known MOOC providers such as Coursera [73]–[76] and edX [77]. These MOOCs follow a similar way of providing the content of the courses, evaluating the received knowledge and, most importantly, sharing the same goal.
- **Language learning** websites and applications. Their main goal is to improve the language proficiency of their users. For example, one of the portals matching the criteria is Duolingo, a popular language-learning platform that applies gamification techniques [78].
- **Traditional formal online learning**. Here we include learning management systems, educational software platforms as well as virtual learning environments whose main goal is to support teaching and learning. There are several articles discussing the benefits of applying these systems such as Moodle[16] [79] and WebCT[17] [80].
- **New applications**. In this group, we classify other online portals that do not fit in any of the previous categories. Here we include all *new applications* developed for learning [81] which can be classified neither as MOOCs nor as language learning or traditional learning. Moreover, in this group, some tools facilitate online learning. For instance, the authors of [82] presented an intelligent mobile learning system optimised for consuming lecture videos in both MOOCs and flipped classrooms. Another interesting example is an E-book system that allows students to preview their lessons before the class and to write questions. Moreover, they can take notes, mark part of a page as important content during the class and review the learning content after classes [83]. On the other hand, during the class, a teacher can monitor how many students are viewing the same page as the teacher, whether they are following the explanation or whether they are reading previous or subsequent pages [84].

4. Social Networks

Social networks are defined by Boyd *et al.* [85] as web-based services that allow individuals to (1) construct a public or semi-public profile; (2) articulate a list of other users with whom they share a connection; and (3) view and traverse their list of connections. These services have proved to be an exceptionally useful tool to interact with other people [86]. This has led many researchers to use them to find various trends, among which we can highlight Facebook[18] [87], [88], Twitter[19] [41] and LinkedIn[20] [89].

By way of example, the authors of [90] worked on understanding the use of *social networks* and *video games* among teenagers. They concluded that the respondents felt the desire to satisfy their needs for pleasure and communication by using these environments, without assuming they were willing to merge it with the educational process. However, several studies proved that this is not an unachievable goal. For example, Boukes [91] investigated how the use of Twitter and Facebook affected citizens' knowledge acquisition. In addition, there is research [92] done on analysing the effect of *social networks* engagement on cognitive and social skills.

5. Summary

Fig. 3 presents the distribution of the environments across the selected articles. The *content sharing & consumption* environment prevails in most published studies since there is a significant amount of research done on this topic covering various platforms.

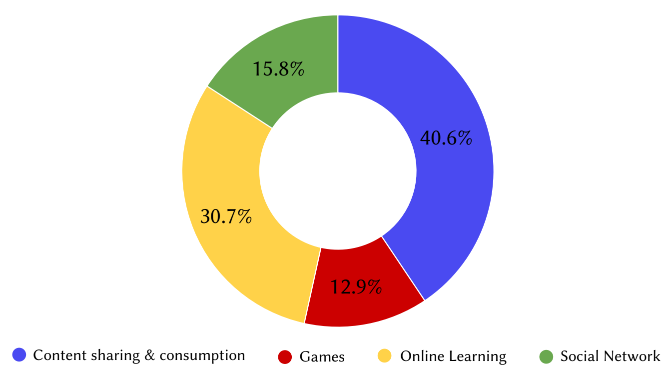


Fig. 3. Environments measuring and/or developing capabilities.

B. How are the Data for the Analysis Accessed? (RQ2)

The action 'to access data' is generally understood as the ability to retrieve, modify, copy or move data from different sources. The analysis revealed that there are four main ways of accessing data in the publications that we selected:

[16] https://moodle.org/

[17] https://blackboard.com/

[18] https://facebook.com/

[19] https://twitter.com/

[20] https://linkedin.com/

1. **Direct access**. Having *direct access* to data means that the researchers own or were provided with the data. For example, this can happen if the researcher is working for the company [93], has direct access to the systems' database [94] or is asking the organisation that owns the data for permission to access them [69].
2. **Application Programming Interface (API)**. An *API* is a set of functions that allows building and integrating the software of applications. There is a variety of public APIs that we can interact with. For instance, the Twitter API21 enables programmatic access to Twitter in unique and advanced ways that can be used to analyse, learn from and interact with tweets, direct messages, users and other key Twitter resources. For example, data were accessed through an *API* by Bouguessa et al. [95] and Bigonha et al. [44]. Finally, some studies have focused on other APIs such as the Graph API of Facebook [96], Reddit API [50], Yahoo! Answers API [53] or GitHub REST API [97], amongst others.
3. **Public domain** data. Data are available in the *public domain* where researchers can easily access them. For example, the authors of [98] collected hundreds of thousands of weblog files from Coursera using data mining techniques. Similarly, Han et al. [99] developed their own web crawling software since Pinterest does not provide an official API for data collection, and Calefato et al. [100] developed a custom web scraper for Apache Software Foundation. However, it is worth mentioning that scraping large-scale social media data from the web requires a high degree of engineering skills and computational resources [101].
4. **Open data set**. Data published as an *open data set* means that anyone can freely access these data. The difference between *public domain* data and *open data set* is that the latter is structured, well-maintained data, and released under certain public licenses that specify how others can use it. Moreover, *open data sets* are easier to access; they are already clean and ready for analysis. For this reason, there is a significant amount of papers [40], [102], [103], which simply downloaded *open data sets* for conducting their research.

Fig. 4 summarises the results of the data access across the papers. As hypothesised, our findings show that a significant number of publications had *direct access* to data or used *public domain* data. Most papers either decided to use the generated data to objectively analyse the outcome and the benefits of the product or chose to use already available data in the *public domain.*

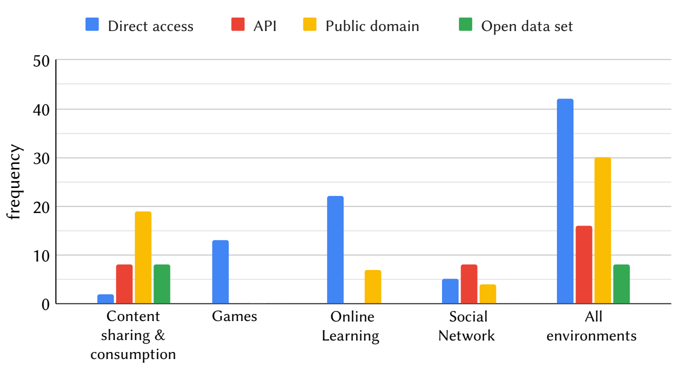


Fig. 4. Data access distribution across the selected articles.

C. What Types of Data are Found Across the Environments? (RQ3)

Initially, we aimed to examine diverse technology-mediated environments that can yield rich data sets generated by users in these environments. Therefore, data play a significant role in our survey because these data hold the potential to infer valuable information about users' competencies and capabilities. Given the aforementioned methods to access data, the developed capabilities of users were retrieved by analysing various types of data, including:

1. **Textual** data type. It is the most common type of data used among the selected publications. Most of the work (e.g., [43], [44], [60], [104]–[109]) mentioning its use has in common that it employed *textual* data generated directly by users – questions and answers, posts, tweets or source code. At the same time, other work [110], [111] generated or extracted *textual* data describing the users' statistics, including active time, the number of completed lessons, amount of lessons per day, test scores, test solutions, successes, difficulties, course progress, etc.
2. **Clickstream** data type. It represents the mouse clicks made by the user when interacting with the web environment. This is a widespread data type observed across environments – there are several studies in games, e.g. [63], [65], [112], using students' *clickstream* log files to explore their behaviour. Another example is [66], where authors extracted different behavioural features from students' evidence trace files logged by the game server such as their actions, time and performance on each specific assessment task, as well as general behavioural features not specific to assessment tasks and students' pre-test scores. Analysing this kind of data could be useful for improving online education for both teachers and students [113] by understanding how students use course resources, what contributed to their persistence and what advanced or hindered their achievements [114]. Another powerful use of *clickstream* data is described by the authors of [115], who detected cheating in MOOCs using a rule model based on heuristics.
3. **Audiovisual** data type. These data include videos, images and audio files. These data can be retrieved from any type of environment and, with the right methods, they can be informative. For example, the study described in [116] was carried out using *audiovisual* data for automated prediction and analysis of job interview performance while the authors of [81] created a system called Social Skills Trainer. It consists of a virtual avatar that recognises the user's speech and language information and can provide feedback to users to improve their storytelling skills.
4. **Biometric** data type. It is a body characteristic that can be measured or calculated. Theoretically, these data can be retrieved across different contexts; however, sophisticated instrumentation is needed to measure these data. For example, it is common to use sensors to detect eye-gaze or wearables to measure heartbeat and electrodermal activity (EDA). Therefore, researchers need to aim specifically to acquire such data as part of their research process since these data cannot be accessed easily. We found evidence of *biometric* data only in *online learning* environments for research purposes. In our context, biometric data could be useful for a better understanding of student engagement or stress resistance.

   The most representative example of the research evaluating *biometric* data is the work carried out by Peng et al. [117]. The authors proposed that, by monitoring Heart Rate Variability (HRV), it is possible to detect a person's cognitive performance. The experiment consisted in measuring the HRV of participants while they were taking part in the discussions and, through this kind of data, evaluating the discussion skills. Next, several models were adopted, such as Logistic Regression, Support Vector Machine and Random Forest. The results proved that participants' HRV data could effectively evaluate the answer quality of Q&A segments as an automatic evaluation of participants' discussion performance. This method outperformed compared with using traditional Natural Language Processing such as semantic analysis.

[21] https://developer.twitter.com/en/support/twitter-api

A visualisation of data types per environment is illustrated in Fig. 5. We can note that the vast majority of publications in environments such as *content sharing & consumption* and *social networks* use textual data while *video games* and *online learning* environments primarily generate *clickstream* data. However, all the environments mentioned previously might contain *audiovisual* data including images, videos, audios and *biometric* data such as heart rate [82] or fingerprints. Also, in general, we see a clear trend of the superiority of the *textual* data across most environments. At the same time, it is curious to mention that *online learning* environments can produce all types of data. This is probably because, for this group, we refer not only to traditional formal online learning and MOOCs but also to language learning platforms and all applications aiming to educate their users. These categories are very innovative; therefore, we could intend to use all possible data types.

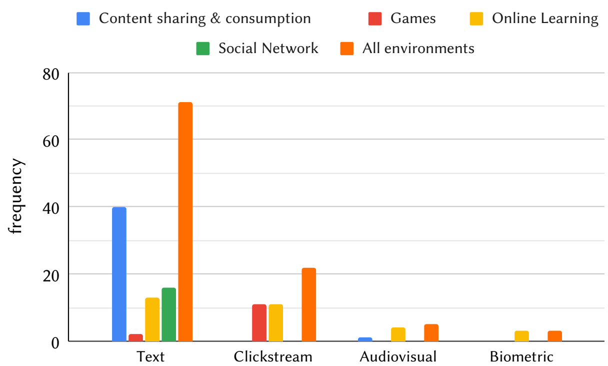


Fig. 5. Data types per environment.

D. What Methods or Techniques are Applied to Infer Competencies and Capabilities? (RQ4)

After exploring what types of data were retrieved across the publications, our goal was to examine the methods that were applied for their analysis. Accordingly, we found that several groups of methods used for analysing the data emerged. They include:

1. **Statistics**. They encompass different mathematical analyses covering many methods, tests and metrics, including Poisson models [118], Mann-Whitney U tests [119], Analysis of Variance (ANOVA) [70], Spearman's Rank Correlation Coefficient [46], Student's t-tests [102], among many others.
2. **Machine Learning (ML)**. It is an Artificial Intelligence application by means of which systems can automatically learn and improve from experience without being explicitly programmed for a specific goal. Some of the most common *ML* algorithms include NN, Support Vector Machine (SVM) [120], Naïve Bayes, Logistic Regression [89] or Random Forest [97]. Many of the analysed publications, e.g. [82], [105], [117], apply a number of these algorithms to compare their performance.

   The authors of [100] performed a quantitative analysis of data from the code commits and email messages contributed by the developers working on the large-scale distributed projects of Apache Software Foundation. They aimed to find evidence that personalities can explain developers' behaviour. The authors applied a Principal Component Analysis (PCA) to group developers with similar personalities, and cluster analysis was performed to reveal developers resembling each other but also differing from the rest. For building the contribution likelihood model, a logistic regression model was used. One of the conclusions was, for example, that the propensity to trust others turned out to be positively influential on the result of code reviews in distributed projects.
3. **Network analysis**. It is an analytical method used to evaluate relationships between nodes that are a part of a connected network. Bouguessa et al. [121] stated that most of the existing approaches attempting to discover experts model the environment as a graph in which the nodes represent users and the edges represent the interactions between them. In this way, the authority score is generally measured through graph-based ranking algorithms such as PageRank, Hyperlink-Induced Topic Search (HITS), InDegree Algorithm, etc. [122].

   In turn, the authors of [123] described the idea underlying the HITS algorithm as follows: "A good authority is one that is pointed to by many good hubs, and a good hub is one that points to many good authorities." [p. 238] In simple words, the quality of a page as an authority depends on the quality of the pages that point to it as hubs and vice versa. The HITS algorithm was used by Jurczyk et al. [54] for predicting experts in Q&A portals, and its effectiveness was proved by performing a large-scale empirical evaluation.
4. **Natural Language Processing (NLP)**. It combines various techniques from different computer science areas, linguistics, and artificial intelligence to interpret, process and analyse human language, often on a large scale. One of the papers using *NLP* is [50], where authors performed a semantic analysis by using a text analysis software called Linguistic Inquiry and Word Count, which counts words that belong to psychologically meaningful categories.
5. **Experimental design**. This approach can be used when researchers are interested in evaluating the impact that specific design decisions or characteristics can have on different outcomes. For example, they can conduct an experiment dividing participants into two groups, where the experimental group is the one that tests a new feature, whereas the control group does not test it. For example, Lesser [64] aimed to determine the effectiveness of digital game-based learning compared to other teaching methods related to music education. An experimental study consisting of test and control groups together with in-depth interviews led to the following results: students who had access to educational *video games* combined with the assistance of an instructor achieved higher mean scores compared to students who had access to either *video games* without instruction or instruction without *video games*. This author suggested that educational *video games* may be potentially used as an effective tool in the music classroom to teach musical concepts and skills as well as to increase student motivation, engagement and a hands-on approach to learning.

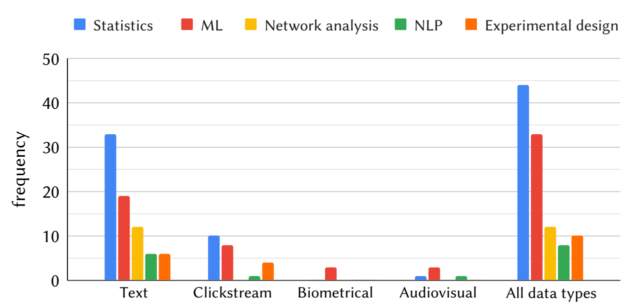


Fig. 6. Methods per data type.

A visualisation of data types per method is illustrated in Fig. 6. It shows that many publications applied statistical methods and ML models. This is because statistics and ML are broad fields that include a variety of different subfields. Next, we analysed how the applied methods spread across the data types. As we see, on *textual* data, all the identified methods were performed. This can be explained by the fact that many more publications focus on this type of data, and therefore there are many more examples of it. Besides, using *textual* data by default implies the possibility of applying various methods. Moreover, we can notice that *ML* is applied to all types of data. With a huge development of this field and the improvements of its methods, their utilisation across various data types has become increasingly widespread.

E. What Competencies and Capabilities are Measured and/or Developed Across the Environments? (RQ5)

For this RQ, we examined diverse technology-mediated environments with the ability to generate rich data sets through the users' interaction. We observed that these data could be used to explicitly or implicitly perform a data-driven evaluation of capabilities. In this section, we summarise the main competencies and capabilities that emerged from the survey, namely, **expertise**, **language proficiency**, **soft skills** and patterns of **behaviour** that were targeted across the different studies.

1. Expertise

Expertise is defined by Herling [124] as the optimal level at which a person is able and/or expected to perform within a specialised realm of human activity. This broad group of capabilities includes:

- **Computer science expertise**. It is defined as proficiency in different programming languages, libraries or tools. The studies aiming to evaluate such skills mainly focused on portals highly related to the field of computer science such as GitHub [97], [125]–[127], StackOverflow [51], [128], [129] or both of them simultaneously [130]. In general, many researchers concluded that information technology capabilities are key drivers to achieve superior customer relations and innovation [131].
- **Topical authority** in a selected topic. The most suitable portals for finding this kind of *expertise* are forums and Q&A websites since they already focus on a particular topic and their users generate a substancial amount of data suitable for the analysis. This analysis revealed that the following portals have been used for this research objective: Yahoo! Answers [121], [132], Quora [49], [133], Reddit [103], AskMe forum[22] [134], TurboTax Live Community[23] [135], MedHelp[24] [136] and Tianya Wenda[25] [137].

  There are many examples of successful work using data to detect *topical authority* in a selected topic with significant results. The most representative one is the research conducted by Abdaoui et al. [138]. The authors aimed to detect posts written by medical experts in health forum discussions. They managed to collect more than 28,000 textual messages from two specialised websites. Through these data, a supervised learning approach to distinguish posts written by medical experts and by patients in health forums was followed. This research shows that it is possible to detect topical experts.
- **Learning outcomes metrics**. We can find this kind of *expertise* only in *online learning* environments since these are the only ones providing tasks to students and evaluating the results. For example, Brinton et al. [139] presented two frameworks whose purpose is to represent video-watching clickstreams: one based on the sequence of events created and another on the sequence of viewed videos. The authors extracted students' actions such as reflecting (repeatedly playing and pausing) and revising (plays and skip backs). The authors concluded that some of this behaviour is significantly associated with user performance in online learning.

The aforementioned portals generate a significant amount of data, which can allow the detection of potential experts. Although this large volume of user-generated content is a potential strength, it also makes the problem of finding authoritative users for a given topic challenging [57]. According to Riahi et al. [39], experts are often not provided with questions matching their expertise and, therefore, new questions may not be matched with an expert properly; and hence they end up without receiving a proper answer. For this reason, improving expertise finding algorithms would be useful to enhance the user experience in these portals.

2. Language Proficiency

Language proficiency refers to a person's ability to correctly use a certain language in terms of grammar, fluent speaking, lexical understanding, etc. We have found two main approaches to infer this capability:

- In environments specifically designed for developing language skills, i.e., language-learning portals such as Duolingo or Babbel, whose main purpose is to help users to fully learn a language through specifically tailored online activities and courses.
- When researchers collect data from other environments where users can exhibit evidence of *language proficiency* capabilities. For example, the authors in [140] investigated whether Twitter could support creative writing development and in [141] English language learners' use of Instagram was explored.

While we found multiple studies detecting improvements in learning foreign languages when learners use various *social networks*, these studies frequently used *experimental design* to measure the impact of *social networks* on the use of language learning [87], [142]. Therefore, they were not directly using the data from learners to evaluate their language capabilities, and consequently, these studies are only weakly connected to our goal.

By way of conclusion, we did not find many studies that firmly fit our survey for this particular capability. One of those that fit well, however, is a spoken *language proficiency* assessment system called Dolphin [143]. Its goal is to automatically evaluate students' phonological fluency and semantic relevance by analysing students' video clip and verbal fluency tasks. The results proved that Dolphin could provide more opportunities to practice and improve their oral language skills, and at the same time, it could reduce teachers' grading burden. The experiments demonstrated the effectiveness in model accuracy, system usage, teacher satisfaction rating and other metrics.

3. Soft Skills

Soft skills are described as a combination of interpersonal and social skills [144]. They are used to indicate personal transversal competencies and personality traits that characterise relationships between people [145]. We found several studies that can be divided into the following *soft skills* capabilities:

- **Executive control skills**. They are defined by Strobach et al. [146] as the control and management of other cognitive processes as well as cognitive skills, working memory and attention skills. One of the studies exploring the relationship between executive control skills and action *video games* is [146]. In this work, video gamers were shown to improve performance in dual-task and task switching situations in comparison with nongamers. Another sample study measuring cognitive skills, working memory and attention skills is the work carried out by Alloway et al. [92]. They measured the impact of *social networks* engagement on cognitive skills and social connectedness.
- **Creativity**. It can be understood as the ability to generate ideas that could be beneficial for problem-solving, communication and

[22] https://ask.metafilter.com/

[23] https://ttlc.intuit.com/

[24] http://medhelp.org/

[25] http://wenda.tianya.cn/

entertainment. The authors of [147] stated that creativity could be inculcated, encouraged and trained. For this purpose, they developed a digital game-based learning system to foster students' creativity.

- **Problem-solving skills**. They represent a range of attitudes and thinking skills that are used to find solutions to problems [148]. Chu et al. [149] proposed a game-based development approach for improving these skills. They conducted an experiment in an elementary school natural science course aiming to evaluate the performance of their approach. Finally, it was proved that the proposed game development-based learning approach could effectively promote the students' problem-solving skills.
- **Critical thinking**. It is defined as 'the intellectually disciplined process of actively and skilfully conceptualising, applying, analysing, synthesising and/or evaluating information gathered from or generated by, observation, experience, reflection, reasoning or communication, as a guide to belief and action' [150]. Chootongchai et al. [79] stated that it is crucial to have a range of thinking and innovation skills, including critical thinking, collaboration and communication, to be successful at work and in life more generally. Accordingly, they developed an online learning system to enhance thinking and innovation skills for higher education learners.
- **Social skills**. We could only find one example of a study evaluating social skills. The authors of [81] developed a dialogue system called Automated Social Skills Trainer that can decrease human anxiety and discomfort in social interaction and help acquire social skills through human-computer interaction. The system includes a virtual avatar that recognises user speech as well as language information and gives feedback to users to help them improve their social skills.

4. Behaviour

Within this group, we mainly consider various behavioural patterns, including engagement, influence or dropout, amidst others. We cannot name them as capabilities, but they are essential for our survey because, as stated in Section B, through publications describing *behaviour*, we can learn of additional studies that hold the potential to perform a data-driven evaluation of competencies.

5. Summary

The distribution of the developed capabilities per environment is represented in Fig. 7.

We observe that all the stated environments prove to measure capabilities of a different nature. We also see that *content sharing & consumption* environment prevails in the development of expertise. This is because, in such environments, we consider many forums and similar websites where it is possible to observe experts from different fields.

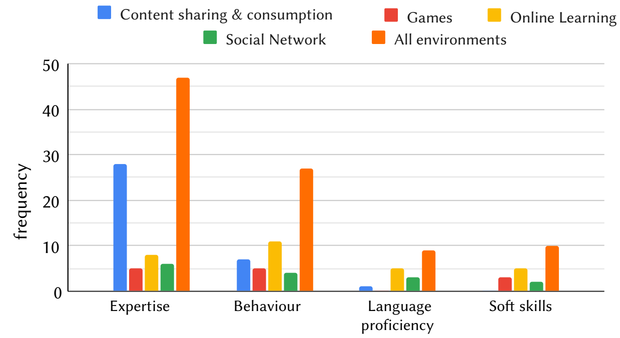


Fig. 7. Capabilities per environment.

F. Are the Findings Across the Publications Validated? (RQ6)

One of the most significant parts of the research reviewed in our survey is the validation of the results. Validation is intended to ensure that the proposed methods and their results proved satisfactory by conducting appropriate experiments. According to our survey topic, only the validation made by humans was considered since it has high reliability and validity.

A close inspection revealed that up to 81.4% of all the analysed studies do not validate their results. After a careful analysis of the publications that did not validate their results, we can conclude that the results' validation was beyond their research goals. Another possible explanation could be that it was not feasible to validate the results. For example, it is not a trivial task to verify that soft skills such as adaptability or stress management skills were obtained.

Regarding the publications that did validate their findings, we saw that almost all of them verify the expertise search results (see Fig. 8). Expertise has an understandable way of validation. The most common one is described in [151]. This paper presented an approach to identify potential experts. The most noteworthy aspect of this work is that the authors also proposed a method to detect users likely to become experts in the future through their behaviour and estimation of their motivation and ability to help others. After building the models and having the results, a human evaluation was performed. The authors asked community managers to evaluate the potential experts identified by the algorithm, and the analysis revealed that there is a high agreement between the human evaluation and the performed algorithms.

Another interesting approach to performing validation is through another portal. For example, the authors of [130] followed a two-step approach. The first step was to measure developers' commit activity on GitHub by considering both the quantity and the continuity of their contributions in isolated projects over time. The second step was to evaluate the generated developers' expertise profiles against recognised answering activity on StackOverflow via a data set of users that were active both on GitHub and on StackOverflow.

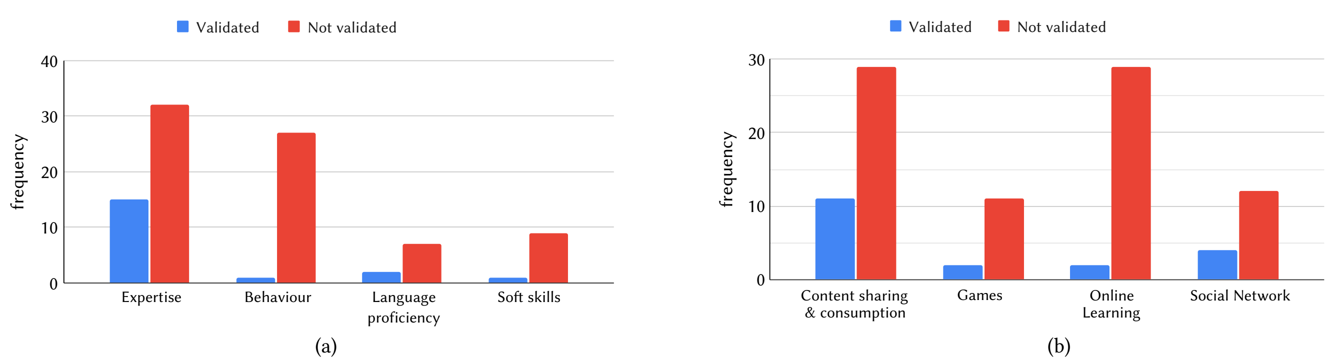


Fig. 8. Validation results: (a) per capability and (b) per environment.

To sum up, it is important to conclude that we did not find many studies performing manual validation of their results. This could be related to the fact that human validation requires much more time and effort and, therefore, the researchers decided to give preference to other methods. As we can also see in Fig. 8, validation by humans was performed mainly in *expertise finding* through the *content sharing & consumption* environment.

G. What are the Main Limitations or Challenges Faced by the Authors of the Studies? (RQ7)

Limitations show potential weak points of the study, and researchers can encounter them due to constraints in research design or methodology. We can group the limitations that the authors faced as follows:

- **Data**. Data were crucial for our survey because our goal was to explore diverse multimedia environments generating rich data sets through the users' interaction and where these data can be used to perform a data-driven evaluation of capabilities. Therefore, we examined the limitations related to data in much detail and reached several conclusions. Firstly, data access limitation is connected with difficulties encountered during the data retrieval, such as being limited in accessing data or not having the right to do so. One of the articles dealing with data access limitation is [152]. In this study, the authors claimed that many researchers who used data extracted from platforms with APIs faced the same limitation. Secondly, the authors of [151] stated that for expertise finding, they only considered a single Q&A site with a narrow purpose and an active team of professionals behind the scene. At the same time, a vast majority of studies, e.g. [44], [88], [104] aim to use more data of the already selected portal or to apply their method in a data set from a new portal as a part of theirfuture work. One more issue of not having large enough data sets can be that results cannot be accurate and reliable enough [93]. Therefore, it is vital to have enough data and to know how to deal with inaccurate, misleading or biased data [56]. In total, 13 publications raised this limitation.
- **Methods**. This limitation was also crucial since researchers' results highly depend on the applied methods to measure the capabilities. There are several studies [93] stating that one of their main limitations was the way in which the selected method was applied. For example, Zhu et al. [137] stated that their category relevancy-based authority ranking approach needed a more accurate and stable method for parameter selection. Moreover, there is another type of work claiming that more analyses [43] or methods should have been applied. In total, we found eight publications with this limitation.
- **Labelling**. Labelling data comprehensively and efficiently is a widely needed but challenging task [153]. However, manual labelling of an unknown data set for ML is a tedious task for humans [154]. Manual labelling can be necessary if there is no automated data preprocessing system [155] and, by default, it can lead to limitations related to human reliability and possible consequences of human errors or oversights. Another issue is the scalability of manual labelling since it is a hard task when the amount of data is enormous. Moreover, human evaluation may also have biases because different raters may consider different criteria [132]. Finally, the time spent in the process of labelling directly translates into the high costs associated with research projects [156]. In total, we have found two publications facing this limitation.

In conclusion, we can point out that the weakest parts of the selected articles are the lack of data while the methodology could be improved in some studies. However, only 23 publications raised the limitations mentioned above; the rest of the articles mainly discussed future directions, not mentioning weak points.

V. Discussion

In this section, we first present a summary and discussion of our main findings. Next, we provide a discussion that goes beyond those findings. Finally, we will present the implications of our research and the limitations of the selected approach.

A. Key Findings

Here we summarise the main outcomes of our research work, highlighting the most important parts of our survey results.

First of all, four environments, namely, *content sharing & consumption*, *games*, *online learning* and *social networks* emerged from the coding process. Across these environments, we observed measurement and/or development of various capabilities such as *expertise*, *language proficiency* and *soft skills* as well as various behavioural patterns including engagement, influence or dropout, which we grouped as *behaviour*. The most striking result that emerged is that all environments are significantly correlated with all the capabilities stemming from the survey. Further analysis showed that the *content sharing & consumption* environment prevails in the publications. Similarly, strong evidence of the development of expertise was found in this environment.

In an attempt to perform this analysis, we first extracted ways of accessing data in the selected publications, that is, using data published as an *open data set*, using an *API*, using data from the *public domain* or having *direct access* to data. The last two substantially prevailed across the analysed studies. Next, four types of data emerged: *textual*, *clickstream*, *audiovisual* and *biometric*. Remarkably, the vast majority of publications in environments such as *content sharing & consumption* and *social networks* used *textual* data while *video games* and *online learning* environments primarily generated *clickstream* data. After exploring what types of data were found, we aimed to examine the data analysis methods. According to the selected articles, such analysis was conducted through various methods, namely, *ML*, *Network Analysis*, *NLP*, *statistics* and *experimental design*. Our study provided further evidence that all the methods mentioned above have been applied to *textual* data, but that the other types of data are not as flexible in terms of methods. Ultimately, we discussed whether the authors validated the results and what limitations they found in their work. We did not find many studies performing manual validation of their results; however, among those that did, validation of *expertise* finding through the *content sharing & consumption* environment was the most frequent. Lastly, we concluded that the main limitations raised as part of the selected articles could be the lack of data or the methods applied to them.

Fig. 9 shows a summary of the environments and their categorisation with the corresponding data types, their access type and applied methods to validate different competencies and capabilities. We present the results based on the surveyed studies; however, the existence of other types of multimedia environments remains an open question.

B. Extending Beyond the Results

We explored several digital environments that can generate rich data sets through the users' interaction and where data can be used to explicitly or implicitly perform a data-driven evaluation of competencies and capabilities. From our survey, we can extract several general characteristics that environments needed to have for data-driven evaluation of capabilities. First of all, publications within the scope of our survey explored environments that can generate large amounts of data. These data can be accessed either with *direct access* to data or in the *public domain*, but at the same time, it is possible to do it through an *API* or to download an *open data set*. Either way, data of different

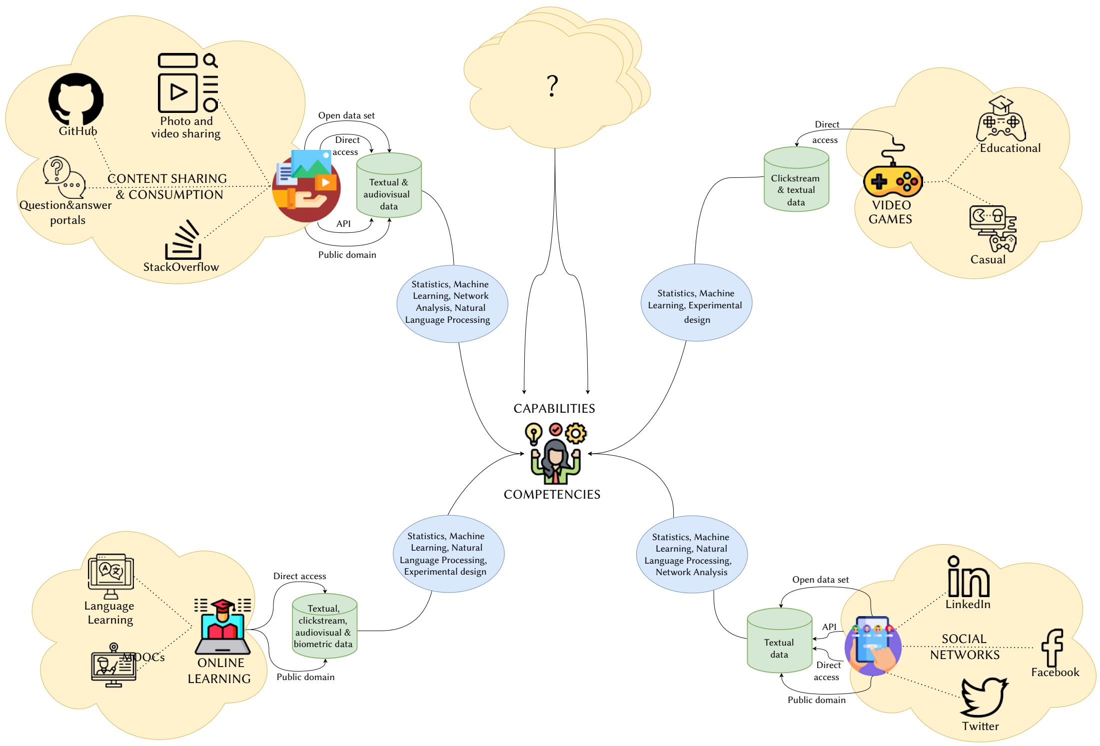


Fig. 9. Overall summary of environments, data access, data type and methods.

types were accessed, where *clickstream* and *textual* data formats prevail over *audiovisual* and *biometric* data. Data access and applied methods are based on the most common types of data; therefore, the rest might prove to be more challenging. Accordingly, other digital environments not found as part of this survey but which fit these criteria also represent potential opportunities to achieve this goal.

As a case in point, we would like to mention several specific environments found in our survey that do not measure capabilities but hold the potential to do so. One of them is Netflix[26] – a media streaming platform described in detail in [157]. Unlike publications within the scope of our survey, this study described the Netflix recommendations system, which helps its users make better decisions. Nevertheless, the authors used vast amounts of data that describe what each Netflix user watches, in what way s/he does it (e.g., the device, time of the day, day of the week, the intensity of watching), the place in which each video was discovered and even the recommendations shown but not played in each session. We believe that these data could hold the potential to evaluate various competencies and capabilities such as language proficiency, among others.

Another example is the social payments platform Venmo[27], which was one of the most surprising findings of our research. It does not match any of the environments detected during the review process, and it does not fit, a priori, the survey. However, we found one publication [152] aiming to understand users' changes in behaviour over time. Its authors accessed nearly 340 million transactions. Each transaction consists of the transaction ID, sender, receiver, message, time created, amongst other kinds of metadata generated from public posts. Since the authors have already explored users' behaviour, we believe that this portal has the potential for conducting the research according to the goal of our work.

Furthermore, interactive museums turned out to be another unexpected environment that attracted the attention of some researchers. The authors of [158] measured learners' engagement by using multi-channel data such as eye-tracking, facial expression, posture and interaction logs. These data were captured from visitors' interactions with a fully-instrumented version of a tabletop science exhibit for environmental sustainability called FutureWorlds. We consider it as an environment with ample opportunities for evaluating its users' competencies and capabilities.

C. Implications/Limitations

Our initial objective was to develop a survey on previous work that has performed a data-driven evaluation of competencies across different multimedia environments. The rationale is that the world is shifting towards a focus on capabilities instead of content. Therefore, this issue is of the utmost importance as it will lead society to better adapt to the jobs of the future. Thus, we have forseen several implications of our results. While the research that we have found shows that it is methodologically feasible to measure competencies based on the data generated in those multimedia environments, it is unclear how to translate these findings into the formal education ecosystem. Some possibilities might include a more frequent utilisation of educational games and other digital environments as part of the classroom activities so that teachers can receive information regarding their students' capabilities to provide personalised feedback. Therefore, more research, products and validations in formal educational settings will be required during the next decade. In that sense, our survey could become a starting point for further research, since it confirms the methodological viability of these approaches. Moreover, it examines new multimedia data-rich environments and their opportunities to support the development of lifelong and lifewide 21st-century capabilities. The current study has some theoretical and practical implications and limitations which will be outlined in this section.

1. Theoretical Implications

Any data set invariably constitutes a biased representation of the population [159]. Moreover, there are unfair practices against

[26] https://netflix.com/

[27] https://venmo.com/

members of vulnerable or underrepresented groups, which include the explicit use of protected data attributes such as age or gender, as well as indirect discrimination that occurs when group status is exploited inadvertently [160], [161]. Other biases in data may involve race or ethnicity. The authors of [162], for example, showed that a widely used algorithm, typically used for industry-wide approaches and which is affecting millions of patients, exhibits significant racial bias. Therefore, more work is required for ensuring the fairness and equity of the methods applied in these studies.

2. Practical Implications

We should like to discuss the research challenges and gaps that arise today within this topic. It is essential to mention that we observed a considerable potential based on many studies spanning diverse multimedia environments. We found that every single study applied different methodologies and processes to transform data into capabilities. This means that there is a significant effort invested within this process in the studies. This problem can potentially be improved by proposing a base framework that can be applied to infer capabilities based on data while generalising well across different digital environments; to the best of our knowledge, this kind of framework does not yet exist. Accordingly, we believe that our work can serve as a knowledge basis for building it. However, we have detected that there are not many examples of using *audiovisual* or *biometric* data in the studies that we explored. This could lead to having insufficient evidence of the development of methods that use those data types to measure some capabilities across various environments. Moreover, easy identification of an individual with very few data points leads to the restricted ethical use of data and their purposes. Therefore, while having the potential to be useful, it might not be possible to use some data given the aforementioned issue and the fact that users' privacy must be respected.

3. Limitations

The most important limitation of this work that could influence the obtained results lies in the search process. This is because the present study only investigated the environments in which we could find evidence of measuring competencies and capabilities that emerged from the coding process (see Section D). Even though we covered the most relevant environments in the context of the data-driven evaluation of competencies and capabilities, there could be other environments we are not aware of, and consequently, they are not included in our study. Despite this fact, we assume that we found evidence of capabilities evaluation across all the environments discussed throughout this survey.

4. Novel Contributions

The strength of the current paper is that we identified and reviewed studies that have been able to use different types of data and analyses to infer a range of competencies and capabilities in the four multimedia environments that emerged as part of the survey. As we mentioned in Section II, all previous studies explored how to develop and measure competencies based on only one specific type of online platform. Our work has provided a more complete picture of the multiple existing multimedia environments that can be used for evaluating different competencies and capabilities. What we learned in the survey is a key starting point for the potential change in the educational and training systems, suggesting new data-driven assessment possibilities that represent new steps forward to provide personalised feedback. At the same time, our work may motivate other researchers to perform additional experiments to learn of new digital environments holding the potential to measure and/or develop various capabilities.

VI. Conclusions and Future Work

This work represents a groundbreaking analysis of current literature examining diverse technology-mediated environments that can generate rich data sets through the users' interaction and where data can be used to perform a data-driven evaluation of competencies and capabilities. This is the first time, as far as we know, that this kind of research was conducted. Despite facing an ill-defined area, this study deeply enhanced our current understanding of this open research line. In this regard, we provided an overview of the existing research as well as concluded that all the environments we discussed (*content sharing & consumption*, *video games*, *online learning* and *social networks*) proved their ability to generate rich data sets through the users' interaction. We found evidence that all these environments are highly correlated with the measurement and/or development of various capabilities such as *expertise*, *language proficiency* and *soft skills*. According to the over one hundred surveyed studies, this measurement was done with the application of different methods (*ML, Network Analysis, NLP, statistics* and *experimental design*), which we also discussed in detail.

We believe that our survey encompasses numerous new approaches that confirm the viability of performing data-driven evaluations of competencies and capabilities. We are confident that based on our results, it is possible to develop a framework that can generalise well to different environments, data types and capabilities, and that this can help to conduct additional research by re-applying the framework in future studies. Accordingly, more research is needed to be able to transfer these ideas into formal education settings with new innovative products. In the future, teachers will be able to use such products to better assess the capabilities of their students and to provide personalised feedback. On the other hand, there is a need to develop this research evaluating the algorithmic bias issues as well as being respectful of students' privacy. Our future work will focus on exploring several of these multimedia environments with the aim of developing our own algorithms for measuring the capabilities. More specifically, we will focus on the measurement of expertise and soft skills across different environments, trying to analyse different types of data by applying various methods.

Appendix

A. Acronyms

Acronym	**Reference abbreviation**
MOOC	Massive Open Online Course
Q&A	Question-and-Answer format
CQA	Community Question Answering
RQ	Research Questions
OVP	Online Video Platform
API	Application Programming Interface
HRV	Heart Rate Variability
ANOVA	Analysis of Variance
ML	Machine Learning
NN	Neural Network
SVM	Support Vector Machine
HITS	Hyperlink-Induced Topic Search
NLP	Natural Language Processing
EDA	Electrodermal activity
PCA	Principal Component Analysis
N/A	Not Applicable
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses

B. Results of the Coding Process

Title	Environment (RQ1)	Data access (RQ2)	Data (RQ3)	Methods (RQ4)	Skills (RQ5)	Validation (RQ6)
[134]	CSC[1]	Public domain	Text	Network analysis	Expertise	No
[49]	CSC	Public domain	Text	ML	Expertise	No
[54]	CSC	Public domain	Text	Network analysis	Expertise	No
[57]	CSC	N/A	Text	Statistics	Expertise	No
[151]	CSC	Public domain	Text	ML	Expertise	Yes
[129]	CSC	Open data set	Text	ML	Expertise	No
[103]	CSC	Open data set	Text	Statistics	Expertise	No
[133]	CSC	Public domain	Text	Statistics	Expertise	No
[122]	CSC	Public domain	Text	Network analysis	Expertise	No
[128]	CSC	Open data set	Text	Statistics	Expertise	No
[138]	CSC	Public domain	Text	ML	Expertise	Yes
[55]	CSC	N/A	Text	Statistics	Expertise	No
[135]	CSC	Public domain	Text	ML	Expertise	No
[39]	CSC	N/A	Text	Network analysis, NLP	Expertise	No
[60]	CSC	API	Text	N/A	Expertise	Yes
[61]	CSC	API	Text	Statistics	Expertise	Yes
[121]	CSC	Public domain	Text	Network analysis	Expertise	Yes
[126]	CSC	API	Text	ML	Expertise	Yes
[136]	CSC	Public domain	Text	ML	Expertise	No
[132]	CSC	Public domain	Text	Network analysis	Expertise	Yes
[59]	CSC	Public domain	Text	ML	Expertise	No
[125]	CSC	API	Text	Statistics	Expertise	Yes
[137]	CSC	Public domain	Text	Network analysis	Expertise	Yes
[53]	CSC	API	Text	Network analysis	Expertise	Yes
[109]	CSC	Public domain	Text	Network analysis	Expertise	Yes
[50]	CSC	API	Text	NLP	Behaviour	No
[97]	CSC	API	Text	ML	N/A	No
[58]	CSC	Public domain	Text	Statistics	N/A	No
[46]	CSC	Open data set	Text	Statistics	Behaviour	No
[56]	CSC	Public domain	Audiovisual	Statistics	Expertise	No
[108]	CSC	Public domain	Text	Statistics	Expertise	No
[42]	CSC	N/A	Text	Statistics	Behaviour	No
[127]	CSC	Public domain	Text	NLP	N/A	No
[99]	CSC	Public domain	Text	Statistics	Behaviour	No
[51]	CSC	Open data set	Text	Statistics	Behaviour	No
[102]	CSC	Open data set	Text	Statistics	Behaviour	No
[157]	CSC	Direct access	Text	ML	N/A	No
[101]	CSC	API	Text	N/A	N/A	N/A
[141]	CSC	Direct access	Text	Statistics	Language[2]	No
[40]	CSC	Open data set	Text	Network analysis, ML	Behaviour	No
[130]	CSC	Open data set	Text	Statistics	Expertise	No
[93]	Video games	Direct access	Clickstream	ML	Expertise	No
[69]	Video games	Direct access	Clickstream	ML	Expertise	No
[71]	Video games	Direct access	Clickstream	Statistics	Behaviour	No
[65]	Video games	Direct access	Clickstream	ML	Expertise	Yes
[70]	Video games	Direct access	Clickstream	Statistics	Behaviour	No
[66]	Video games	Direct access	Clickstream	ML	Expertise	No
[146]	Video games	Direct access	Clickstream	Statistics, Experiment[3]	Soft skills	No
[147]	Video games	Direct access	Clickstream	Experiment	Soft skills	No
[149]	Video games	Direct access	Clickstream	Experiment	Soft skills	Yes
[155]	Video games	Direct access	Text	Statistics	Behaviour	No
[64]	Video games	Direct access	Text	Experiment	Expertise	No

Title	Environment (RQ1)	Data access (RQ2)	Data (RQ3)	Methods (RQ4)	Skills (RQ5)	Validation (RQ6)
[63]	Video games	Direct access	Clickstream	ML	Behaviour	No
[112]	Video games	Direct access	Clickstream	Statistics	Behaviour	No
[100]	Online Learning	Public domain	Text	ML	Soft skills	No
[98]	Online Learning	Public domain	Clickstream	Statistics	Expertise	No
[78]	Online Learning	Public domain	Text	Statistics	Language	No
[82]	Online Learning	Direct access	Biometric	ML	Expertise	No
[81]	Online Learning	Direct access	Audiovisual	ML	Soft skills	No
[117]	Online Learning	Direct access	Biometric	ML	Soft skills	No
[73]	Online Learning	N/A	Clickstream	NLP	Expertise	No
[143]	Online Learning	Direct access	Audiovisual	ML	Language	No
[142]	Online Learning	Direct access	Text	Experiment	Language	Yes
[110]	Online Learning	Direct access	Text	Statistics, Experiment	Language	No
[83]	Online Learning	Direct access	Text	Statistics	Behaviour	No
[76]	Online Learning	Direct access	Clickstream	ML	Behaviour	Yes
[139]	Online Learning	Direct access	Clickstream	Statistics	Expertise	No
[74]	Online Learning	Direct access	Audiovisual	NLP	Soft skills	No
[114]	Online Learning	Direct access	Clickstream	Statistics	Soft skills	No
[75]	Online Learning	N/A	Text	Statistics	Expertise	No
[116]	Online Learning	Direct access	Audiovisual	ML	Expertise	No
[48]	Online Learning	Direct access	Text	NLP	Behaviour	No
[106]	Online Learning	Public domain	Text	Network analysis	Behaviour	No
[158]	Online Learning	Direct access	Biometric	ML	Behaviour	No
[111]	Online Learning	Direct access	Text	Statistics	Language	No
[47]	Online Learning	Public domain	Clickstream	ML	Behaviour	No
[120]	Online Learning	Public domain	Clickstream	ML	Expertise	No
[118]	Online Learning	Direct access	Clickstream	Statistics	Behaviour	No
[84]	Online Learning	Direct access	Clickstream	Experiment	Behaviour	No
[77]	Online Learning	Public domain	Clickstream	Statistics	Behaviour	No
[113]	Online Learning	Direct access	Clickstream	Statistics	Behaviour	No
[119]	Online Learning	Direct access	Text	Statistics	Expertise	No
[80]	Online Learning	Direct access	Text	Statistics	N/A	No
[79]	Online Learning	Direct access	Text	Statistics, Experiment	N/A	No
[94]	Online Learning	Direct access	Text	Statistics, Experiment	Behaviour	No
[52]	Social Network	Public domain	Text	Statistics	Expertise	No
[95]	Social Network	API	Text	ML	Expertise	Yes
[107]	Social Network	Direct access	Text	ML	Expertise	Yes
[140]	Social Network	API	Text	ML	Language	No
[41]	Social Network	Public domain, API	Text	ML	Expertise	No
[96]	Social Network	API	Text	ML	Soft skills	No
[89]	Social Network	Direct access	Text	ML	Expertise	Yes
[88]	Social Network	Public domain	Text	Statistics	Language	No
[105]	Social Network	API	Text	ML	Behaviour	No
[45]	Social Network	API	Text	Statistics	Behaviour	No
[87]	Social Network	Direct access	Text	Statistics, Experiment	Language	Yes
[43]	Social Network	API	Text	NLP	Behaviour	No
[44]	Social Network	API	Text	Network analysis	Behaviour	No
[91]	Social Network	Direct access	Text	Statistics	Expertise	No
[104]	Social Network	Public domain	Text	NLP	N/A	No
[92]	Social Network	Direct access	Text	Statistics	Soft skills	No
[152]	N/A	API	Text	Statistics	Behaviour	No

[1] Content sharing & consumption
[2] Language proficiency
[3] Experimental design

Acknowledgment

This study was partially funded by the Spanish Government grants IJC2020-044852-I and RYC-2015-18210, co-funded by the European Social Fund, as well as by the COBRA project (10032/20/0035/00), granted by the Spanish Ministry of Defence and by the SCORPION project (21661-PDC-21), granted by the Seneca Foundation of the Region of Murcia, Spain.

References

[1] B. S. Bloom, "The new direction in educational research: Alterable variables," *The Journal of Negro Education*, vol. 49, no. 3, pp. 337–349, 1980.

[2] C. R. Wolfe, "Creating informal learning environments on the world wide web," in *Learning and teaching on the World Wide Web*, Elsevier, 2001, pp. 91–112.

[3] S. Downes, *et al.*, "New technology supporting informal learning," *Journal of emerging technologies in web intelligence*, vol. 2, no. 1, pp. 27–33, 2010.

[4] A. Nandi, M. Mandernach, "Hackathons as an informal learning platform," in *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, SIGCSE '16, New York, NY, USA, 2016, p. 346–351, Association for Computing Machinery.

[5] J. Maldonado-Mahauad, M. Pérez-Sanagustín, R. F. Kizilcec, N. Morales, J. Munoz-Gama, "Mining theory-based patterns from big data: Identifying self-regulated learning strategies in massive open online courses," *Computers in Human Behavior*, vol. 80, pp. 179 – 196, 2018, doi: https://doi.org/10.1016/j.chb.2017.11.011.

[6] M. Anshari, Y. Alas, L. S. Guan, "Developing online learning resources: Big data, social networks, and cloud computing to support pervasive knowledge," *Education and Information Technologies*, vol. 21, no. 6, pp. 1663–1677, 2016.

[7] R. Kizilcec, C. Brooks, "Diverse Big Data and Randomized Field Experiments in Massive Open Online Courses," in *The Handbook of Learning Analytics*, C. Lang, G. Siemens, A. F. Wise, D. Gaševic Eds., Alberta, Canada: Society for Learning Analytics Research (SoLAR), 2017, pp. 211–222, 1 ed.

[8] R. Eynon, "The rise of big data: what does it mean for education, technology, and media research?," *Learning, Media and Technology*, vol. 38, no. 3, pp. 237– 240, 2013, doi: 10.1080/17439884.2013.771783.

[9] J.-E. Mai, "Big data privacy: The datafication of personal information," *The Information Society*, vol. 32, no. 3, pp. 192–199, 2016.

[10] V. Mayer-Schönberger, K. Cukier, *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt, 2013.

[11] J. Van Dijck, "Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology," *Surveillance & Society*, vol. 12, no. 2, pp. 197–208, 2014.

[12] H.-U. Otto, H. Ziegler, "Capabilities and education," *Social Work & Society*, vol. 4, no. 2, pp. 269–287, 2006.

[13] D. Kember, D. Y. Leung, R. S. Ma, "Characterizing learning environments capable of nurturing generic capabilities in higher education," *Research in Higher Education*, vol. 48, no. 5, p. 609, 2007.

[14] M. Pinzone, P. Fantini, S. Perini, S. Garavaglia, M. Taisch, G. Miragliotta, "Jobs and skills in industry 4.0: An exploratory research," in *Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing*, Cham, 2017, pp. 282–288, Springer International Publishing.

[15] A. Smith, J. Anderson, "AI, robotics, and the future of jobs," *Pew Research Center*, vol. 6, p. 51, 2014.

[16] N. P. Stromquist, *Education in a globalized world: The connectivity of economic power, technology, and knowledge*. Rowman & Littlefield, 2002.

[17] C. Redecker, M. Leis, M. Leendertse, Y. Punie, G. Gijsbers, P. Kirschner, S. Stoyanov, B. Hoogveld, "The future of learning: New ways to learn new skills for future jobs," *Results from an online expert consultation. Technical Note JRC60869, JRC-IPTS, Seville*, 2010.

[18] J. C.-Y. Sun, R. Rueda, "Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education," *British Journal of educational technology*, vol. 43, no. 2, pp. 191– 204, 2012.

[19] R. Nagarajan, R. Prabhu, "Competence and capability: A new look," *International Journal of Management*, vol. 6, no. 6, pp. 7–11, 2015.

[20] R. Hipkins, "Competencies or capabilities," *He Whakaaro An, Se2*, vol. 3, pp. 55–57, 2013.

[21] R. E. Boyatzis, *The competent manager: A model for effective performance*. John Wiley & Sons, 1982.

[22] R. Henderson, I. Cockburn, "Measuring competence? exploring firm effects in pharmaceutical research," *Strategic management journal*, vol. 15, no. S1, pp. 63–84, 1994.

[23] J. F. Lozano, A. Boni, J. Peris, A. Hueso, "Competencies in higher education: A critical analysis from the capabilities approach," *Journal of Philosophy of Education*, vol. 46, no. 1, pp. 132–147, 2012.

[24] V. Vathanophas, "Competency requirements for effective job performance in Thai public sector," *Contemporary management research*, vol. 3, no. 1, p. 45, 2007.

[25] J. Macnamara, "Competence, competencies and/or capabilities for public communication? a public sector study," *Asia Pacific Public Relations Journal*, vol. 19, pp. 16–40, 2018.

[26] P. Morgan, "The concept of capacity," *European Centre for Development Policy Management*, pp. 1–19, 2006.

[27] L. Vincent, "Differentiating competence, capability and capacity," *Innovating Perspectives*, vol. 16, no. 3, pp. 1–2, 2008.

[28] M. Z. Al-Taie, S. Kadry, A. I. Obasa, "Understanding expert finding systems: domains and techniques," *Social Network Analysis and Mining*, vol. 8, no. 1, p. 57, 2018.

[29] O. Husain, N. Salim, R. A. Alias, S. Abdelsalam, A. Hassan, "Expert finding systems: A systematic review," *Applied Sciences*, vol. 9, no. 20, p. 4250, 2019.

[30] I. Srba, M. Bielikova, "A comprehensive survey and classification of approaches for community question answering," *ACM Transactions on the Web (TWEB)*, vol. 10, no. 3, pp. 1–63, 2016.

[31] T. M. Connolly, E. A. Boyle, E. MacArthur, T. Hainey, J. M. Boyle, "A systematic literature review of empirical evidence on computer games and serious games," *Computers & Education*, vol. 59, no. 2, pp. 661 – 686, 2012, doi: https://doi.org/10.1016/j.compedu.2012.03.004.

[32] E. A. Boyle, T. Hainey, T. M. Connolly, G. Gray, J. Earp, M. Ott, T. Lim, M. Ninaus, C. Ribeiro, J. Pereira, "An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games," *Computers & Education*, vol. 94, pp. 178 – 192, 2016, doi: https://doi.org/10.1016/j.compedu.2015.11.003.

[33] T. Hainey, T. M. Connolly, E. A. Boyle, A. Wilson, A. Razak, "A systematic literature review of games- based learning empirical evidence in primary education," *Computers & Education*, vol. 102, pp. 202– 223, 2016.

[34] X. Wei, N. Saab, W. Admiraal, "Assessment of cognitive, behavioral, and affective learning outcomes in massive open online courses: A systematic literature review," *Computers & Education*, vol. 163, p. 104097, 2021, doi: https://doi.org/10.1016/j.compedu.2020.104097.

[35] M. J. Page, D. Moher, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, *et al.*, "PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews," *BMJ*, vol. 372, 2021.

[36] R. Huelin, I. Iheanacho, K. Payne, K. Sandman, "What's in a name? systematic and non-systematic literature reviews, and why the distinction matters," *The evidence*, pp. 34–37, 2015.

[37] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," in *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, EASE '14, New York, NY, USA, 2014, Association for Computing Machinery.

[38] J. Randolph, "A guide to writing the dissertation literature review," *Practical Assessment, Research, and Evaluation*, vol. 14, no. 1, p. 13, 2009.

[39] F. Riahi, Z. Zolaktaf, M. Shafiei, E. Milios, "Finding expert users in community question answering," in *Proceedings of the 21st International Conference on World Wide Web*, WWW '12 Companion, New York, NY, USA, 2012, p. 791–798, Association for Computing Machinery.

[40] D. Movshovitz-Attias, Y. Movshovitz-Attias, P. Steenkiste, C. Faloutsos, "Analysis of the reputation system and user contributions on a question answering website: StackOverflow," in *2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013)*, 2013, pp. 886–893, IEEE.

[41] Y. Xu, D. Zhou, S. Lawless, "Inferring your expertise from Twitter: Integrating sentiment and topic relatedness," in *2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, 2016, pp. 121–128, IEEE.

[42] L. Akritidis, D. Katsaros, P. Bozanis, "Identifying the productive and influential bloggers in a community," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 5, pp. 759–764, 2011.

[43] Y. Bae, H. Lee, "Sentiment analysis of Twitter audiences: Measuring the positive or negative influence of popular twitterers," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 12, pp. 2521–2535, 2012.

[44] C. Bigonha, T. N. Cardoso, M. M. Moro, M. A. Gonçalves, V. A. Almeida, "Sentiment-based influence detection on Twitter," *Journal of the Brazilian Computer Society*, vol. 18, no. 3, pp. 169–183, 2012.

[45] M. Cha, H. Haddadi, F. Benevenuto, P. K. Gummadi, *et al.*, "Measuring user influence in Twitter: The million follower fallacy," *Icwsm*, vol. 10, no. 10-17, p. 30, 2010.

[46] H. U. Khan, A. Daud, "Finding the top influential bloggers based on productivity and popularity features," *New Review of Hypermedia and Multimedia*, vol. 23, no. 3, pp. 189–206, 2017.

[47] S. Nagrecha, J. Z. Dillon, N. V. Chawla, "MOOC dropout prediction: Lessons learned from making pipelines interpretable," in *Proceedings of the 26th International Conference on World Wide Web Companion*, WWW '17 Companion, Republic and Canton of Geneva, CHE, 2017, p. 351–359, International World Wide Web Conferences Steering Committee.

[48] N. Dmoshinskaia, "Dropout prediction in MOOCs: using sentiment analysis of users' comments to predict engagement.," Master's thesis, University of Twente, 2016.

[49] S. Patil, K. Lee, "Detecting experts on Quora: by their activity, quality of answers, linguistic characteristics and temporal behaviors," *Social network analysis and mining*, vol. 6, no. 1, p. 5, 2016.

[50] D. Choi, J. Han, T. Chung, Y.-Y. Ahn, B.-G. Chun, T. T. Kwon, "Characterizing conversation patterns in Reddit: From the perspectives of content properties and user participation behaviors," in *Proceedings of the 2015 ACM on Conference on Online Social Networks*, COSN '15, New York, NY, USA, 2015, p. 233–243, Association for Computing Machinery.

[51] J. Yang, K. Tao, A. Bozzon, G.-J. Houben, "Sparrows and owls: Characterisation of expert behaviour in StackOverflow," in *User Modeling, Adaptation, and Personalization*, Cham, 2014, pp. 266–277, Springer, Springer International Publishing.

[52] E. Malherbe, M.-A. Aufaure, "Bridge the terminology gap between recruiters and candidates: A multilingual skills base built from social media and linked data," in *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2016, pp. 583–590, IEEE.

[53] G. Zhou, S. Lai, K. Liu, J. Zhao, "Topic-sensitive probabilistic model for expert finding in question answer communities," in *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, CIKM '12, New York, NY, USA, 2012, p. 1662–1666, Association for Computing Machinery.

[54] P. Jurczyk, E. Agichtein, "Discovering authorities in question answer communities by using link analysis," in *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, 2007, pp. 919–922.

[55] N. Raj, L. Dey, B. Gaonkar, "Expertise prediction for social network platforms to encourage knowledge sharing," in *Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology - Volume 01*, WI-IAT '11, USA, 2011, p. 380–383, IEEE Computer Society.

[56] A. Raikos, P. Waidyasekara, "How useful is YouTube in learning heart anatomy?," *Anatomical sciences education*, vol. 7, no. 1, pp. 12–18, 2014.

[57] A. Pal, A. Herdagdelen, S. Chatterji, S. Taank, Chakrabarti, "Discovery of topical authorities in Instagram," in *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, Republic and Canton of Geneva, CHE, 2016, p. 1203–1213, International World Wide Web Conferences Steering Committee.

[58] J. T. Hertel, N. M. Wessman-Enzinger, "Examining Pinterest as a curriculum resource for negative integers: An initial investigation," *Education Sciences*, vol. 7, no. 2, p. 45, 2017.

[59] A.-M. Popescu, K. Y. Kamath, J. Caverlee, "Mining potential domain expertise in Pinterest," in *UMAP workshops*, 2013.

[60] J. Oliveira, M. Viggiato, E. Figueiredo, "How well do you know this library? mining experts from source code analysis," in *Proceedings of the XVIII Brazilian Symposium on Software Quality*, SBQS'19, New York, NY, USA, 2019, p. 49–58, Association for Computing Machinery.

[61] R. Saxena, N. Pedanekar, "I know what you coded last summer: Mining candidate expertise from GitHub repositories," in *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17 Companion, New York, NY, USA, 2017, p. 299–302, Association for Computing Machinery.

[62] P. A. Martínez, M. J. Gómez, J. A. Ruipérez-Valiente, G. Martínez Pérez, Y. J. Kim, "Visualizing educational game data: A case study of visualizations to support teachers," in *Learning Analytics Summer Institute Spain 2020: Learning Analytics. Time for Adoption?*, Jun 2020, pp. 97–111, CEUR Workshop Proceedings.

[63] E. Harpstead, T. Zimmermann, N. Nagapan, J. J. Guajardo, R. Cooper, T. Solberg, D. Greenawalt, "What drives people: Creating engagement profiles of players from game log data," in *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '15, New York, NY, USA, 2015, p. 369–379, Association for Computing Machinery.

[64] A. J. Lesser, *Video game technology and learning in the music classroom.* PhD dissertation, Teachers College, Columbia University, 2019.

[65] P. Kantharaju, K. Alderfer, J. Zhu, B. Char, B. Smith, S. Ontañón, "Tracing player knowledge in a parallel programming educational game," 2019.

[66] F. Chen, Y. Cui, M.-W. Chu, "Utilizing game analytics to inform and validate digital game-based assessment with evidence-centered game design: A case study," *International Journal of Artificial Intelligence in Education*, vol. 30, no. 3, pp. 481–503, 2020.

[67] J. A. Ruipérez-Valiente, M. Gaydos, L. Rosenheck, Y. J. Kim, E. Klopfer, "Patterns of engagement in an educational massive multiplayer online game: A multidimensional view," *IEEE Transactions on Learning Technologies*, 2020.

[68] Y. J. Kim, J. A. Ruipérez-Valiente, "Data-driven game design: The case of difficulty in educational games," in *European Conference on Technology Enhanced Learning*, 2020, pp. 449–454, Springer.

[69] M.-T. Cheng, Y.-W. Lin, H.-C. She, "Learning through playing Virtual Age: Exploring the interactions among student concept learning, gaming performance, in-game behaviors, and the use of in-game characters," *Computers & Education*, vol. 86, pp. 18–29, 2015.

[70] J. Kang, M. Liu, W. Qu, "Using gameplay data to examine learning behavior patterns in a serious game," *Computers in Human Behavior*, vol. 72, pp. 757– 770, 2017.

[71] W. Westera, R. Nadolski, H. Hummel, "Serious gaming analytics: What students' log files tell us about gaming and learning," *International Journal of Serious Games*, vol. 1, Jun. 2014, doi: 10.17083/ijsg.v1i2.9.

[72] T. Daradoumis, R. Bassi, F. Xhafa, S. Caballé, "A review on massive e-learning (MOOC) design, delivery and assessment," in *2013 Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, 2013, pp. 208–213.

[73] S. Crossley, L. Paquette, M. Dascalu, D. S. McNamara, R. S. Baker, "Combining click-stream data with NLP tools to better understand MOOC completion," in *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, LAK '16, New York, NY, USA, 2016, p. 6–14, Association for Computing Machinery.

[74] R. Zhao, V. Li, H. Barbosa, G. Ghoshal, M. E. Hoque, "Semi-automated 8 collaborative online training module for improving communication skills," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 2, pp. 1–20, 2017.

[75] R. Reddick, "Using a Glicko-based algorithm to measure in-course learning," *International Educational Data Mining Society*, 2019.

[76] X. Wang, D. Yang, M. Wen, K. Koedinger, C. P. Rosé, "Investigating how student's cognitive behavior in MOOC discussion forums affect learning gains.," *International Educational Data Mining Society*, 2015.

[77] J. Kim, P. J. Guo, D. T. Seaton, P. Mitros, K. Z. Gajos, R. C. Miller, "Understanding in-video dropouts and interaction peaks in online lecture videos," in *Proceedings of the First ACM Conference on Learning @ Scale Conference*, L@S '14, New York, NY, USA, 2014, p. 31–40, Association for Computing Machinery.

[78] D. Huynh, L. Zuo, H. Iida, "Analyzing gamification of "Duolingo" with focus on its course structure," in *International Conference on Games and Learning Alliance*, 2016, pp. 268–277, Springer.

[79] S. Chootongchai, N. Songkram, "Design and development of seci and moodle online learning systems to enhance thinking and innovation skills for higher education learners," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 13, no. 03, pp. 154–172, 2018.

[80] B. Hightower, C. Rawl, M. Schutt, "Collaborations for delivering the library to students through WebCT," *Reference Services Review*, 2007.

[81] H. Tanaka, S. Sakti, G. Neubig, T. Toda, H. Negoro, H. Iwasaka, S. Nakamura, "Automated social skills trainer," in *Proceedings of the 20th International Conference on Intelligent User Interfaces*, IUI '15, New York, NY, USA, 2015, p. 17–27, Association for Computing Machinery.

[82] P. Pham, J. Wang, "Attentivelearner: Improving mobile MOOC learning via implicit heart rate tracking," in *Artificial Intelligence in Education*, Cham, 2015, pp. 367–376, Springer, Springer International Publishing.

[83] C. Yin, F. Okubo, A. Shimada, M. Oi, S. Hirokawa, H. Ogata, "Identifying and analyzing the learning behaviors of students using e-books," in *Doctoral Student Consortium (DSC) - Proceedings of the 23rd International Conference on Computers in Education, ICCE 2015*, 2015, pp. 118–120, Asia-Pacific Society for Computers in Education.

[84] A. Shimada, K. Mouri, H. Ogata, "Real-time learning analytics of e-book operation logs for on-site lecture support," in *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*, 2017, pp. 274–275, IEEE.

[85] D. M. Boyd, N. B. Ellison, "Social network sites: Definition, history, and scholarship," *Journal of computer-mediated communication*, vol. 13, no. 1, pp. 210–230, 2007.

[86] J. Pastor-Galindo, M. Zago, P. Nespoli, S. López Bernal, A. Huertas, M. Pérez, J. A. Ruipérez- Valiente, G. Martínez Pérez, F. Gómez Mármol, "Spotting political social bots in Twitter: A use case of the 2019 spanish general election," *IEEE Transactions on Network and Service Management*, vol. 17, pp. 2156– 2170, 10 2020, doi: 10.1109/TNSM.2020.3031573.

[87] E. Özdemir, "Promoting EFL learners" intercultural communication effectiveness: a focus on Facebook," *Computer Assisted Language Learning*, vol. 30, no. 6, pp. 510–528, 2017.

[88] W. Orawiwatnakul, S. Wichadee, "Achieving better learning performance through the discussion activity in Facebook," *Turkish Online Journal of Educational Technology-TOJET*, vol. 15, no. 3, pp. 1–8, 2016.

[89] X. Yan, J. Yang, M. Obukhov, L. Zhu, J. Bai, S. Wu, Q. He, "Social skill validation at LinkedIn," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '19, New York, NY, USA, 2019, p. 2943–2951, Association for Computing Machinery.

[90] B. Muros-Ruiz, Y. Aragón-Carretero, A. Bustos- Jiménez, "Youth's usage of leisure time with video games and social networks," *Comunicar: Revista Científica de Comunicación y Educación*, vol. 20, no. 40, pp. 31–39, 2013.

[91] M. Boukes, "Social network sites and acquiring current affairs knowledge: The impact of Twitter and Facebook usage on learning about the news," *Journal of Information Technology & Politics*, vol. 16, no. 1, pp. 36–51, 2019.

[92] T. P. Alloway, R. G. Alloway, "The impact of engagement with social networking sites (SNSs) on cognitive skills," *Computers in Human Behavior*, vol. 28, no. 5, pp. 1748–1754, 2012.

[93] M. Niemelä, T. Kärkkäinen, S. Äyrämö, M. Ronimus, U. Richardson, H. Lyytinen, "Game learning analytics for understanding reading skills in transparent writing system," *British Journal of Educational Technology*, 02 2020, doi: 10.1111/bjet.12916.

[94] A. I. Wang, A. Lieberoth, "The effect of points and audio on concentration, engagement, enjoyment, learning, motivation, and classroom dynamics using Kahoot!," in *European Conference on Games Based Learning*, vol. 20, 2016, Academic Conferences International Limited.

[95] M. Bouguessa, L. B. Romdhane, "Identifying authorities in online communities," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 6, no. 3, pp. 1–23, 2015.

[96] G. Guerrero, E. Sarchi, F. Tapia, "Predict the personality of Facebook profiles using automatic learning techniques and BFI test," in *New Knowledge in Information Systems and Technologies*, Cham, 2019, pp. 482–493, Springer International Publishing.

[97] R. Nielek, O. Jarczyk, K. Pawlak, L. Bukowski, R. Bartusiak, A. Wierzbicki, "Choose a job you love: predicting choices of GitHub developers," in *2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, 2016, pp. 200–207, IEEE.

[98] A. Cohen, U. Shimony, R. Nachmias, T. Soffer, "Active learners' characterization in MOOC forums and their generated knowledge," *British Journal of Educational Technology*, vol. 50, no. 1, pp. 177–198, 2019.

[99] J. Han, D. Choi, A.-Y. Choi, J. Choi, T. Chung, T. T. Kwon, J.-Y. Rha, C.-N. Chuah, "Sharing topics in Pinterest: Understanding content creation and diffusion behaviors," in *Proceedings of the 2015 ACM on Conference on Online Social Networks*, COSN '15, New York, NY, USA, 2015, p. 245–255, Association for Computing Machinery.

[100] F. Calefato, F. Lanubile, B. Vasilescu, "A large- scale, in-depth analysis of developers' personalities in the apache ecosystem," *Information and Software Technology*, vol. 114, pp. 1–20, 2019.

[101] J. Baumgartner, S. Zannettou, B. Keegan, M. Squire, J. Blackburn, "The Pushshift Reddit dataset," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 14, 2020, pp. 830–839.

[102] B. Vasilescu, V. Filkov, A. Serebrenik, "StackOverflow and GitHub: Associations between software development and crowdsourced knowledge," in *2013 International Conference on Social Computing*, 2013, pp. 188–195, IEEE.

[103] W. H. Lim, M. J. Carman, S.-M. J. Wong, "Estimating relative user expertise for content quality prediction on Reddit," in *Proceedings of the 28th ACM Conference on Hypertext and Social Media*, HT '17, New York, NY, USA, 2017, p. 55–64, Association for Computing Machinery.

[104] C. Mhamdi, M. Al-Emran, S. A. Salloum, *Text Mining and Analytics: A Case Study from News Channels Posts on Facebook*, pp. 399–415. Cham: Springer International Publishing, 2018.

[105] W. Xing, F. Gao, "Exploring the relationship between online discourse and commitment in Twitter professional learning communities," *Computers & Education*, vol. 126, pp. 388–398, 2018.

[106] T. Hecking, I.-A. Chounta, H. U. Hoppe, "Investigating social and semantic user roles in MOOC discussion forums," in *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, LAK '16, New York, NY, USA, 2016, p. 198–207, Association for Computing Machinery.

[107] A. Pal, S. Counts, "Identifying topical authorities in microblogs," in *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, WSDM '11, New York, NY, USA, 2011, p. 45–54, Association for Computing Machinery.

[108] Z.-J. Yang, J. Lin, Y.-S. Yang, "Identification of network behavioral characteristics of high-expertise users in interactive innovation: The case of forum autohome," *Asia Pacific Management Review*, 2020, doi: https://doi.org/10.1016/j.apmrv.2020.06.002.

[109] H. Zhu, H. Cao, H. Xiong, E. Chen, J. Tian, "Towards expert finding by leveraging relevant categories in authority ranking," in *Proceedings of the 20th ACM international conference on Information and knowledge management*, 2011, pp. 2221–2224.

[110] C. MACLEOD, "Evaluating student use of Duolingo, an online self—study platform," *JOURNAL OF ATOMI UNIVERSITY FACULTY OF LITERATURE*, no. 54, pp. A49–A67, 2019.

[111] S. Loewen, D. Crowther, D. R. Isbell, K. M. Kim, J. Maloney, Z. F. Miller, H. Rawal, "Mobile-assisted language learning: A Duolingo case study," *ReCALL*, vol. 31, no. 3, pp. 293–311, 2019.

[112] M. Liu, J. Lee, J. Kang, S. Liu, "What we can learn from the data: A multiple-case study examining behavior patterns by students with different characteristics in using a serious game," *Technology, Knowledge and Learning*, vol. 21, no. 1, pp. 33–57, 2016.

[113] Y. Sun, C. Xin, "Using Coursera clickstream data to improve online education for software engineering," in *Proceedings of the ACM Turing 50th Celebration Conference - China*, ACM TUR-C '17, New York, NY, USA, 2017, Association for Computing Machinery.

[114] L. Breslow, D. E. Pritchard, J. DeBoer, G. S. Stump, A. D. Ho, D. T. Seaton, "Studying learning in the worldwide classroom research into edX's first MOOC," *Research & Practice in Assessment*, vol. 8, pp. 13–25, 2013.

[115] G. Alexandron, J. A. Ruipérez-Valiente, Z. Chen, P. J. Muñoz-Merino, D. E. Pritchard, "Copying@Scale: Using harvesting accounts for collecting correct answers in a MOOC," *Computers & Education*, vol. 108, pp. 96–114, 2017.

[116] I. Naim, M. I. Tanveer, D. Gildea, M. E. Hoque, "Automated prediction and analysis of job interview performance: The role of what you say and how you say it," in *2015 11th IEEE international conference and workshops on automatic face and gesture recognition (FG)*, vol. 1, 2015, pp. 1–6, IEEE.

[117] S. Peng., K. Nagao., "Automatic evaluation of presenters' discussion performance based on their heart rate," in *Proceedings of the 10th International Conference on Computer Supported Education - Volume 1: CSEDU*, 2018, pp. 27–34, INSTICC, SciTePress.

[118] A. Shimada, Y. Taniguchi, F. Okubo, S. Konomi, H. Ogata, "Online change detection for monitoring individual student behavior via clickstream data on e-book system," in *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, LAK '18, New York, NY, USA, 2018, p. 446–450, Association for Computing Machinery.

[119] F. J. Gutierrez, J. Simmonds, N. Hitschfeld, C. Casanova, C. Sotomayor, V. Peña Araya, "Assessing software development skills among k-6 learners in a project-based workshop with scratch," in *Proceedings of the 40th International Conference on Software Engineering: Software Engineering Education and Training*, ICSE-SEET '18, New York, NY, USA, 2018, p. 98–107, Association for Computing Machinery.

[120] C. G. Brinton, M. Chiang, "MOOC performance prediction via clickstream data and social learning networks," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, 2015, pp. 2299–2307, IEEE.

[121] M. Bouguessa, B. Dumoulin, S. Wang, "Identifying authoritative actors in question-answering forums: The case of Yahoo! Answers," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '08, New York, NY, USA, 2008, p. 866–874, Association for Computing Machinery.

[122] W.-C. Kao, D.-R. Liu, S.-W. Wang, "Expert finding in question-answering websites: A novel hybrid approach," in *Proceedings of the 2010 ACM symposium on applied computing*, SAC '10, New York, NY, USA, 2010, p. 867–871, Association for Computing Machinery.

[123] A. Borodin, G. O. Roberts, J. S. Rosenthal, P. Tsaparas, "Link analysis ranking: algorithms, theory, and experiments," *ACM Transactions on Internet Technology (TOIT)*, vol. 5, no. 1, pp. 231–297, 2005.

[124] R. W. Herling, "Operational definitions of expertise and competence," *Advances in Developing Human Resources*, vol. 2, no. 1, pp. 8–21, 2000, doi: 10.1177/152342230000200103.

[125] A. Santos, M. Souza, J. Oliveira, E. Figueiredo, "Mining software repositories to identify library experts," in *Proceedings of the VII Brazilian Symposium on Software Components, Architectures, and Reuse*, SBCARS '18, New York, NY, USA, 2018, p. 83–91, Association for Computing Machinery.

[126] J. A. E. Montandon, L. L. Silva, M. T. Valente, "Identifying experts in software libraries and frameworks among GitHub users," in *Proceedings of the 16th International Conference on Mining Software Repositories*, MSR '19, 2019, p. 276–287, IEEE, IEEE Press.

[127] C. Hauff, G. Gousios, "Matching GitHub developer profiles to job advertisements," in *2015 IEEE/ACM 12th Working Conference on Mining Software Repositories*, 2015, pp. 362–366.

[128] M. S. Faisal, A. Daud, A. U. Akram, R. A. Abbasi, N. R. Aljohani, I. Mehmood, "Expert ranking techniques for online rated forums," *Computers in Human Behavior*, vol. 100, pp. 168 – 176, 2019, doi: https://doi.org/10.1016/j.chb.2018.06.013.

[129] D. van Dijk, M. Tsagkias, M. de Rijke, "Early detection of topical expertise in community question answering," in *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '15, New York, NY, USA, 2015, p. 995–998, Association for Computing Machinery.

[130] E. Constantinou, G. M. Kapitsaki, "Identifying developers' expertise in social coding platforms," in *2016 42nd Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, 2016, pp. 63–67.

[131] M. A. Vogel, *Leveraging information technology competencies and capabilities for a competitive advantage.* PhD dissertation, 2005.

[132] D.-R. Liu, Y.-H. Chen, W.-C. Kao, H.-W. Wang, "Integrating expert profile, reputation and link analysis for expert finding in question-answering websites," *Information processing & management*, vol. 49, no. 1, pp. 312–329, 2013.

[133] Z. Zhao, L. Zhang, X. He, W. Ng, "Expert finding for question answering via graph regularized matrix completion," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 4, pp. 993–1004, 2014.

[134] A. Omidvar, M. Garakani, H. R. Safarpour, "Context based user ranking in forums for expert finding using wordnet dictionary and social network analysis," *Information Technology and Management*, vol. 15, no. 1, pp. 51–63, 2014.

[135] A. Pal, F. M. Harper, J. A. Konstan, "Exploring question selection bias to identify experts and potential experts in community question answering," *ACM Transactions on Information Systems (TOIS)*, vol. 30, no. 2, pp. 1–28, 2012.

[136] V. V. Vydiswaran, M. Reddy, "Identifying peer experts in online health forums," *BMC medical informatics and decision making*, vol. 19, no. 3, p. 68, 2019.

[137] H. Zhu, E. Chen, H. Xiong, H. Cao, J. Tian, "Ranking user authority with relevant knowledge categories for expert finding," *World Wide Web*, vol. 17, no. 5, pp. 1081–1107, 2014.

[138] A. Abdaoui, J. Azé, S. Bringay, N. Grabar, P. Poncelet, "Expertise in french health forums," *Health informatics journal*, vol. 25, no. 1, pp. 17–26, 2019.

[139] C. G. Brinton, S. Buccapatnam, M. Chiang, H. V. Poor, "Mining MOOC clickstreams: On the relationship between learner behavior and performance," 2015.

[140] A. Bozkurt, B. Aydin, A. Taşkıran, Koral Gümüşoğlu, "Improving creative writing skills of EFL learners through microblogging," *The Online Journal of New Horizons in Education*, vol. 6, pp. 88–98, 06 2016.

[141] T. Gonulal, "The use of Instagram as a mobile-assisted language learning tool," *Contemporary Educational Psychology*, vol. 10, pp. 309–323, 07 2019, doi: 10.30935/cet.590108.

[142] R. Vesselinov, J. Grego, "Duolingo effectiveness study," *City University of New York, USA*, vol. 28, 2012.

[143] Z. Liu, G. Xu, T. Liu, W. Fu, Y. Qi, W. Ding, Y. Song, C. Guo, C. Kong, S. Yang, et al., "Dolphin: A spoken language proficiency assessment system for elementary education," Apr 2020, pp. 2641–2647, ACM.

[144] J. Dixon, C. Belnap, C. Albrecht, K. Lee, "The importance of soft skills," *Corporate Finance Review*, vol. 14, pp. 35–38, May 2010.

[145] B. Cimatti, "Definition, development, assessment of soft skills and their role for the quality of organizations and enterprises.," *International Journal for quality research*, vol. 10, no. 1, 2016.

[146] T. Strobach, P. A. Frensch, T. Schubert, "Video game practice optimizes executive control skills in dual- task and task switching situations," *Acta Psychologica*, vol. 140, no. 1, pp. 13–24, 2012.

[147] H.-S. Hsiao, C.-S. Chang, C.-Y. Lin, P.-M. Hu, "Development of children's creativity and manual skills within digital game-based learning environment," *Journal of Computer Assisted Learning*, vol. 30, no. 4, pp. 377–395, 2014.

[148] D. R. Woods, A. N. Hrymak, R. R. Marshall, P. E. Wood, C. M. Crowe, T. W. Hoffman, J. D. Wright, P. A. Taylor, K. A. Woodhouse, C. K. Bouchard, "Developing problem solving skills: The McMaster problem solving program," *Journal of Engineering Education*, vol. 86, no. 2, pp. 75–91, 1997.

[149] H.-C. Chu, C.-M. Hung, "Effects of the digital game-development approach on elementary school students' learning motivation, problem solving, and learning achievement," *International Journal of Distance Education Technologies (IJDET)*, vol. 13, no. 1, pp. 87–102, 2015.

[150] M. Scriven, R. Paul, "Defining critical thinking. the critical thinking community: foundation for critical thinking," 2007.

[151] A. Pal, R. Farzan, J. A. Konstan, R. E. Kraut, "Early detection of potential experts in question answering communities," in *International Conference on User Modeling, Adaptation, and Personalization*, 2011, pp. 231–242, Springer.

[152] C. Unger, D. Murthy, A. Acker, I. Arora, A. Chang, "Examining the evolution of mobile social payments in Venmo," in *International Conference on Social Media and Society*, SMSociety'20, New York, NY, USA, 2020, p. 101–110, Association for Computing Machinery.

[153] M. Yuan, L. Zhang, X.-Y. Li, H. Xiong, "Comprehensive and efficient data labeling via adaptive model scheduling," in *2020 IEEE 36th International Conference on Data Engineering (ICDE)*, 2020, pp. 1858–1861, IEEE.

[154] D. Papp, G. Szűcs, Z. Knoll, "Machine preparation for human labelling of hierarchical train sets by spectral clustering," in *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2019, pp. 157–162, IEEE.

[155] J. Kang, S. Liu, M. Liu, *Tracking Students' Activities in Serious Games*, pp. 125–137. Cham: Springer International Publishing, 2017.

[156] M. Oczak, K. Maschat, D. Berckmans, E. Vranken, J. Baumgartner, "Can an automated labelling method based on accelerometer data replace a human labeller?–postural profile of farrowing sows," *Computers and Electronics in Agriculture*, vol. 127, pp. 168–175, 2016.

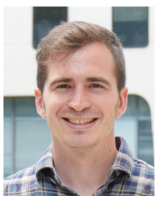
[157] C. A. Gomez-Uribe, N. Hunt, "The Netflix recommender system: Algorithms, business value, and innovation," *ACM Transactions on Management Information Systems (TMIS)*, vol. 6, no. 4, pp. 1–19, 2015.

[158] A. Emerson, N. Henderson, J. Rowe, W. Min, S. Lee, J. Minogue, J. Lester, "Investigating visitor engagement in interactive science museum exhibits with multimodal Bayesian hierarchical models," in *Artificial Intelligence in Education*, Cham, 2020, pp. 165– 176, Springer International Publishing.

[159] A. K. Cooper, S. Coetzee, "On the ethics of using publicly-available data," in *Responsible Design, Implementation and Use of Information and Communication Technology*, Cham, 2020, pp. 159–171, Springer International Publishing.
[160] C. Kuhlman, L. Jackson, R. Chunara, "No computation without representation: Avoiding data and algorithm biases through diversity," *arXiv preprint arXiv:2002.11836*, 2020.
[161] M. Feldman, S. A. Friedler, J. Moeller, C. Scheidegger, S. Venkatasubramanian, "Certifying and removing disparate impact," in *Proceedings of the 21st ACM SIGKDD international conference on knowledge discovery and data mining*, 2015, pp. 259–268.
[162] Z. Obermeyer, B. Powers, C. Vogeli, S. Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations," *Science*, vol. 366, no. 6464, pp. 447–453, 2019.

Sofia Strukova

She received the B.Sc. degree in computer science from Moscow Power Engineering Institute, Russia and M.Sc. degree in Big Data from the University of Murcia, Spain where she is currently pursuing the Ph.D. degree with the Department of Information and Communications Engineering.



José A. Ruipérez-Valiente

He received the B.Eng. degree in telecommunications from Universidad Católica de San Antonio de Murcia, and the M.Eng. degree in telecommunications and the M.Sc. and Ph.D. degrees in telematics from Universidad Carlos III of Madrid while conducting research with Institute IMDEA Networks in the area of learning analytics and educational data mining. He was a postdoctoral associate with MIT. He has received more than 20 academic/research awards and fellowships, has published more than 90 scientific publications in high impact venues, and participated in over 20 funded projects and contracts. He is currently an Assistant Professor at the University of Murcia. More info at https://webs.um.es/jruiperez

Félix Gómez Mármol

He is a researcher in the Department of Information and Communications Engineering at the University of Murcia, Spain. His research interests include cybersecurity, internet of things, machine learning and bio-inspired algorithms. He received a M.Sc. and Ph.D. in computer engineering from the University of Murcia. More info at: https://webs.um.es/felixgm