Hybrid Human-AI Driven Open Personalized Education

Von der Fakultät für Elektrotechnik und Informatik
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des akademischen Grades
Doktor rerum naturalium
(abgekürzt: Dr. rer. nat.)
genehmigte Dissertation

von Herrn
Mohammadreza Tavakoli, M.Sc.
geboren am
03.12.1989
in
Behbahan, Iran

2023

1. Referent:	Prof. Sören Auer
2. Referent:	Prof. Stefan T. Mol

Tag der Promotion:	06.06.2023

Abstract

Attaining those skills that match labor market demand is getting increasingly complicated as prerequisite knowledge, skills, and abilities are evolving dynamically through an uncontrollable and seemingly unpredictable process. Furthermore, people's interests in gaining knowledge pertaining to their personal life (e.g., hobbies and life-hacks) are also increasing dramatically in recent decades. In this situation, anticipating and addressing the learning needs are fundamental challenges to twenty-first century education. The need for such technologies has escalated due to the COVID-19 pandemic, where online education became a key player in all types of training programs. The burgeoning availability of data, not only on the demand side but also on the supply side (in the form of open/free educational resources) coupled with smart technologies, may provide a fertile ground for addressing this challenge. Therefore, this thesis aims to contribute to the literature about the utilization of (open and free-online) educational resources toward goal-driven personalized informal learning, by developing a novel Human-AI based system, called *eDoer*.

In this thesis, we discuss all the new knowledge that was created in order to complete the system development, which includes 1) prototype development and qualitative user validation, 2) decomposing the preliminary requirements into meaningful components, 3) implementation and validation of each component, and 4) a final requirement analysis followed by combining the implemented components in order develop and validate the planned system (eDoer).

All in all, our proposed system 1) derives the skill requirements for a wide range of occupations (as skills and jobs are typical goals in informal learning) through an analysis of online job vacancy announcements, 2) decomposes skills into learning topics, 3) collects a variety of open/free online educational resources that address those topics, 4) checks the quality of those resources and topic relevance using our developed intelligent prediction models, 5) helps learners to set their learning goals, 6) recommends personalized learning pathways and learning content based on individual learning goals, and 7) provides assessment services for learners to monitor their progress towards their desired learning objectives. Accordingly, we created a learning dashboard focusing on three *Data Science* related jobs and conducted an initial validation of eDoer through a randomized experiment. Controlling for the effects of prior knowledge as assessed by the pretest, the randomized experiment provided tentative support for the hypothesis that learners who engaged with personal eDoer recommendations attain higher scores on the posttest than those who did not. The hypothesis that learners who received personalized content in terms of format, length, level of detail, and content type, would achieve higher scores than those receiving non-personalized content was not supported as a statistically significant result.

Keywords: Personalized Education, Open Education, Artificial Intelligence

Zusammenfassung

Der Erwerb von Qualifikationen, die der Nachfrage auf dem Arbeitsmarkt entsprechen, wird immer komplizierter, da sich die erforderlichen Kenntnisse, Fähigkeiten und Fertigkeiten in einem unkontrollierbaren und scheinbar unvorhersehbaren Prozess dynamisch entwickeln. Darüber hinaus hat in den letzten Jahrzehnten auch das Interesse der Menschen am Erwerb von Wissen für ihr persönliches Leben (z. B. Hobbys und Life-Hacks) dramatisch zugenommen. In dieser Situation sind die Antizipation und das Eingehen auf die Lernbedürfnisse grundlegende Herausforderungen für die Bildung des 21. Jahrhunderts. Der Bedarf an solchen Technologien wurde durch die COVID-19-Pandemie noch verstärkt, bei der die Online-Bildung eine Schlüsselrolle in allen Arten von Schulungsprogrammen spielte. Die wachsende Verfügbarkeit von Daten, nicht nur auf der Nachfrageseite, sondern auch auf der Angebotsseite (in Form von offenen/freien Bildungsressourcen) in Verbindung mit intelligenten Technologien, kann einen fruchtbaren Boden für die Bewältigung dieser Herausforderung bieten. Daher zielt diese Arbeit darauf ab, einen Beitrag zur Literatur über die Nutzung von (offenen und kostenlosen Online-) Bildungsressourcen für zielgerichtetes personalisiertes informelles Lernen zu leisten, indem sie ein neuartiges Mensch-KI-basiertes System namens *eDoer* entwickelt.

In dieser Arbeit wird das gesamte neue Wissen diskutiert, das zur Fertigstellung der Systementwicklung geschaffen wurde. Dazu gehören 1. die Herausarbeitung eines Konzeptes, 2. die Entwicklung eines Prototyps und die qualitative Nutzervalidierung, 3. die Zerlegung der vorläufigen Anforderungen in sinnvolle Komponenten, 4. die Implementierung und Validierung der einzelnen Komponenten und 5. eine abschließende Anforderungsanalyse, gefolgt von der Kombination der implementierten Komponenten, um das geplante System (eDoer) zu entwickeln und zu validieren.

Alles in allem macht unser unser vorgeschlagenes System folgendes: 1. Es leitet die Qualifikationsanforderungen für eine breite Palette von Berufen (da Qualifikationen und Arbeitsplätze typische Ziele des informellen Lernens sind) durch eine Analyse von Online-Stellenausschreibungen ab; 2. zerlegt die Qualifikationen in Lernthemen; 3. sammelt eine Vielzahl von offenen/freien Online-Bildungsressourcen, die diese Themen ansprechen; 4. prüft die Qualität dieser Ressourcen und die Relevanz der Themen mit Hilfe der von uns entwickelten intelligenten Vorhersagemodelle; 5. hilft den Lernenden bei der Festlegung ihrer Lernziele; 6. empfiehlt personalisierte Lernpfade und Lerninhalte auf der Grundlage der individuellen Lernziele; und 7. bietet Bewertungsdienste für die Lernenden, um ihre Fortschritte auf dem Weg zu den gewünschten Lernzielen zu überwachen. Dementsprechend haben wir ein Lern-Dashboard erstellt, das sich auf drei Berufe im Bereich der Datenwissenschaften konzentriert, und eine erste Validierung von eDoer durch ein randomisiertes Experiment durchgeführt. Das randomisierte Experiment, bei dem die Auswirkungen des durch den Vortest ermittelten Vorwissens kontrolliert wurden, lieferte eine vorläufige Unterstützung für die Hypothese, dass Lernende, die sich mit persönlichen eDoer-Empfehlungen beschäftigten, im Posttest bessere Ergebnisse erzielten als diejenigen, die dies nicht taten. Die Hypothese, dass Lernende, die personalisierte Inhalte in Bezug auf Format, Länge, Detaillierungsgrad und Art des Inhalts erhielten,

höhere Punktzahlen erzielen würden als diejenigen, die nicht personalisierte Inhalte erhielten, wurde nicht als statistisch signifikantes Ergebnis unterstützt.
Keywords: Personalisierte Bildung, offene Bildung, künstliche Intelligenz

Preface

When I started applying for Ph.D. positions in 2018, I was only thinking of really top-class universities in Europe. Suddenly, I found a position, not from the institutes on my list. I found the topic extremely interesting. It was about education. I always believe education is one of the most important pillars, if not the most important one, of saving the world from wars, environmental/economical issues, etc. Unfortunately, the position deadline was already over, but I gave it a try. In the end, I was selected to work with the team for 18 months. I had a lot of stress when I started working on the project because I thought I would have only 18 months to show my skills. However, the first time I visited my supervisors (Prof. Dr. Sören Auer and Dr. Gábor Kismihók), I found out that there was no need for stress. They were nice, friendly, and supportive in a way that I had the freedom to choose my way besides receiving their productive feedback. So, I would like to thank Sören and Gábor.

My direct supervisor, Gábor and I started working on the project in 2019, and after only a few months, I became sure that the idea would not be just my thesis project. We were developing a sustainable system that had the capability to improve the education level for many people around the world. Gábor's outstanding attitude toward leading our team not only taught me many skills, but also made me sure that the project would proceed in the right direction. Slowly, we built a team working on the system. Now, we are not just teammates; we are close friends with each other. At least, I feel I work with my close friends Abdolali Faraji and Mohammadreza Molavi on the project which we have many hopes for and is part of our life. So, I would like to thank them for this great team. Moreover, I want to thank Dr. Stafan T. Mol, because of his amazing support and feedback. He was in a way, my third supervisor. I learned a lot from him not only about my thesis, but also about how to be a well-structured honest researcher.

Finally, I want to thank my family who is clearly key people in my life. I start with my wife (Sahar Rabet) who has been my first go-to person for any problem I face as she always provides me with effective solutions. I believe I have the best parents in the world; they have always been a great support through all stages of my life, childhood, education, etc. I also want to thank my great sisters Farima, Farida, and Maryam who have had the greatest impact on my life as they are my role model for any decision I make.

Contents

CHAPTER 1

Introduction

With the clock ticking on the United Nations' Sustainable Development Goal pertaining to quality education, the time is ripe to develop cost-effective, scalable, and sustainable means to match the exponentially growing array of open educational resources to (the needs of) individual learners, regardless of socio-economic status and/or demographic background. Indeed, where top-quality educational resources were once solely accessible to the privileged few, the growing trend of opening up such resources, together with related technological developments has created opportunities to distribute and disseminate such educational resources more equitably, inclusively, and effectively, to all those who seek them.

In this chapter, first, we discuss the motivations of this thesis. Afterward, we explain the challenges we needed to handle and illustrate our research questions toward proposing our solution. Finally, we explain the thesis overview including its structure, contributions, and list of publications.

1.1 Motivation

Although altruism and the "feel-good factor" have been identified as some of the main drivers of the movement to open up educational resources [1], to date the word *open* has remained more of a legal designation, than a harnessed potential. There are benefits associated with tapping into the vast array of Free-Online and Open Educational Resources (hereafter called OERs) that go beyond just making them accessible to people who may otherwise not be able to access education. First, in light of the burgeoning amount of publicly available textual data [2], there are opportunities for more explicitly mapping educational content to the demands of the labor market (as an important basis for learners' goals), therewith enhancing learners' motivation, learning effectiveness, and employability. Indeed, to date, efforts at personalizing educational content to learners are often backward-looking (i.e. where learners came from) as opposed to forward-looking (where they are going). Second, greater and greater demands are being placed on teachers, not only in terms of the ICT (Information and Communications Technology) heavy teaching methods they need to master, but also in terms of increasing student numbers and courses they may have to teach [3]. As we shall illustrate later, the ability of students to identify their needs and be recommended OERs based on where they stand and where they are going may complement traditional courses and may ultimately serve to make teachers' workloads more manageable. Third, many educational curricula crush students' self-directed learning, proactivity, sense of control, and autonomy by dictating what is to be learned and when it is to be learned, without providing learners with a sense of the bigger picture, or why they are learning what they are having to learn. The information asymmetry that this entails, means that all too often students are just passive receivers of education, as opposed to them taking guided decisions and expending motivated effort toward shaping their own future. It is against this backdrop that we started working on designing and constructing a vehicle that can connect learners to the educational contents that they seek and/or need regardless of their geographic location, demographic characteristics, and/or formal educational qualifications.

On the other hand, recent decades have seen educational environments changing dramatically in response to the increasing demand for online personalized learning [4]. There is a growing need for online personalized educational services because of 1) the rapid evolution in both the quantity and quality of skills demand [5–7], 2) the gap between knowledge (and skills) that job markets require and the training that formal educational programs offer [8–11], 3) the global challenges for work and education due to the emergence of the COVID-19 pandemic [12], and 4) increasing interest in gaining knowledge pertaining to their personal interests (e.g., hobbies, "Do It Yourself" (DIY), and life-hacks) [13, 14].

According to the above-mentioned reasons, we are facing exponential growth in educational services (mainly offering OERs) that are being produced and disseminated on an unprecedented scale, and published in different contexts (e.g. location, language, discipline, expertise level, and format) [15, 16]. However, the heterogeneity and (lack of) targeted distribution of these educational contents leads to a number of problems for learners that limit their usefulness. First, learners may not understand which components they need to learn to fulfill skill (or knowledge) requirements [17]. Second, even if they knew what it was that they were seeking to learn, learners are unlikely to be able to distinguish between high-quality and low-quality educational resources. In sum, confronted with an abundance of learning materials, learners may be overwhelmed and will unlikely be able to plot and follow their own effective learning path without directional guidance in the form of personalized educational recommendations. Due to these obstacles, many learners around the world prefer to learn from well-known developed courses (e.g., on *Coursera*). However, this is against personalization, as all the learners have to choose and pass the same materials, disregarding their preferences, background, goals, etc.

1.2 Challenges

According to the above-mentioned issues, educational services should be tailored for offering personalized, dynamic learning experiences. In this respect, open education becomes a key facilitator in many areas, including personal skills development [18]. OERs are also gaining popularity as content sources for open education [19]. Major free educational resource repositories have large amounts of regularly updated learning content in a wide range of content areas. Therefore it is surprising that despite their growing capacity, these platforms still underperform when offering personalized learning services. As an example, users must consult and search through several repositories (with different interfaces) manually in order to find appropriate learning content.

Due to the lack of high-quality personalized services like search and recommendation, the popularity of free-online/open educational services has been limited in most user groups (typically educators or lifelong learners) [19–23]. Analyzing the previous efforts to build such personalized educational systems by drawing on the rapidly growing amount of OERs [24, 25] revealed the lack of high-quality OER metadata, and effective quality control processes [26]. These issues seriously curtail the accessibility of OERs, and consequently deployment of high-quality, open personalized educational services.

At the same time, to help learners be up-to-date about the knowledge they require to achieve, a number of taxonomies (e.g. ESCO and O*NET) exist that may be leveraged to provide information about occupations, skills, and knowledge. However, most of these taxonomies are updated through a largely manual process, meaning they are time-consuming and expensive to construct, and also susceptible to being outdated [27]. Alternatively, AI-based methods (e.g., text-based) algorithms can be developed to extract those topics that are on the one hand manifest in corpora of job vacancies, and on the other, covered by existing educational materials in an effort to help learners to build their learning path [15–17, 28]. However, these AI-based models have their own disadvantages like 1. not being 100% accurate, and 2. not being generalizable for different knowledge areas [15].

1.3 Research Questions

In light of the above-mentioned points, open learning systems which support personalized education should be implemented. These systems need to offer up-to-date high-quality curricula in order to help learners proceed toward their goals (e.g., achieving skills required for their current/desire job). Therefore, the main research questions of this thesis are:

- *RQ1.* What are the key learner requirements for a goal driven (labor market based) personalized educational system?

- *RQ2.* How property analysis of educational contents can help in building educational recom-

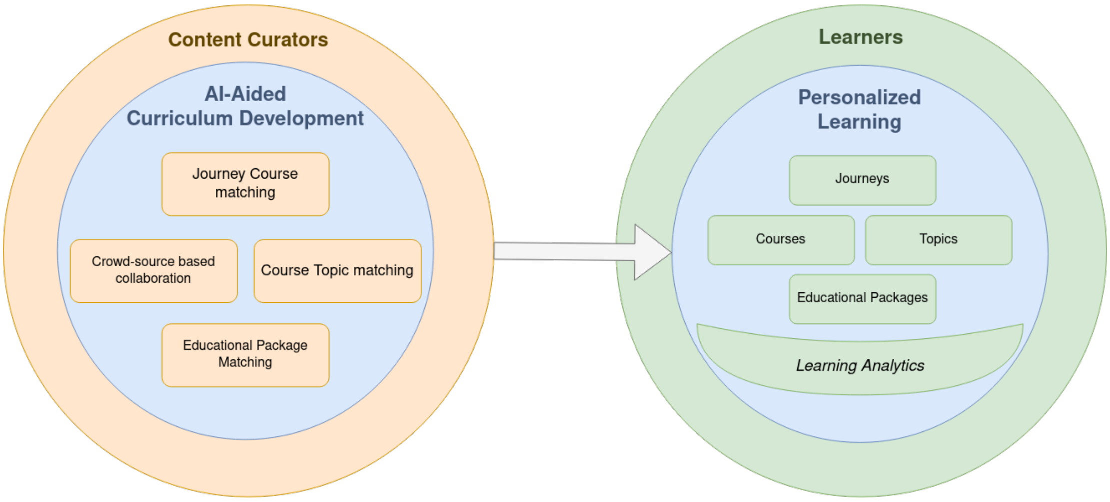


Figure 1.1: Our proposed system's main parts: curriculum development (left) and personalized learning dashboard (right)

mender systems?

- *RQ3.* How to build automatic models which can accurately predict the quality of OERs using their metadata and/or content?
- *RQ4.* How to extract learning topics from educational resources, and consequently, decompose knowledge-areas/skills into meaningful learning components?
- *RQ5.* How can we empower educational service providers to create, validate and maintain up-to-date curricula by using AI methods and crowdsourcing techniques?
- *RQ6.* How can we combine the answers to the above-mentioned questions to develop a personalized open educational system for learners?

1.4 Thesis Overview

To build our solution that addresses the above-mentioned research questions, we have developed an open, community-based learning and curriculum development method, which consists of two main components, as shown in Figure 1.1:

1. **Intelligent and scalable curriculum development.** The curriculum development component helps content curators to build, update, and maintain a wide range of curricula with the help of AI-generated recommendations. The curricula can be broken down into a four-level structure. On the highest level, one can define (1) learning ***journeys*** (or high-level goals), which consist

of relevant (2) ***courses*** (or skills) covered by (3) ***learning topics*** that target some learning objective. Learning topics, finally, are explained by (4) ***educational packages***, which consist of one or multiple (O)ERs. Our AI algorithms aid content curators with recommendations at the intersections of this four-level structure. First, our AI recommends potential courses for learning journeys (journey-course matching). This recommendation is followed by assistance in defining those learning topics that cover a course (course-topic matching). Lastly, our AI proposes high-quality educational packages for learning topics (topic-package matching). Moreover, crowd-sourcing techniques have been implemented to support collaboration among content curators.

2. **Personalized learning dashboard.** This dashboard empowers learners to make use of the above-mentioned quality-controlled four-level knowledge base. Through this dashboard learners (1) get information about available journeys, courses, and topics, which fit their interests and learning needs. They can (2) set their learning goals, and (3) receive personalized learning pathway recommendations, based on their context (preferences) and the available content. Finally (4) learners can monitor/assess their progress toward their learning objectives.

1.4.1 Thesis Structure

The structure of this thesis is as follows: The background knowledge needed for reading this thesis is provided in Chapter 2. Subsequently, we cover the state-of-the-art of the related areas to this thesis project in Chapter 3. Afterward, we explain our first step toward proposing our personalized goal-driven educational system in Chapter 4. After that, we discuss our findings on the process of analyzing the educational resources toward creating educational recommendations in Chapter 5. Our methods on metadata-based quality scoring and evaluation of OERs are explained in Chapter 6. In Chapter 7, we discuss our approach for decomposing knowledge/skills to meaningful learning topics and how to recognize those topics from educational resources. The idea on how to make our solution scalable by creating a Human-AI based curriculum development system is discussed in Chapter 8. Finally, we illustrate the final step of building our personalized goal-driven learning system using the developed components in Chapter 9. Moreover, we cover other investigations which are done using the outcome of this thesis, but not directly related to our target research questions, as horizontal aspects of this thesis in Chapter 10.

1.4.2 List of Publications and Contributions

Table 1.1 shows the publications, their contributions toward our research questions, and the chapter in which we explain their methods. Moreover, the publications which are not part of the main thesis objectives, are covered in Chapter 10.

Table 1.1: Outcomes of this thesis project

Target Research Question: Contribution	Explained in	Publications
RQ1: Building the prototype of labor market based goal driven personalized learning to get preliminary feedback from stakeholders (i.e. learners and instructors, and managers).	Chapter 4	**1.** Tavakoli, M., Mol, S.T. and Kismihók, G., 2020. Labour market information driven, personalized, OER recommendation system for lifelong learners. In the 12th Computer Supported Education Conference (CSEDU). **2.** Tavakoli, M., Faraji, A., Mol, S.T. and Kismihók, G., 2020. OER recommendations to support career development. In 2020 IEEE Frontiers in Education Conference (FIE).
RQ2 & RQ3: Developing a recommender system that analyzes educational videos' properties to offer the most fit resources to learners.	Chapter 5	**3.** Tavakoli, M., Hakimov, S., Ewerth, R. and Kismihok, G., 2020. A recommender system for open educational videos based on skill requirements. In 2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT).
RQ3: Creating metadata-based quality scoring and prediction models for educational resources.	Chapter 6	**4.** Tavakoli, M., Elias, M., Kismihók, G. and Auer, S., 2021. Metadata analysis of open educational resources. In LAK21: 11th International Learning Analytics and Knowledge Conference. **5.** Tavakoli, M., Elias, M., Kismihok, G. and Auer, S., 2020. Quality prediction of open educational resources a metadata-based approach. In 2020 IEEE 20th international conference on advanced learning technologies (ICALT).
RQ4: Building a topic extraction model which 1. recognizes the topics that should be covered for in each knowledge-area/skill, and 2. extract the covered learning topics from an educational resources	Chapter 7	**6.** Molavi, M., Tavakoli, M. and Kismihók, G., 2020. Extracting Topics from Open Educational Resources. In European Conference on Technology Enhanced Learning (EC-TEL).
RQ5: Creating a curricula development system based on AI methods and crowdsourcing techniques.	Chapter 8	**7.** Tavakoli, M., Faraji, A., Molavi, M., Mol, S.T. and Kismihók, G., 2022. Hybrid Human-AI Curriculum Development for Personalised Informal Learning Environments. In LAK22: 12th International Learning Analytics and Knowledge Conference.
RQ6: Building and evaluating the personalized educational system that helps learners 1. set their learning goals, 2. receive personalized learning path, and 3 assess their knowledge.	Chapter 9	**8.** Tavakoli, M., Faraji, A., Vrolijk, J., Molavi, M., Mol, S.T. and Kismihók, G., 2022. An AI-based open recommender system for personalized labor market driven education. Advanced Engineering Informatics Journal.

CHAPTER 2

Background

In this chapter, we provide a general insight into various concepts that are essential in this thesis. Since building a technology-based educational product is the main goal of this project, at first, we cover the *Educational Technology* concept in Section 2.1. After that, in Section 2.2, we discuss the *Open Education* concept as being free and/or open is a key part of our product. Afterward, as a recommendation service, we introduce the *Recommnder Systems* in Section 2.3. Also, developing a goal-driven product leads us to cover the *Goal Setting* concept (Section 2.4). Finally, we explain the most important methodologies that are used in the area of educational technologies, including *Artificial Intelligence* and *Crowdsourcing*, in Section 2.5.

2.1 Educational Technology and Online learning [29]

Based on the *Association for Educational Communications and Technology (AECT)*, *Educational Technology* is "the study and ethical practice of facilitating learning and improving performance by creating, using and managing appropriate technological processes and resources". The goal of educational technology is to facilitate the process of learning and improve the learners' performance by integrating technology into education in a positive manner. Therefore, in educational technology, we are able to make use of different disciplines such as computer science, communication, education, psychology, and sociology.

Online education, as one of the main mediums for educational technology, has received a lot of attention from learners in recent decades. *Western Behavioral Sciences Institute in La Jolla, California* helped online education to be emerged in 1982 by opening its *School of Management and Strategic Studies*. In 2008, the *Council of Europe* announced that e-learning has the potential to improve education together with equality across the EU. 2015 was the first year that private nonprofit organizations overtook for-profits in the number of registered online learners. However, public universities still enrolled the highest number of online students. Online education is not only successful in establishing a method for distance learning, but also it offers a powerful approach to make this type of learning more efficient by connecting instructors and learners in an online environment. On the other hand, learners growing up in this digital age tend to perform their everyday tasks online which has made online learning a trend in the 21st century. From these pieces of information, it can be concluded that the number of learners using online education is on a steady increase.

Furthermore, with the start of the COVID-19 pandemic, many schools and universities across the world were forced to close, which led to a situation that more and more learners and instructors migrated from traditional classrooms to online learning environments. The huge change which was forced by the pandemic resulted in a situation where by the of 2021, 189 million learners enrolled in courses on *Coursera*, as a well-known educational platform, while this number was 76 million in 2019 [30]. This showed that, in only two years, the number of online learners almost got tripled on this platform.

2.2 Open Education [31]

Open education, as an educational movement founded on openness, broadens access to the learning and training traditionally offered through formal education systems. Researches have shown that scientific/economic progress, technology, pedagogy, and related socioeconomic developments have a symbiotic relationship with open and distance education.

Open education has origins, especially in higher education, to the 17th century by *John Amos Comenius*, who proposed open access to education as a core goal. The postwar era of the 1960s and 1970s faced a "world-wide crisis in education" as there was a huge gap between what education systems offered and the demand for higher education in an era of scientific and economic prosperity. This phenomenon emphasized the need for an approach that would support a much larger and diversified group of lifelong learners which led to the establishment of open and distance education systems globally.

There are various examples of institutional practices in line with open education which has been decreasing the education entry barriers, for example, eliminating academic admission requirements. The Open University in Britain, Athabasca University in Canada, and the Open University of Catalonia, in Spain, are among the institutes which offer this type of education. Furthermore, Open Educational Resources (OER), Massive Open Online Courses (MOOCs), and OpenCourseWare are among the most recent and visible approaches to open education. However, there can be costs for acquiring certificates.

Since open education usually occurs at different times and places for most learners across the world, it is usually served through online learning platforms. Subsequently, various educational technologies need to be tailored to enhance open-based learning programs. These technologies can be personalized search/recommendation services, assessment, and question and answering. Although there is a high potential for implementing open education systems, there are still barriers regarding the development of such systems. These barriers include, but not limited to, lack of:

- administrative oversight
- automatic quality assurance services for educational resources
- personalized search and recommendation services
- automatic metadata extraction services

The above-mentioned issues have made the process of up-to-date open curriculum development [32] and pathway creation difficult. The reason is that there are millions of educational resources online, and using these resources manually without automatic services is getting more and more complicated. Therefore, there have been efforts on implementing curricula development systems which try to support authors/instructors with automatic intelligent services. These services consist of services on educational requirement analysis, objective design, structure creation, content creation/selecting, and curriculum review.

2.3 Recommender Systems [33]

A *Recommender System* is an information filtering system that provides suggestions for items that are most relevant to a particular user. Recommender systems are particularly useful when a user needs to choose the item(s) among a huge number of potential options such as buying a product, selecting music, or filtering news to read.

Besides recommending the most accurate items to users, there are a number of factors that are considered in recommender systems:

- ***Diversity.*** In some use-cases, users prefer to receive recommendations with a higher intra-list diversity (e.g., movies from different genres).
- ***Privacy.*** As creating profiles is a common technique in building recommender systems, privacy standards should be followed to prevent users' sensitive information from being revealed. This issue sometimes leads to a situation where there are trade-offs between personalization and privacy issues.
- ***User demographics.*** Researches have shown that user demographics have effects on the users' satisfaction with their received recommendations. For instance, users' geographical location, age, and gender can be considered in order to generate more useful recommendations.

- ***Serendipity.*** Serendipity shows "how surprising a recommendation is". For instance, recommending a horror movie to a user who is interested in horror movies is accurate, but as it is an obvious recommendation, it has not high serendipity. However, recommending another genre (e.g., Sci-fi) to the user may result in less accurate but with high serendipity items which can lead to 1. more interesting items, and 2. improved recommendation algorithms as users' feedback/opinion can be retrieved on the new areas.

- ***Explainability.*** Users should be able to trust the recommender systems. One well-known way to achieve this trust is by explaining how and/or why a set of recommendations are generated.

2.4 Goal setting [34]

Goal setting is developing an action plan which motivates and guides a person/group toward a goal. Goal setting is one of the key components of personal-development and management literature. The well-known studies by *Edwin A. Locke* and his colleagues have shown difficult specific goals result in significantly higher performance than easy goals, no goals, or even the setting of an abstract goal such as asking people to do their best.

Setting goals can affect outcomes in different ways like:

- ***Choice.*** Goals may help people to narrow their attention and focus on the path toward their goals.

- ***Effort.*** Goals may encourage people to put more effort on their goal-related tasks.

- ***Goal commitment.*** People who have goals may perform better as they are committed to reaching certain goals.

- ***Feedback.*** Providing feedback for goal setters is one of the key features to have successful goals. Feedback cannot be provided without goals, and in the same way, goals can not be reached without providing feedback. Therefore, goals and feedback are highly interrelated.

In recent decades, there have been many researches on the effects of goal-setting in educational environments which show that goal setting can improve the learning process. These researches have also shown that in online learning, students who have a better understanding of the tasks, set more detailed goals in general which leads to higher performance.

2.5 State-of-the-art Methologies

In this section, we cover the common methodologies in the area of education which are used in this thesis. We start with *Artificial Intelligence* (*AI*), followed by the *Crowdsourcing* techniques.

2.5.1 Artificial Intelligence [35]

Artificial intelligence (AI) is intelligence provided by machines, as opposed to natural intelligence, has been defined as the field of study that refers to any system which takes actions in order to maximize the benefit of users by receiving data from its environment. There are a lot of applications derived from AI such as search engines (e.g., *Google*), human speech analysis (e.g., Siri and Alexa), self-driving cars (e.g., Tesla), text mining (e.g., sentiment analysis of the Amazon product reviews), game systems (e.g., chess computers).

Artificial intelligence was founded as an academic discipline in 1956, and there have been various experiences (e.g., waves of optimism, followed by disappointment and the loss of funding) since that time. In the 21st century, mathematical-statistical methods have been helping Machine Learning (as a key component of AI) to become highly successful in solving the target problems in the business industry, medical field, etc.

Traditionally, there are two main Machine Learning method categories:

- ***Supervised learning.*** These methods require labels for the input data by humans, and include two main varieties: 1. *Classification*, and 2. *Regression*. Classification, like spam detection, is applied to determine what label a new data belongs to by using a number of example data from different labels and learning to classify. Regression, like predicting tomorrow's temperature, is the attempt to build a function that maps inputs into continuous outputs.

- ***Unsupervised learning.*** In this category, methods find patterns in a stream of input without any label. For instance, we can find existing clusters in employees' salaries.

2.5.2 Crowdsourcing [36]

The word *Crowdsourcing* consists of "crowd" and "outsourcing". It often involves digital platforms which divide the work between participants to achieve a cumulative result. Crowdsourcing offers several advantages such as improved costs, speed, quality, flexibility, scalability, and diversity. As an example, *Wikipedia* as a Non-profit organization has used crowdsourcing to develop common goods. As another example, *Amazon Mechanical Turk*, as a commercial platform, matches tasks submitted by requesters to workers who perform them.

Jeff Howe and Mark Robinson introduced the word crowdsourcing in 2006 in order to describe how businesses were using the internet to "outsource work to the crowd", which quickly led to the portmanteau "crowdsourcing". Crowdsourcing can either be taken place as an explicit or an implicit way. Explicit crowdsourcing lets users work directly on specific tasks. While implicit crowdsourcing means that users solve a problem as a side effect of the task(s) they are doing.

Difficulty-wise, there can be a spectrum of crowdsourcing activities that can be done by people with different expertise levels, from amateurs to experts. Also, reward-wise, participants are sometimes compensated monetarily with prizes, and in some cases, rewards can be achieved through kudos or intellectual satisfaction.

CHAPTER 3

Related Work

In this chapter, we explain the related researches to this thesis project. To do this, we held literature reviews on each of our research questions. Accordingly, all of the related areas (research questions) are covered in the following sections. At first, we describe the existing conceptual challenges toward building a goal (labor market) driven open personalized education (relates to *RQ1*, *RQ2*, and *RQ3*) in Section 3.1. As the next target area, we had to focus on the metadata analysis of educational resources (relates to *RQ3*). The outcome of this investigation is explained in Section 3.2. In order to create an educational path, we had to automatically analyze the existing learning topics in educational resources (relates to *RQ4*). The state-of-the-art on this subject is described in Section 3.3. Finally, in order to build a dynamic up-to-date educational system, we decided to investigate the area of automatic curriculum development (relates to *RQ5*), which is reported in Section 3.4.

In each section, at first, we discuss the importance of the target area. After that, we categorize the existing researches on the target area based on their methods. Afterward, we cover each category by explaining its containing research papers. At the end of each section, we briefly summarize the mentioned researches, and describe our findings from the covered area. Also, we conclude the results of our literature review investigation in Section 3.5.

3.1 Labor Market Driven Personalized OER Recommendation

In this part, we cover the area of building personalized OER recommendations toward labor market driven education. We cover this part as the following two problems: 1. how to dynamically match between jobs with their required skills, and 2. how to recommend the most relevant OER to learners?

3.1.1 Matching between Jobs and Skills

Having access to reliable labor market information on skills and jobs is not easy. Currently, only several governments or inter-governmental organizations (the most prominent actors are the US Government, European Commission, or Singapore) attempt to build skill inventories and occupational taxonomies (such as ESCO, ISCO, or O*NET). Although these taxonomy-building efforts have created a stable basis for basic skill analytics (inter-skill relationships, high-level matching to competencies and occupations), most of these resources are created and maintained by human experts in several time-consuming steps, which makes them expensive and also susceptible to out-dating [27]. It is therefore not surprising that more and more commercial and research attempts target new ways to obtain real-time labor market information about skills, using and analyzing alternative data sets like job vacancy announcement text, resume text, or social media data. These attempts can be clustered into the following three main categories:

Semantic-based Methods. This approach builds on ontologies to reveal and organize components of jobs (e.g. skills, tasks) [7, 37, 38]. These methods provide meaningful information for stakeholders (i.e. structure of existing jobs, skills, and their relationships), however, their dynamicity is limited, since building and maintaining ontologies to cover a wide range of occupations and skills, are currently done manually (by subject matter experts), which is a very costly and time-consuming exercise [39].

Text Mining and Machine Learning Methods. A number of studies analyze online vacancy announcements to classify job components (e.g. skills, tasks) according to existing, static taxonomies (e.g. ESCO). This is done to update taxonomies and provide fresh information about labor markets. Most of these papers try to extract features from the vacancies by applying embedding techniques (e.g. word2vec and doc2vec)[6], Topic Modeling techniques (e.g. LDA) [6, 40], TFIDF [41] and afterward, use classification techniques such as Logistic Regression, SVM, and Random Forest [42, 43] or calculate distance [41] to assign job vacancies to their closest job class. Furthermore, a number of papers are focusing on using Text Mining and clustering techniques to find relationships between skills and jobs, and to calculate similarity measures [9, 27]. These papers build vectors for skills using embedding techniques (e.g., Bag of Words) and apply clustering techniques such as K-Means [27] to find the structure of related skills and jobs. Contrary to Ontology-based systems, given that a powerful model is constructed, these methods can automatically extract the required information from job vacancies. However, the identification of such general models remains challenging.

Content Analysis. Several papers focus on specific job areas, and collect related job vacancies from various sources (e.g. job boards, and newspapers). Subsequently, they apply content analysis techniques such as counting the number of skills occurrence and skills co-occurrence in order to provide insights about skills in the investigated job area [44–46]. Although these methods are successful when finding and identifying required skills in a given job area, in most cases, they cannot scale. The reason is that mostly these studies use static lists of jobs and skills in their focus areas, which results in a "tunnel vision" and fails to detect new, emerging job components.

3.1.2 OER Recommendation

The area of OER recommendation systems has enormous development potential. The available literature on OER based content recommendations to learners is currently limited [23] and there is no signal that factors related to typical lifelong learning goals (skills, jobs) play any role here. To structure recent developments, we clustered available studies into the following four categories:

Heuristic Method. [20] examines the Cold Start problem [47] in the case of new micro OERs. The paper defines rules, based on recommended sequences of learning objects (e.g. some learning objects should be learned before others) using an existing ontology and calculates a Violation Degree according to the rules. The more a learning path violates the rules, the higher the Violation Degree is. Subsequently, the system recommends and adds new OERs into users' learning paths, based on minimizing the violation degree.

Semantic and Ontology Based Methods. [24] builds an ontology for learners, learning objects, and their environments to establish similarity measures between learning objects. This is done in order to update learning objects' properties and provide diverse and adaptive recommendations. Some studies make use of ontologies and open source RDF data to leverage semantic content, and define recommendation algorithms suitable for linked data [21, 23, 25]. Moreover, [22] tries to define an open linked vocabulary to describe user profiles, in order to facilitate recommendations.

Social Network Analysis. [48] uses social networks to build graphs of OERs and learners. Therefore, it finds tweets that have valid URLs, and builds a graph, based on the co-occurrences of tweets' hashtags. They also build a similar graph with users, based on their *mentions* and *retweet*. Finally, they recognize important and influential hashtags, and use density and centrality measures from the graphs to provide recommendations.

Machine Learning. [25] attempts to classify users (and their demographic features) with the help of Decision Trees and Naive Bayes algorithms to recommend them OERs. Furthermore, [49] uses Document Clustering and LSA in order to find similar OERs and use them for recommendations.

3.1.3 Findings on Labor Market Driven Personalized OER Recommendation

Based on the state of the art, it is clear that 1) it is worthwhile and timely to consider labor market information to define learning goals; 2) Efforts to decompose jobs into components suitable for educational purposes are still in their infancy, and 3) the area of OER recommendation systems is an under-researched area, with a number of challenges from a technical (e.g. available algorithms, data integration, scalability) perspective.

3.2 Metadata Analysis of OERs

According to the importance of *metadata* in building recommender systems, in this section, we study the existing researches on the area of educational resources metadata. First, we explain different dimensions and metrics in the area of educational resources and after that, we cover the efforts on improving the quality of educational resources' metadata.

3.2.1 Defining Dimensions and Metrics for Metadata

Currently, the following dimensions have been proposed to determine the quality of OER metadata: *completeness, accuracy, provenance, consistency, coherence, timeliness*, and *accessibility* [50]. Ochoa and Duval [51] have defined a set of calculated metrics based on the dimensions, which have been widely reused by researchers addressing OERs' metadata quality [52]. Moreover, they evaluated the metrics regarding *completeness and accuracy* on *425* OERs from the ARIADNE Learning Object Repository [53]. Palaez and Alarcon [54] have evaluated the completeness and consistency of OERs metadata based on Ochoa and Duval's metrics [51] and the standardized domain values (e.g., language should be according to *ISO 639-111* language standard).

3.2.2 Improving the Quality of Metadata

To have high-quality metadata, some methods have been developed in order to help authors and experts in providing metadata for OERs. A process for improving the metadata quality of OERs was developed to support domain experts with metadata creation; the process introduces qualitative methods (e.g., online peer review of metadata) and tools (e.g., metadata quality assessment grid) in the various phases when it comes to populating metadata in OER repositories [55]. Furthermore, a higher level of metadata quality analysis was applied to help metadata creators to assess and improve the quality of metadata [56]. They exploit linked open data to discover and analyze connectivity between metadata records. Accordingly, they used network statistics (e.g., the density of graphs) to calculate the relationship between the metadata records in terms of their attributes (e.g. subject) and values. Their study was applied to six large digital library collections and they discussed several improvements that can help users find related resources.

3.2.3 Findings on Metadata Analysis of Educational Resources

Based on the state-of-the-art, although there are several attempts regarding assessing and improving OER metadata, most of these efforts are either conceptual [57], or focusing only on a few dimensions [58, 59]. Furthermore, currently, there are no scalable solutions available [53], which limits the capability of existing approaches, when it comes to OER metadata quality assessment and improvement [60]. Therefore, it is clear that there is a significant need for improving the discoverability, usability, and reusability of OERs with the help of intelligent metadata quality assessment [60]. Also, our recent research [26] showed that there is a close relationship between metadata quality and content quality of educational resources. As a result, we conclude that: it is worthwhile and timely to analyze OER metadata and build metadata-based quality prediction models which not only improve OER-based services, but also facilitate the quality control processes of OERs.

3.3 Topic Analysis of OER

One of the key features of educational resources is their target topics. These target topics are important for classifying the resources and subsequently, recommending the most fit resources to each learner. Also, topic analysis of educational resources helps us to extract the topics that should be covered in order to learn a knowledge area. In this section, we categorize the existing researches on the topic analysis of educational resources based on their methods, and cover the researches in each category.

3.3.1 Semantic-based Methods

A number of studies use semantic methods and structured representation of data (such as taxonomies) to extract topics from educational resources. For instance, [15] proposed a framework that combines semantic classification, taxonomies, and graph structures to extract topics and detect their relationships. As another example, [16] extracts topics from slides and videos in *Massive Open Online Courses* (e.g. Coursera) on the basis of their components (e.g. header, body, and footer) using *Explicit Semantic Analysis* (*ESA*). [61] uses semantic technologies and statistical relevancy ranking methods to extract topics in online educational forums to help learners and tutors to align their discussions. Moreover, [62] collected educational resources from the area of *Natural Language Processing* (*NLP*), and built topic taxonomies based on expert annotation and pre-defined criteria (e.g. the existence of a Wikipedia page) in order to create a recommender system for NLP learners.

3.3.2 Text Mining Methods

Studies in this group analyze educational text and use text-related machine learning methods to detect topics in educational resources. For example, [63] built an associative classifier to extract keywords from distance learning resources in social media. [64] calculated sentence similarity in web learning resources by computing relations between words in sentences, then applying clustering techniques on sentences, and finally constructed summary sentences as topic descriptors for the extracted themes (sentence clusters). Furthermore, [28] created a system, which collects domain-specific content from online learning systems. They extracted domain-specific terms by creating *Generalised Suffix Tree* (*GST*) from resources' text and detected repeated sub-sequences as candidate terms to provide topic-specific recommendations for learners.

3.3.3 Findings on Topic Analysis of Educational Resources

According to the literature, it is clear that 1) automatic topic extraction methods from OERs are progressing, and therefore, 2) it is essential to put more emphasis on automatic OER metadata generation, especially when it comes to the identification of *Covered Topics* in OERs. Thus, 3) providing reliable metadata on *covered-topics* is a key requirement for providing high-quality OER-based services such as search and recommendation.

3.4 Curriculum Development in Online Environments

The extremely fast changes that we face in learning goals (e.g., skills need from the labor market) and learning methods (towards *online* learning platform) have emphasized the need for dynamic curriculum systems. Here, we investigate the most recent efforts, including *AI-based* and *Crowdsourcing-based* methods, that aid the dynamic curricula development process.

3.4.1 Artificial Intelligence Based Curriculum Development

AI has been aiding curriculum development predominantly by using *Machine Learning* [26, 65] and *Text Mining* [17, 66] methods. [65] proposed an educational program model (i.e. prerequisite, content, expected outcome) based on labor market demand using *AI back-propagation* concept in order to help learners to up-skill themselves toward their current or desired job. Although they claim that the result of their curriculum model is promising, they focused on the single area of *Internet of Things* (*IoT*). [66] created knowledge graphs of *Open Educational Resources* (*OERs*) by applying *Natural Language Processing* (*NLP*) techniques to help authors to deliver their content to the proper audience. However, they only focused on content level and not on higher level learning goals. Moreover, some researches [26, 67] proposed approaches that perform an automatic quality assessment for educational resources in order to help content providers to filter out low-quality content from their resources list. [17] introduces a novel method that uses *Latent Dirichlet Allocation* (*LDA*) [68] algorithm to extract topics covered by specific educational resources, in order to build learning pathways. Their application only focused on building a curriculum for educating *text mining* to learners.

3.4.2 Crowdsourcing Based Curriculum Development

As user participation in online learning platforms increases [69], "the wisdom of the crowd" has great potential for both teachers and learners. "Crowdsourcing for education" has been used for content-creation [70] and also for sharing practical and theoretical knowledge [71] on large scale. [72] concluded that including the crowd's opinion in the process of education is useful when it comes to building scalable and personalized curricula. [73] integrated crowdsourcing into the pedagogical paradigm as "Crowdlearning". They suggested that including students in the creation of the curriculum would not only increase the amount of content produced, but also improve the depth and performance of learning. However, crowdsourcing has its own flaws as well. Participation time and its effectiveness for complex tasks are two of the most noted issues [74, 75].

3.4.3 Findings on Curriculum Development in Online Educational Environments

While crowdsourcing is generally useful in content-creation for curriculum development, and also contributes to a system's scalability, the time consumed by a participant can cause motivation problems and prevent their effective participation [74]. On the other hand, AI can help us to automate some tedious tasks such as quality assessment and learning topic extraction from educational resources. Therefore, novel curriculum development methods are needed, which utilize the benefits of both AI and crowdsourcing approaches to generate dynamic and scalable personalized curricula.

3.5 Findings through the Literature Review Steps

Based on the literature review, it is clear that:

- Online educational services that offer open and online educational resources to learners may address the pressing need for inclusive, equitable, and effective online education, while at the same time becoming ever more viable as technology is developing.
- The time is ripe to focus greater research attention on personalized high-quality educational resource recommendations. This is especially important for learners in further education, who need help to build their own learning trajectories toward their desired jobs.
- Having mechanisms (especially to get help from AI and crowds) to develop, maintain, and update high-quality inclusive curricula plays a key role in offering personalized goal-driven learning

Figure 3.1 summarizes the outcome of this chapter which includes the following information for each area (investigated research questions) 1. the target area, 2. the existing methods to tackle the issues, and 3. our findings through our investigations in the area. Our conclusion from the whole literature review is also available in the figure.

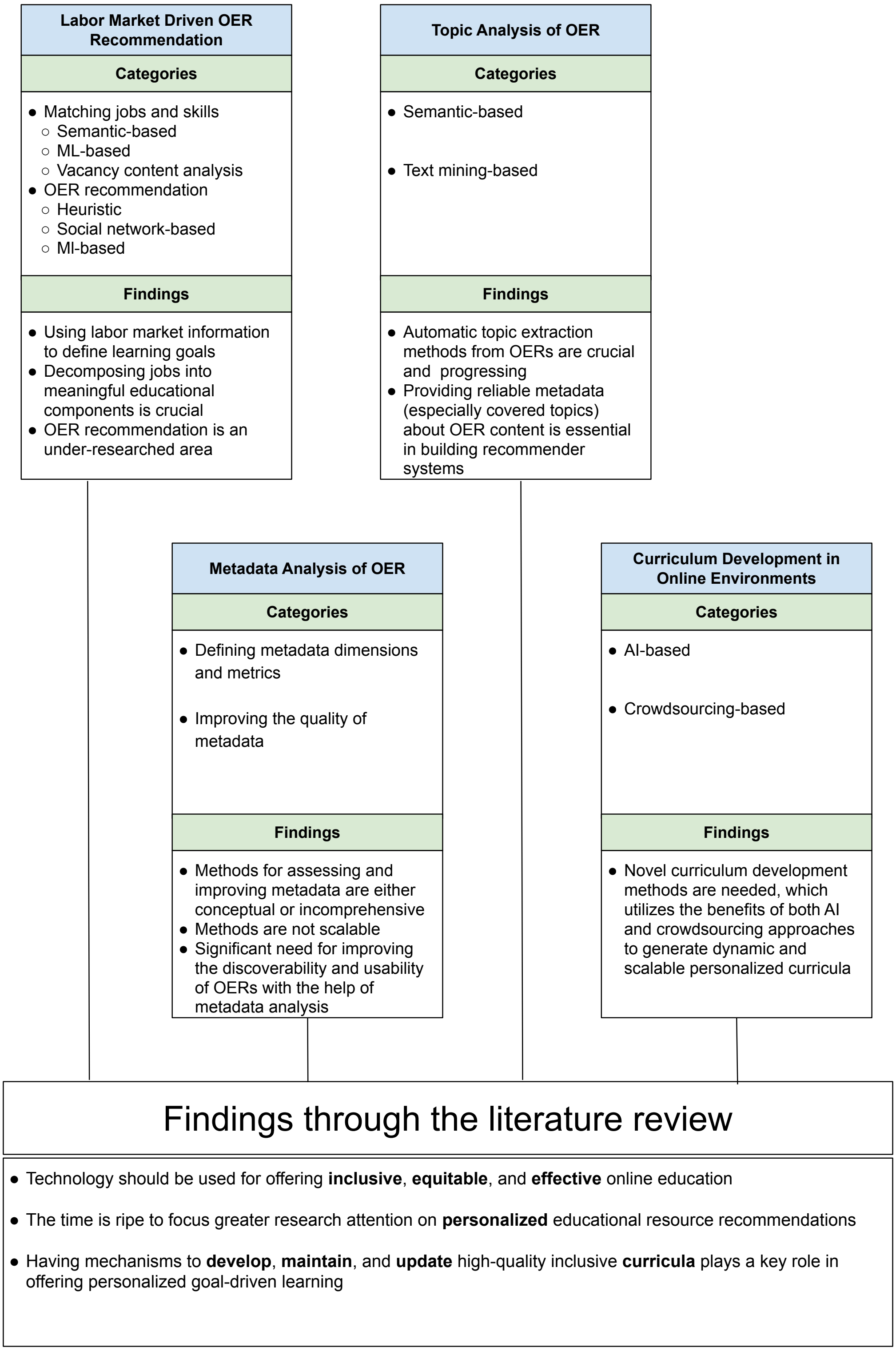


Figure 3.1: Components of our Labour Market Intelligence (LMI) based OER recommender

CHAPTER 4

Prototype of Goal-driven Personalized Learning Environment[1]

In this chapter, we suggest a novel method to aid lifelong learners to access relevant OER based learning content to master skills demanded in the labor market. Our approach 1) applies Text Classification and Text Mining methods on vacancy announcements to decompose jobs into meaningful skill components, which lifelong learners should target; and 2) creates a hybrid OER Recommender System prototype to suggest personalized learning content for learners to progress toward their skill targets. For the first evaluation of this prototype, we focused on two job areas: *Data Scientist*, and *Mechanical Engineer*. We applied our skill extractor approach and provided OER recommendations for learners targeting these jobs. To clarify our idea, in this chapter, first, we explain the steps toward implementing the proposed system. After that, we illustrate how we evaluated our software prototype through semi-structured interviews in order to recognize the key requirements for further developing our system.

4.1 METHODS

In this section, first, we discuss the data collection step regarding collecting job vacancies data. After that, we illustrate how skill extraction from job vacancies is done in order to create a job-skill matching algorithm. Finally, we explain our personalized OER recommendation algorithm to help learners achieve their target skills.

[1] This chapter has been published as follows: Tavakoli, M., Mol, S.T. and Kismihók, G., 2020. Labour market information driven, personalized, OER recommendation system for lifelong learners. In the 12th Computer Supported Education Conference (CSEDU).

4.1.1 Data Collection

For the prototyping, we used a crawled sample dataset from Monster.com containing 22,000 job vacancies[2]. We used 80% of our dataset for training and cross-validation and 20% of them as our test set. Moreover, for our OER recommendation, we have used APIs, provided by the following OER providers: SkillsCommons[3] and Wisc-Online[4].

4.1.2 Labor Market Intelligence (LMI)

Extracting Skills from Job Vacancies

Since our aim was to avoid any dependency on existing taxonomies (which are updated slowly), we put existing methods classifying jobs and skills into predefined classes aside, and created a dynamic job-skill matching mechanism to detect skill changes in jobs quickly. As the first step, we constructed a model to find skill-related sentences in job vacancies. After an exploratory analysis, we concluded that large number of vacancies do not contain a "Required Skills" section. Therefore, in order to build our model, we selected vacancies with an explicit "Required Skills" section and run the following preprocessing procedure on each of those vacancies:

- Deletion of unimportant characters, punctuations, and bullet-points
- Removal of irrelevant stop words
- Removal of conjunctions, articles, and prepositions
- Sentence Tokenization
- Lowercase Conversion
- Lemmatization

Altogether we obtained more than 60,000 sentences with this method. This corpus included both sentences, which were mentioned in a "Required Skills" section (we set their label to 1), and also sentences mentioned in other sections in vacancies (we set their label to 0). As a result, we got around 15,000 sentences related to "Required Skills" labeled as 1, and around 45,000 sentences not related to "Required Skills" labeled as 0. Subsequently, we applied embedding techniques on word-level n-grams, and built sentence vectors with averaging word/n-gram embeddings and using Multinomial Logistic Regression model to minimize the classification error[5]. It should be mentioned that word-level n-gram applies the n-gram concept on the character level and find the most common sequences of characters.

[2]The dataset is accessible from: `https://www.kaggle.com/PromptCloudHQ/us-jobs-on-monstercom`
[3]https://www.skillscommons.org/
[4]https://www.wisc-online.com/
[5]We used FastText Library in Python for our classification task [76]

Table 4.1: User Properties.

Property	Values	Note
Selected job	Existing jobs	Selected by users
Skills-Levels	[0..100] for Skills	Determined by users
Personal information	Location, Gender, Education	Entered by users
Pref_Resources	[0..100] for Resources	Higher tendency→ higher value
Pref_Length	[0..100]	*Preferred_long* and *Preferred_short*
Pref_Check	[0..100]	Prefer assured → closer to 100
Pref_Accessibility	[0..100]	Prefer higher accessibility → closer to 100

Therefore, vectors are created for each of the extracted sequences of characters and it helps us build vectors for new words (skills), based on our existing vector for the new word's sequences of characters (e.g. building an initial vector for *Mechatronic* based on existing vectors which are extracted from *Elecronic* and *Mechanic*). Applying our model to the test dataset resulted in the detection of 88.7% balance accuracy (including precision and recall) of skill-related sentences. Finally, we used TFIDF weighting to detect skill terms in skill-related sentences. It should be mentioned that we used *Minimum Document Frequency* of 3 as the cut-off point in order to handle typing errors and remove rare words.

Calculating Skills' Importance for Jobs

To calculate the importance of particular skills associated with jobs in a specific geographical location, we calculated the rate of skill occurrence in the previous 6 months at the given job location. After normalizing the rates, we use a simple decay function to compute the new importance score, which combines the previous importance scores and the new rates with more weight on the new rates.

4.1.3 Recommending OERs

Method for Initializing Learners' Properties

Table 4.1 depicts learners' properties in our OER recommender prototype. During the initialization of a new user, we capture known properties entered by users (i.e. Personal Information, Skill Level List, and Selected Job), and also a number of properties without values (i.e. Resource scores, Length scores, Quality scores, and Accessibility scores). To set an initial value for these unknown properties, we sample similar users, based on the known properties and use the weighted average (based on similarity) of their properties as initial values for unknown properties. This strategy scaffolds the cold start problem of new users. To sample similar users, we use (4.1) to compute the similarity between user i and j where the Similarity Effect function for user i and j in property k is calculated as (4.2).

$$similarity\left(i,j\right) = \frac{\Sigma_{k=known\_properties} sim\_effect\left(i,j,k\right)}{100} \tag{4.1}$$

$$sim\_effect\left(i,j,k\right)=\begin{cases} equal\_val\left(k\right), & \text{same k for i\&j} \\ 0, & \text{otherwise} \end{cases} \tag{4.2}$$

Furthermore, the equality value of property k (equal_val(k)), showing the effect of variable k on similar behavior (rating) by users, is calculated through the following process:

1. We collect user pairs who gave exactly the same ratings for the same OER in the period
2. Compute the ratio of the number of pairs having exactly the same value in property k to the number of all pairs
3. Normalizing the ratios in a way that the sum of all the ratios becomes equal to 100 and the normalized ratio of k is the Equality Value of k

This process is executed regularly, after defining a time period (e.g. after every month).

Method for Updating Learners' Properties

Since we aim to capture learners' preferences quickly and provide relevant OERs according to the changes and improvements in learners' property values, we decided to update user properties after each rating action on any of the recommended OERs. This is done by using a real-time updating process that, according to the rating score and the properties of the recommended OERs (i.e. length, quality, accessibility), updates the properties of the users. As a consequence, if a learner is satisfied (dissatisfied) with a given OER, we will encourage (discourage) the properties (see details in the next section) of that particular OER for that learner. For instance, if a user is dissatisfied with a long OER (e.g. it takes 10 weeks to complete), we will update the *Preferred Long* property of the user and decrease its value in order to provide shorter OERs in the future. Along the same line, with assigning positive ratings to accessible OERs, learners can enhance their accessibility criterion and increase their *Preferred Accessibility* value to receive content with accessibility support (critical for instance for visually impaired learners [77]).

OER Properties

Table 4.2 shows OER properties. Based on existing literature, we selected *Level*, *Length*, *Quality*, and *Accessibility* as important properties of OERs [78–80]. When assigning a value to a particular OER property, first we extract and order all existing values assigned to that property, then classify them, and count the number of classes. Based on the number of classes, we assign a value between 0 and 100 to that property. For instance, we take property *Level*, we extract 3 values (beginner, intermediate and advanced - 3 classes), and as a result, we set the value for beginner OERs to 0, intermediate OERs to 50, and advanced OERs to 100.

Table 4.2: OER Properties.

Property	Values	Note
Resource	Repositories	E.g. SkillCommons, Wisc-Online
Skill	Existing skills	Based on subjects
Author	Full name	The provider
URL	URL	Web address of OERs
Length	[0..100]	How_long and How_short
Level	[0..100]	Higher level $\rightarrow$ closer to 100
Quality	[0..100]	More quality assurance $\rightarrow$ closer to 100
Accessibility	[0..100]	More accessibility $\rightarrow$ closer to 100
Relevance	[0..100]	Decreased if defined *Irrelevant*

Method for Initializing OER Properties

For each OER, using its known properties, we attempt to identify similar OERs. For instance, if we know *Skill* and *Author* of a new OER, we identify all other OERs provided by the same author and the same skill target, compute their average values, and set the initial property values accordingly.

Method for Updating OER Properties

Detecting OER properties is a slow process in the beginning since change happens when users alter their rating pattern. This happens usually when they are confronted with new OERs. Therefore, we run the updating process after a specific time period (e.g. once each month). To adjust the properties (except Relevance) of each OER, at first, we collect all related users and their ratings in the given time period. Afterward, we compute the property values for the OER as X in order to minimize (4.3) using Gradient Descend, where θ_i is the property vector of user i and Y_i is the satisfaction rate of user i.

$$LossFunction = \Sigma_{i=users}|\theta_i^T * X - Y_i| \tag{4.3}$$

This strategy of using all recent ratings in updating OER properties enhances the diversity in our recommendations. All learners contribute to calculating these OER properties (for each OER they studied) through their individual evaluations. Users can also rate OERs as irrelevant. As a consequence, the *Relevance* property of an OER *o* is calculated as (4.4) where the *total_recom(o)* shows the number of times that OER *o* has been recommended to users and *irrelev_count(o)* is the number of times that *o* has been determined as *Irrelevant*. Finally, OERs with a Relevance Value less than the average in relation to a specific skill, are marked as Irrelevant (for that skill only), and therefore will not be recommended (for that skill) anymore.

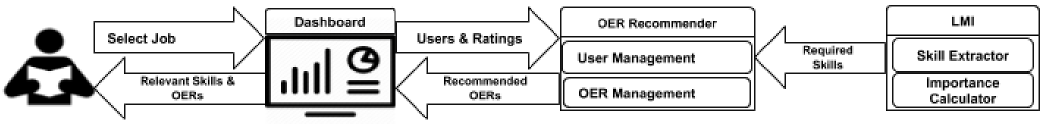


Figure 4.1: Components of our labor Market Intelligence (LMI) based OER recommender

$$relevancy\left(o\right) = \frac{total\_recom\left(o\right) - irrelev\_count\left(o\right)}{total\_recom\left(o\right)} \tag{4.4}$$

Recommendation Algorithm

For recommending an OER to a learner, we calculate Cosine Similarity between the properties of candidate OERs (which are related to the skill level of any user) and the properties of the user. The system will recommend an OER with the lowest distance between those two. Since we update user properties in a real-time process and update OER properties after a predefined period, for recommending the best match for a user, we only need to find an OER, which has the closest properties to the user. Furthermore, the Rating Sparsity problem (i.e. users rate only a few OERs) is one of the most important issues when building recommender systems. In our case users and OERs have a mutual contribution to calculating properties, which intends to eliminate the effects of Rating Sparsity. Even if an OER has a limited amount of ratings, we can rely on the properties of the learners. On the basis of their ratings on other (similar) OERs, we calculate the properties for OERs suffering from Rating Sparsity.

4.1.4 Recommender Prototype Overview

Learners were confronted with a prototype of our recommender system in a form of a dashboard[6]. Through this dashboard, learners can search for their current or desired job, display the list of required skills, and set their level of expertise for each skill. Subsequently, on the learning tab, the dashboard shows the current expertise levels of the learner, and the links to the recommended OERs. OERs are ordered according to the importance of skills for the selected job. In case a learner thinks that a recommended OER is not related (Irrelevant) or does not find the content engaging, a new recommendation could be generated, without changing the expertise level of the learner. After consulting (learning) a recommended OER, learners are asked to rate their satisfaction with that OER. Finally, the dashboard updates the learner's expertise level and provides an updated recommendation based on the new rating. This is done until the learner masters all required skills at the highest level. Figure 4.1 depicts the building blocks of our proposed approach.

[6]You can find a demo of our prototype from: https://github.com/rezatavakoli/CSEDU2020

4.2 VALIDATION

In this part, we explain how we evaluated our first prototype through semi-structured interviews with experts from the area of *Data Science* and *Mechanical Engineering*. To validate our proposed approach, we conducted semi-structured interviews with subject matter experts in the job areas of *Data Science* and *Mechanical Engineering*. We focused on jobs, which are related to these areas, and randomly selected 100 job vacancies for Data Scientists and 100 job vacancies for Mechanical Engineers from August 2019. Afterward, we applied our skill extraction and importance detection model to select the most important skills in both occupations. To evaluate recommendations, we invited four university instructors with at least 12 years of teaching and 13 years of industrial experience and eight Ph.D. students with a minimum teaching experience of 1 year and a minimum industrial experience of 2 years for a semi-structured interview[7]. Participants gave feedback on our prototype with regard to its general objectives, logic, and potential contribution to individual learning. Each interviewee had to go through the following protocol:

1. Learning about the research problem and the proposed approach - 15 minutes
2. Work with our prototype - 15 minutes
3. Going through a semi-structured interview with the help of a qualitative questionnaire[8] - 30 minutes

During working sessions with our prototype, participants generated more than 150 OER recommendations. 76.9% of these recommendations were useful and relevant to participants' skill levels and properties. 8.2% of the recommended OERs were signaled as irrelevant, and in 14.9% of the cases, participants decided to change the recommended OERs. The results of the interviews are summarised in the following three sections.

4.2.1 Objectives

Interviewees confirmed that there is a potential value in building a labor market information driven OER recommender system. Both instructors and Ph.D. students thought that there are several useful and high-quality OERs available on the Internet, but finding them is complicated and time-consuming. Regarding skill extraction, participants recommended to use alternative data sources, besides vacancy announcements. *Student_2* for example suggested that "you should also use other data sources related to the labor market like CVs and available data about salaries". Moreover, interviewees thought that this approach is extremely useful for job-seekers, job-holders, and people who have clear ideas about their preferred occupations. However, they were skeptical about those learners, who want to focus their attention on a specific skill only.

[7] Detailed profiles of our interview participants are available on: https://github.com/rezatavakoli/CSEDU2020
[8] The questionnaire is available on: https://github.com/rezatavakoli/CSEDU2020

4.2.2 Logic

Participants confirmed that our method to calculate the importance of particular skills in recent job vacancies can potentially help learners to focus on the most important elements of their current or future job. However, as it was also suggested by *Student_1*, a more intelligent decay function, to combine recent and previous skill important values might be desirable. Regarding the self-assessment of learners to set their initial level of expertise, *Instructor_1* suggested to "introduce basic assessment in a form of technical or non-technical questions" for each targeted skill.

4.2.3 Contribution to Learning

Participants emphasized that interacting with learners in order to recognize their preferences (e.g. recommending OERs based on their previous ratings) is one of the most important, novel and engaging components of our proposed approach. *Student_5* recommended to include more properties: "You should capture more learners' properties such as language preferences or type of OERs (e.g. presentation, video)." Moreover, interviewees were convinced that setting specific and personalized goals for each skill in our prototype system has a strong and positive effect on the learning process.

4.3 Outcome of Creating the Prototype of Goal-driven Personalized Learning Environment

In this chapter, we showcased a hybrid OER Recommender system prototype to support individual skill development, targeting concrete, labor market oriented skills and jobs. For this prototype, a skill extraction mechanism has been constructed, which captures skill-related sentences in vacancy announcements with a balanced accuracy of 88.7%. These dynamically generated skills became individual learning objectives and were connected to OER based learning contents. Recommendations were generated through a dashboard, with combining OERs and learner properties. The system prototype was validated with semi-structured interviews. The initial results showed that our proposed approach has the potential to aid lifelong learners to construct their individual learning pathways and progress toward their desired job-related skills. Moreover, participants valued that user properties were critical when formulating recommendations.

CHAPTER 5

Analyzing Educational Resources Toward Building Recommender Systems[1]

As a result of the prototype evaluation in the previous section, we understood that it is reasonable to put more effort into our software concept. Therefore, we decided to decompose the system into different components and started working on them separately. Therefore, we suggest a novel method that analyzes educational resources in order to help learners find relevant open educational videos to master skills demanded in the labor market. We have built a prototype, which 1) applies text classification and text mining methods on job vacancy announcements to match jobs and their required skills; 2) predicts the quality of educational videos according to their features (e.g., length and view count) and target areas; and 3) creates an open educational video recommender system to suggest personalized learning content to learners. In this chapter, we illustrate the data collection steps, followed by the system development procedure. In the end, we explain our validation strategy.

5.1 Data Collection

In this section, we describe the data we collected to build an open educational platform to recommend educational videos. Firstly, we describe the procedure of collecting skills, followed by an explanation on the retrieval of educational videos.

[1]This chapter has been published as follows: Tavakoli, M., Hakimov, S., Ewerth, R. and Kismihok, G., 2020, July. A recommender system for open educational videos based on skill requirements. In 2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT) (pp. 1-5). IEEE.

5.1.1 Skill Collection

The first step for building our recommender was the identification of skills that are correlated with particular jobs. To do this, we used the model developed in Chapter 4. We ran our skill extraction method on 300 randomly crawled data science job vacancies (the context of this chapter), which have been published on Monster.com in December 2019, and obtained a list of skills that learners should focus on for building a career in data science. In total, we extracted 16 important and unique data science skills. We provide a sample skill with other metadata below.

Skill: Python programming

Keywords: python, python programming

Description: Python is an interpreted, high-level, general-purpose programming language.

To find skill descriptions we used Wikipedia python API[2] and crawled the Wikipedia content, which is related to skills[3].

5.1.2 Open Educational Videos

We collected educational videos from two main sources: YouTube and TIB AV portal[4]. YouTube is the most popular platform for hosting any type of video content. The TIB AV-Portal[5] is a dedicated portal for scientific videos from the realms of architecture, chemistry, computer science, engineering and technology, mathematics, etc. and the videos include among others, computer visualizations, learning material, simulations, experiments, interviews, video abstracts, and recordings of lectures and conferences.

We retrieved videos by performing a keyword search on each portal. As explained above, each skill contained a set of keywords. All keywords were used to search and retrieve relevant videos. Videos from both sources might contain transcriptions of audio. YouTube includes them as subtitles, TIB AV Portal shows the body of the transcribed text. Upon availability, we extracted these transcriptions for retrieved videos. Missing transcriptions were obtained by applying Google Cloud Speech[6] on audio files extracted from videos.

Videos contain different types of information depending on the source. We collected/calculated the following metadata from YouTube and TIB AV portal videos:

[2] https://pypi.org/project/wikipedia/

[3] Complete list of skills with their properties is available: https://github.com/rezatavakoli/ICALT2020_recommender

[4] For this prototyping exercise, we used openly available videos, but we disregarded the type of license for our analysis. Nevertheless, licensing will obviously play a role in future implementations.

[5] https://av.tib.eu/

[6] https://pypi.org/project/google-cloud-speech/

YouTube: title, target skill, URL, length, description, transcription, view count, rating, likes, dislikes, relevancy score (assigned according to the rank in the search results), Textual similarity (which is calculated based on the similarity between skill description and video transcription. We explain this calculation in Section 5.2)

TIB AV Portal: title, target skill, URI, description, transcription

For developing our solution, we retrieved 550 videos from YouTube and TIB AV portal, which covered the 16 skills we fetched previously. These videos were presented to six experts in data science (with more than six years of industrial and more than three years of teaching experience in data science-related positions) to annotate whether they fit their target skill, or not. Each video was reviewed by at least three annotators and annotators assigned at least 2 minutes to set the label of each video. The final label was assigned based on a majority vote. In total, the annotators provided labels for 550 videos, where 213 of them fit to a skill (positive label) and 337 did not fit (negative label). The complete list of videos and labels is available for the research community[7].

5.2 Method

The following section provides details about how the analysis of educational videos was performed for building a recommender system that provides the fittest educational resources for learners. Also, we illustrate the learners-recommender interaction at the end of this section.

5.2.1 Fit Prediction

We trained a machine learning model to predict whether a given video fits to a skill or not. As mentioned earlier, we selected 550 videos for 16 skills to annotate whether a video fits a skill or not and annotators provided labels for 213 videos as a fit, and 337 were annotated as not fitting.

A Random Forest model was trained on the annotated data to build a model that outputs a binary decision: match/no-match. The algorithm used the following video features to train our model.

- Length: the length of a video in seconds
- Rating: the user rating, what a video received on a platform
- View count: the number of views on a video
- Relevancy score: the score assigned during the search process based on the video-platforms' results ranking as $\frac{1}{ranking\_position}$

[7] Annotated dataset including videos' properties and labels: https://github.com/rezatavakoli/ICALT2020_recommender

- Level: the pre-defined levels either *beginner*, *intermediate* or *advanced*. The levels are set during the collection process by concatenating the search term with "beginner", "intermediate" or "advanced" to search for videos at different levels.

- Text similarity: the similarity is computed between skill description and video transcription. First, each word in a text is encoded using pre-trained 300-dimensional Glove vectors [81]. Second, we average the vectors of words in a text to get a single vector that represents the whole text. We apply the described method to obtain a vector representation for both video transcription and skill description. Finally, the text similarity is a cosine similarity between the resulting two vectors.

70% of the data was used to train the model and the remaining 30% was used for the evaluation. The classifier achieved an F1 score of **86.3%** in predicting whether a video matches a skill. Additionally, we analyzed the importance of each feature for the classification task. The trained model assigned different importance scores to each feature on the basis of the provided training data. Each feature had a different weight on the decision based on these scores. The weights were calculated by pruning out trees below a particular node (as feature selection). The weights for the selected features were calculated as follows: length: 0.61, rating: 0.10, view count: 0.10, relevancy score: 0.08, level: 0.2, text similarity: 0.09. The model assigned the highest score for the *Length* feature and it was followed by the *Rating* feature. In the following section, we describe how the trained binary classifier can be used within a recommender system.

5.2.2 Recommendation Generation

Our proposed recommender system suggests new content to learners based on different parameters. The goal is to optimize weights for these parameters by increasing learner satisfaction (based on their ratings). The recommender system uses the following parameters:

- Popularity: We calculate the difference between the number of likes and the number of dislikes for each video, group videos by their target skill, and for each group, we normalize the values using Minmax normalization

- Fit probability: The probability of a video fitting a skill (explained in Subsection 5.2.1)

- Length: The length of the videos; we group videos by their target skill, and for each group, we normalize the length using Minmax normalization

- Text similarity: The textual similarity between a video transcription and a skill description

We build a 4-dimensional vector of X where each item in the vector is a value for a parameter mentioned above. We define a vector P as a preference matrix for each user that contains a weight for each parameter in X. The goal is to optimize weights in P for each learner based on previous ratings. In this way, we capture learners' preferences to provide personalized recommendations. The following loss function is used to optimize the weights in P for each user with respect to the generated recommendations.

$$LossFunction = \Sigma_{i=recommendations}|P * X_i - Y_i| \tag{5.1}$$

where X_i is the mentioned 4-dimensional vector of a recommended video i and Y_i is the satisfaction rate of the user for that particular video i. We use Gradient Descend to find the best P for each user and the initial weights in P are set by taking weights from similar users (e.g. with the same job, location, etc.).

Finally, the system generates a recommendation based on optimized weights in the preference matrix P given the parameters of each video for a target skill. The videos are ranked by computing a cosine similarity between their matrix X and the preference matrix P of a user. The video with the highest score is given as a recommendation to a user.

5.2.3 Learning Dashboard

We have built the prototype of our recommender system in the form of a dashboard[8]. The users interact with the dashboard for searching or adding skills they want to master, setting their levels of expertise for each skill, and adding contextual information about their occupation, geographical location and educational level. Subsequently, on the learning tab, the dashboard shows the list of their target skills, the learner's current expertise levels in each of them, and the links to the recommended open educational videos. Learners can watch the recommended content or ask for a new one in case they are not satisfied with the recommendations.

After watching a recommended video, the learner rates her satisfaction with the recommendation. The system changes the learner's expertise level, updates her preference matrix P, and provides a new recommendation based on the new expertise level and preferences. The process continues until the learner reaches the highest mastery level for a particular target skill. Figure 5.1 depicts the building blocks of our proposed approach.

5.3 Validation

To validate our proposed approach, we conducted semi-structured interviews with subject matter experts in the job area of *Data Science*. We randomly selected 300 job vacancies related to data science from Monster.com which have been published in December 2019. Afterward, we collected the required skills for data scientists as described previously in Section 5.1.1.

[8]Demo of our prototype is available: `https://github.com/rezatavakoli/ICALT2020_recommender`

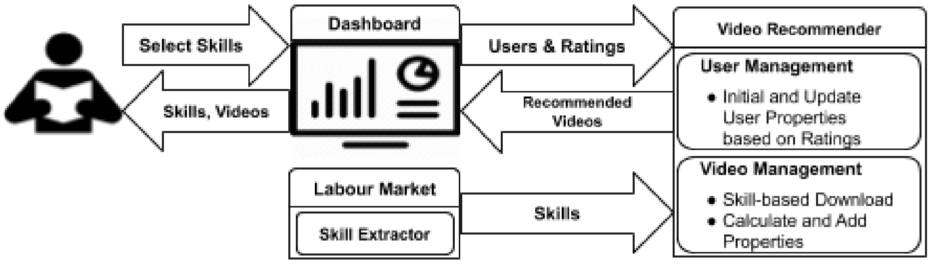


Figure 5.1: Components of our Labor Market based Video Recommender

To validate the proposed recommender system, we invited five university instructors with at least 10 years of teaching and 13 years of industrial experience, and 10 Ph.D. students with a minimum teaching experience of one year and a minimum industrial experience of three years for a semi-structured interview[9].

Participants gave feedback on our prototype with regard to its general objectives, logic, and potential contribution to individual learning. Each participant had to complete the following protocol:

1. Learning about the research problems and the proposed approach - 15 minutes

2. Work with our prototype dashboard - 15 minutes

3. Going through a semi-structured interview with the help of a qualitative questionnaire[10] - 30 minutes

Participants received more than 250 video recommendations while working with our prototype dashboard (Each participant received 15-17 recommendations). 82.8% of these recommendations were signaled as useful and relevant to participants' skill levels and properties. 2.8% of the recommended videos were recognized as irrelevant, and in 14.4% of the cases, participants decided to change the recommended video. The outputs of the interviews are summarised in the three following sections.

5.3.1 Objectives

Interviewees confirmed that there is a potential value in recommending open educational videos based on labor market information. Both instructors and Ph.D. students expressed that although there are several open educational videos on the Internet, finding the most suitable content for learners' preferences is a complicated and time-consuming task. For instance, *Instructor_2*, *Instructor_5*, *Student_10* told that personalization of open educational content recommendation is one of the most important features of our proposed approach. Moreover, *Instructor_3* suggested that we should

[9] Detailed profiles of our interview participants are available on: https://github.com/rezatavakoli/ICALT2020_recommender

[10] The questionnaire is available on: https://github.com/rezatavakoli/ICALT2020_recommender

recognize the level of expertise that the learner needs to achieve in order to prevent over-qualifying and wasting time.

5.3.2 Logic

Participants emphasized that our recommendation model can help learners in finding the most relevant videos covering particular skill areas. *Student_6* suggested that the system needs to take into account the job area of skills, which may result in fine-grained recommendations that target a specific skill for a specific job. For instance, the skill *data visualization* might have different content depending on the job areas such as E-commerce or Bioinformatics. Regarding the recommendation logic, participants thought that suggesting videos based on learners' previous ratings is a novel idea and *Instructor_2* told that the system should use more user properties such as language preferences, province of residency, etc. and also properties of similar users in order to generate better recommendations.

5.3.3 Contribution to Learning

Participants confirmed that engaging with learners based on their preferences may result in better retention rates for learners. *Instructor_4* and *Student_1* valued that setting specific goals and recommending videos to learners accordingly help them focus on their skill targets and could potentially improve their learning performance. Also, *Student_4* and *Student_7* recommended to build a list of topics, which should be associated with skills, and use these topics to improve skills assessments at the beginning (setting initial expertise levels) and also during the learning process (e.g. evaluating knowledge gains after watching videos).

5.4 Outcome of Analyzing Educational Resources Toward Building a Recommender System

In this chapter, we demonstrated a recommender system prototype, which analyzes open educational videos, and built a personalized learning environment, where users can select skills and master them based on labor market information. The recommender was validated with semi-structured interviews with subject matter experts. The initial results showed that participants were satisfied with 82.8% of the generated recommendations.

CHAPTER 6

Metadata-base Scoring and Quality Prediction of Educational Resources[1]

Based on the work depicted in the previous chapter, we realized that since many online learning repositories provide millions of OERs, it is exceedingly difficult for learners to find the most appropriate OER among these resources. Subsequently, precise OER metadata is critical for providing high-quality services such as search and recommendation. Moreover, metadata facilitates the process of automatic OER quality control as the continuously increasing number of OERs makes manual quality control extremely difficult [26]. Therefore, we decided to work on the metadata quality of OERs by performing an exploratory data analysis on the metadata of *8,887* OERs. In this chapter, first, we discuss our data collection and analysis methods. Afterward, we explain how we validated our results, and finally, we analyze the outcome of our effort.

6.1 Data Collection and Research Method

In this part, we cover 1. the data collection step from two educational sources (i.e., SkillsCommons and Youtube), 2. exploratory analysis of the collected resources in order to investigate the importance of each metadata field in the dataset, 3. creating metadata scoring model based on our analysis, and 4. building a machine learning model which considers the metadata fields and our proposed scoring model to predict the general quality of educational resources.

[1]This chapter has been published as follows: Tavakoli, M., Elias, M., Kismihók, G. and Auer, S., 2021, April. Metadata analysis of open educational resources. In LAK21: 11th International Learning Analytics and Knowledge Conference (pp. 626-631).

6.1.1 Data Collection

We have used two datasets to analyze the OERs metadata and evaluate our model. The *SkillsCommons* dataset was used to analyze and train our machine learning model and the *YouTube* dataset was used to evaluate our prediction model.

SkillsCommons. For analyzing the OERs metadata and building the quality prediction model, we retrieved all search results for the terms *Information Technology* and *Health Care* via the SkillsCommons platform API and built our OER metadata dataset [26]. The dataset contains *8,887* OERs metadata[2]. The OER metadata in our sample included the following fields: *url, title, description, date of availability, date of issuing, subject list, target audience-level, time required to finish, accessibilities, language list, and quality control* (i.e., a categorical value that shows if a particular OER went through quality control or not). It should be mentioned that the *quality control* field means manual quality control, and it was set to **with control** if an OER had at least one inspection regarding the Quality of Subject Matter, and at least one inspection regarding the Quality of Online/Hybrid Course Design, otherwise it was set to **without control**.

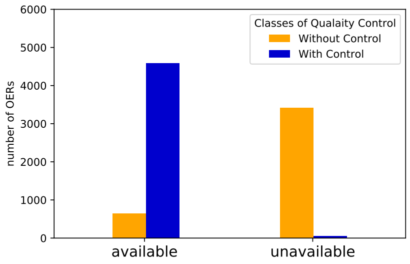
Youtube. To evaluate our proposed model, we selected 16 topics, which are defined by [82] as *Information Technology* related search keywords. In addition, we randomly selected another 16 topics from [83] as *Health Care* related search terms. Afterward, for each of the 32 selected topics in the areas of *Information Technology* and *Health Care*, top videos in Youtube search results were collected[3] using *Pafy* python library[4]. In a Youtube search, the number of top videos appearing in search results depends on the search query topic, and therefore, we can be confronted by a different number of videos as top results. However, we collected at least 10 videos per each search term. In the end, *884* Youtube educational videos were collected for our evaluation step[5]. The video metadata includes the following fields: *url, title, description, number of dislikes, length, number of likes, rating, subject list*, and *number of views*.

6.1.2 Exploratory Analysis of OER Metadata

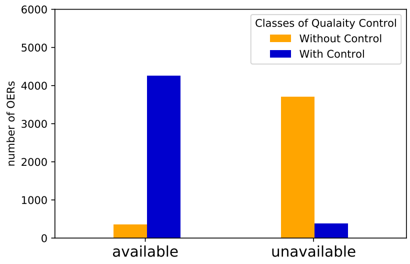
As a point of departure, we used our Skillscommons dataset to explore the availability of different OER metadata elements (i.e., level, language, time required, accessibilities) based on their quality control categories ("with control" or "without control"). The results of the analysis are summarized in Figure 6.1:

- *Level* refers to the learners' expertise or educational level in relation to a specific OER. Figure 6.1(a) illustrates how quality control increases the availability of level metadata.

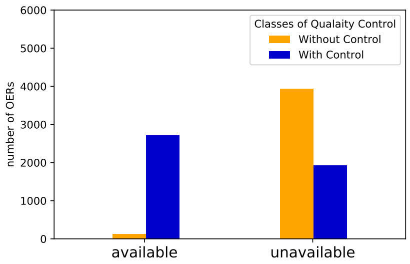
[2]Our *SkillsCommons* dataset is available on: https://github.com/rezatavakoli/ICALT2020_metadata
[3]Our *Youtube* dataset is available on: https://github.com/rezatavakoli/LAK21_metadata
[4]https://pypi.org/project/pafy/
[5]For the current version, we used openly available videos, but we disregarded the type of license for our analysis.



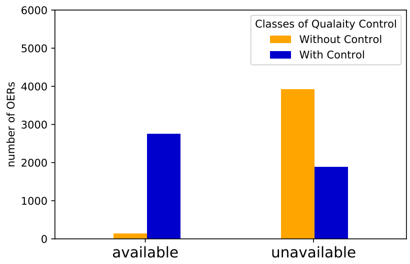
(a) *Level* metadata



(b) *Language* metadata



(c) *Time Required* metadata



(d) *Accessibilities* metadata

Figure 6.1: Analyzing metadata availability with respect to manual quality control

- *Language* refers to the available language versions of an OER. Figure 6.1(b) illustrates the effect of quality control in increasing the availability of language metadata.
- *Time Required* refers to the expected duration needed to complete an OER. Figure 6.1(c) shows that it is more likely that OERs with quality control have this type of metadata.
- *Accessibilities* defines the accessibility guidelines supported by an OER. Figure 6.1(d) illustrates how quality control increases the availability of the accessibility metadata.

To clarify, in each chart, bars on the left show the number of OERs including the particular metadata field, and bars on the right show the number of OERs missing that particular metadata field. Moreover, blue bars are related to the number of OERs with quality control, and orange bars show the number of OERs without quality control. For example, in the left chart of *Level* metadata, you can see more than 4,000 OERs have passed through quality control and also contain Level metadata. At the same time, around 3,000 OERs did not go through quality control, and also do not contain the Level metadata. The plots in Figure 6.1 show a clear improvement in OER metadata quality (i.e., availability) in the OERs which have passed through quality control. Therefore, this improvement can be interpreted as a result of quality control processes. However, as Figure 6.2 shows, the proportion of manual OER quality control has been decreasing over the last years in our dataset. We believe that the growing number of OER providers and contents are among the main reasons for this negative change in the proportion of manual OER quality control.

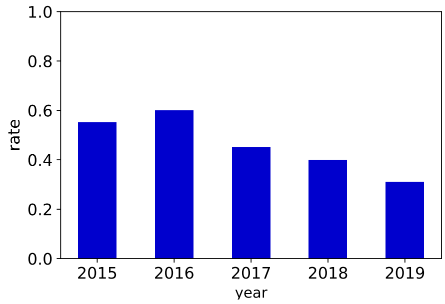
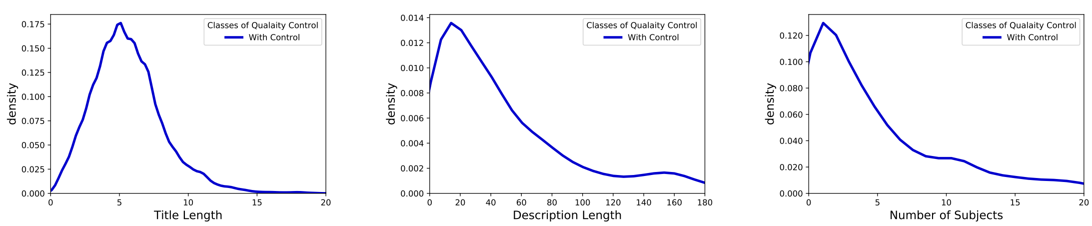


Figure 6.2: Proportion of manual OER quality control

As results of our exploratory data analysis, (1) we can use the OER metadata subset with already existing quality control to define quality benchmarks for metadata elements, and (2) it is desirable to define a method to facilitate the automatic assessment of OER metadata quality, and consequently the assessment of OER content quality. Therefore, as the next step in our analysis, we focused on OERs with quality control and screened the remaining metadata elements (i.e., title, description, and subjects) of these OERs:

- *Title* refers to the title given to an OER. Figure 6.3(a) shows the distribution of title length (as number of words).
- *Description* refers to the content summary of an OER. Figure 6.3(b) illustrates the distribution of description length (as number of words).
- *Subject* refers to the subjects (topics) which an OER addresses. Figure 6.3(c) shows the distribution of subjects (as number of subjects).

The plots in Figure 6.3 show that these features have distributions similar to normal. Therefore, it is possible to fit a normal distribution on them and build a scoring model based on the distribution parameters.



(a) Length of *Title* (word count)

(b) Length of *Description* (word count)

(c) Length of *Subject* (subject count)

Figure 6.3: Metadata analysis of quality controlled OER elements

Table 6.1: OER metadata fields and importance [26]

Type	**Importance Rate** [0-1]	**Normalized Importance Rate** [0-1]	**Rating Function** [0-1]
Title	1	0.17	$\frac{1}{\lceil \lvert x-5.5 \rvert 2.5 \rceil}$
Description	1	0.17	$\frac{1}{\lceil \lvert x-54.5 \rvert 40 \rceil}$
Subjects	0.86	0.145	$\frac{1}{\lceil \lvert x-4.5 \rvert 3.5 \rceil}$
Level	0.98	0.165	If available: 1; else: 0
Language	0.92	0.155	If available: 1; else: 0
Time Required	0.58	0.098	If available: 1; else: 0
Accessibilities	0.59	0.099	If available: 1; else: 0

6.1.3 OER Metadata Scoring Model

In order to build our scoring model, we started with the definition of the importance of each metadata field, and a rating function based on the quality-controlled OERs. Thus, we defined the importance rate of each metadata field based on their availability rate (between 0 and 1) among quality-controlled OERs [26]. For instance, we set the importance rate of the *description* field to 1 as this field was included in all quality-controlled OERs, and we set the importance rate to 0.58 for the *time required* field since *58%* of quality controlled OERs included this metadata field. Accordingly, we normalized the calculated importance rates as *normalized importance rate*.

Moreover, we created a rating function for each field based on quality-controlled OERs, in order to rate metadata values [26]. Regarding the fields *title, description*, and *subjects*, we fitted a normal distribution on their value length, as according to Figure 6.3, they have distributions similar to normal. Afterward, to rate the metadata values based on the properties of controlled OERs, we used the reverse of the Z-score concept [84] for each metadata value. Thus, the closer an OER *title/description/subject* length to the mean of the distributions of quality controlled OERs, the higher is the rate[6]. Regarding the four fields of *level, length, language,* and *accessibility*, we used a Boolean function, which assigns 1 when they have a value and assigns 0 otherwise. The output of these calculations is illustrated in Table 6.1.

Finally, to consider the defined benchmarks in evaluating the quality of OERs' metadata, we defined the following two scoring models [26]:

Availability Model [26]. OER availability score is calculated as (6.1). $norm\_import\_rate\,(k)$ is the *Normalized Importance Rate* of metadata field k. The output indicates the completeness of a given metadata in a weighted summation. The weights here are the normalized important rates. As a consequence, high availability score means that the metadata of a given OER consists of fields with significant importance. Consider an example, when a given $OER1$ has values for the following important metadata *title, description*, and *level*, while $OER2$ contains metadata for *subjects, language,*

[6] It should be mentioned that when a field value is *equal to the mean* or *empty*, the rate will be 1 or 0, respectively.

time required, and *accessibilities*. In our model, $OER1$ gets a higher availability score than $OER2$.

$$avail\_score\left(o\right) = \Sigma_{k=available\_fields} norm\_import\_rate\left(k\right) \tag{6.1}$$

Normal Model [26]. The normal score of an OER o is defined as (6.2): $norm\_import\_rate\left(k\right)$ is the *Normalized Importance Rate* of metadata field k, where *rating(o,k)* is the assigned rating to OER o regarding field k. This score is built on the rating function of the metadata field k. As a result, we can benchmark a given metadata to a predefined standard (In our case, we consider quality controlled OER metadata as the standard). This means that a given OER with similar metadata properties to a standard OER, will obtain a high normal score.

$$norm\_score\left(o\right) = \Sigma_{k=fields} norm\_import\_rate\left(k\right) * rating\left(o,k\right) \tag{6.2}$$

6.1.4 Predicting the quality of OERs based on their metadata

As the next step, we used our scoring models to build a machine learning model to predict the quality of OERs based on their metadata [26]. For this purpose, we extracted *4,651* OERs *with quality control* and classified them as high-quality OER, while labeling the remaining *4,236* OERs as *low quality* OER. Subsequently, we trained a Random Forest classifier on the *SkillsCommons* dataset to build a model that makes a binary decision: *high-quality/low-quality*. We used *80%* of the data as the training set and the remaining *20%* as the test set. As a result, the classifier achieved a *94.6%* F1-score when classifying OERs into one of the two above-mentioned categories[7]. Furthermore, we extracted the importance value (i.e. effect) of each feature on our classification model as: *Availability Score*: 0.32, *Normal Score*: 0.25, *Level Metadata Availability*: 0.23, *Description Length*: 0.10, *Title Length*: 0.05, *Subjects Length*: 0.05.

6.2 Validation

In this section, we report the results of applying our scoring and prediction models on our *Youtube* dataset, including the metadata of *884* educational videos in *32* subjects in the areas of *Information Technology* and *Health Care*. First, we applied our scoring and prediction models on the dataset to classify the videos into two groups: *with control* (higher quality) and *without control* (lower quality)[8].

After classification, we got *477* videos *with control* and *407* videos *without control*. Then, we needed to identify a metric in their metadata to compare the two groups in order to check whether our model detects the groups of videos with higher quality or not. Therefore, we decided to focus

[7]We implemented this classifier in Python. Our steps and results are publicly available on: `https://github.com/rezatavakoli/ICALT2020_metadata`

[8]In order to apply our model, we set our required fields based on the video properties. For instance, we set *level availability* based on the video's title, and set *length availability* to "available" as all videos have length metadata.

Table 6.2: Difference between videos rating of groups

Subject	**Rating Difference**	**Difference Sign**
bioethics	0.15	+
deep learning	-0.15	-
infectious disease	0.14	+
sleep disorder	-0.14	-
apache spark	0.13	+
data mining	0.10	+
allergies	0.09	+
vaccinations	0.08	+
women and nutrition	-0.08	-
data management	0.07	+
SQL language	-0.06	-
brain tumors	0.05	+
big data	0.05	+
cancer prevention	0.05	+
data cleaning	0.05	+
sun awareness	0.05	+
addiction	0.05	+
data visualization	0.04	+
psychology	0.03	+
neural network	0.03	+
apache hadoop	0.03	+
stress management	0.02	+
tensorflow	0.02	+
obesity care	0.02	+
python language	0.02	+
R language	0.02	+
statistics	0.02	+
text mining	0.02	+
machine learning	0.01	+
prostate cancer	0.01	+
eye care	0.01	+
smoking health risks	-0.01	-
Average	**0.05**	**+**

on the video *rating* feature as a quality indicator from the users' perspective, which is calculated based on *likes* and *dislikes*, and one of the most commonly used metrics of quality assessment of videos [85]. Finally, for each of the *32* subjects, we calculated the average of video ratings for each of the predicted groups (*with control* as higher quality and *without control* as lower quality). Table 6.2 shows the subjects, the difference of the average rating between the groups, and the difference sign which specifies whether our model predicted correctly (the "with control" group has higher ratings (shows with +)) or not (shows with -).

As per the results detected by our prediction model, the average rating in a group with higher quality has 0.05 higher video rating than the lower quality group. This is very reasonable considering the standard deviation of ratings in the dataset of *0.25*. To further elaborate, the maximum difference between around *80%* of the ratings is *0.25*. Therefore, dividing them into two groups with a rating difference of *0.05*, emphasizes that our classifier works well in this context. Additionally, in *27* out of *32* subjects (**84.3%**), where our model detected higher quality groups, they had higher ratings.

6.3 Discussion

In this part, we analyze the results we have achieved through 1. collecting educational resources 2. exploratory data analysis on the collected resources, 3. creating metadata scoring, and 4. building a metadata-based quality prediction model.

OER Metadata. Based on the exploratory analysis on our OER dataset, it is clear that there is a strong relationship between OER quality control and metadata quality. Therefore, the more an OER passes the quality control process, the higher the chance of including high-quality metadata is. Accordingly, we can define benchmarks for metadata quality by analyzing the controlled OERs. On the other hand, using metadata quality as a proxy for OER content quality can be beneficial in developing automatic quality control processes for OERs. According to the analysis of quality-controlled OERs, *Title* and *Description* metadata play a key role in publishing OERs, as all of the controlled OERs contain these two fields in their metadata. Moreover, more than *85%* of the controlled OERs include metadata regarding *Language*, *Level*, and *Subject* which shows the importance of these three fields in defining OERs.

Metadata Scoring. Analyzing the importance values in our quality prediction model reveals the effectiveness of our proposed scores for metadata, as the *Random Forest* model assigns the highest value to our *Availability Score* and *Normal Score* features. Therefore, these two proposed indicators illustrate the quality of OER metadata well and can be applied not only for metadata scoring, but also for OER content quality prediction.

Quality Prediction Model. The F1-score of our proposed prediction model (**94.6%**) shows that we can accurately predict the quality of OERs in *SkillsCommons* repository. Our validation step on *Youtube* dataset also supports the generalizability of our model, which can be applied in different repositories and various types of educational resources (e.g. videos, text-based). Moreover, according to the result of our validation step, as our prediction model detected the higher quality groups in *14* (out of *16*) *Information Technology* topics and in *13* (out of *16*) *Health Care* topics, the proposed *Random Forest* prediction model works well in different topic areas.

6.4 Outcome of Metadata-base Scoring and Quality Prediction of Educational Resources

In this study, we used the metadata of a large OER dataset to analyze OER metadata quality and OER quality control processes. Based on our analysis, we created a prediction model to evaluate the quality of OER metadata and as a consequence OER content quality. We deem that our proposed model not only helps OER providers to revisit and think about the importance of the quality of their metadata, but also facilitates the process of OER quality control in general, which is essential in light of the rapidly growing number of OERs. Applying our quality prediction model on the *Skillscommons* dataset showed that it can detect quality-controlled OERs with the F1-score of **94.6%**. We also validated our approach in another context, by applying our scoring and prediction model to open educational videos on *Youtube*. The results show that our approach successfully detects videos with higher user rating values. The validation step indicates that our approach can be used on different OER repositories.

CHAPTER 7

Topic Analysis of Knowledge Areas and Educational Resources[1]

The previous chapter revealed the importance of OER metadata for building educational services. Particularly, OER metadata about covered topics (subjects) is essentially required by learners to build effective learning pathways toward their individual learning objectives. Therefore, in this chapter, we report on a project proposing an OER topic extraction approach by applying text mining techniques. To do this, we illustrate our steps toward building our models including data collection, pre-processing, model training, and validation.

7.1 Data Collection and Pre-processing

In this section, we describe the data that were collected in order to build our topic extraction approach.

Target Skills. To propose the first version of our approach, we extracted important skills for data science by mining relevant job vacancies [4]. In our online job vacancy dataset (from August 2019 to December 2019[2]) the three most important data science skills were 1. *Machine Learning*, 2. *Text Mining*, and 3. *SQL Language*.

OER Resources for Building Topic Models. In order to build our topic models, we collected 123 relevant online lectures (and their transcripts) from *Coursera*[3] and *Khan Academy*[4] related to our target skills (including 67 lectures for machine learning, 27 for text mining, and 29 for sql language).

[1]This chapter has been published as follows: Molavi, M., Tavakoli, M. and Kismihók, G., 2020, September. Extracting Topics from Open Educational Resources. In European Conference on Technology Enhanced Learning (pp. 455-460). Springer, Cham.

[2]the dataset will be added after the blind reviews

[3]https://www.coursera.org/

[4]https://www.khanacademy.org/

OER Resources for Evaluating our Models. To evaluate our proposed model, we used the dataset defined by [82] including 550 educational videos and their properties (e.g. rate, transcript, view-count) from *Youtube/TIB-AV-Portal* in the area of data science.

Pre-processing of OER Resources. Before building the topic extraction model, we applied the following data pre-processing steps on our collected OER transcripts from *Coursera*, *Khan Academy*, and *Youtube/TIB-AV-Portal* to prepare them for our analysis:

- Removal of
  - unimportant characters
  - punctuations
  - links
  - stop words
- Building TF-IDF representation

7.2 Method

In this section, we showcase our process to build and evaluate our topic extraction model. First, we explain the steps towards building topic extraction models for each target skill, followed by the evaluation of each created model.

7.2.1 Building Topic Models

To extract knowledge areas that are covered by particular educational resources and related skills, we used *Latent Dirichlet Allocation* (*LDA*) [68]. *LDA* is a generative probabilistic topic model that considers each document as a distribution of different topics, each topic as a distribution of different words, and tries to extract existing topics together with their distribution of words for a corpus. To set the number of topics that *LDA* extracts, we calculated C_V *Coherence* [86] for different number of topics (between 2 to 10), and selected the topic amount with the highest coherence value. The following sections explain the process of finding the most appropriate value of topic amounts, extracting topics, and assigning a name to topics (done manually, after executing *LDA* and based on the distribution of their words [87]).

Machine Learning. Figure 7.1 illustrates C_V coherence for different numbers of topics (from 2 to 10) in machine learning related educational materials. As you can see, *9* topics lead to the best coherence in topics. Therefore, we set the parameter k of *LDA* to *9* and executed it on the machine learning corpus. Table 7.1 shows the result of the extracted topics, the assigned name for each topic, and *10* selected words in each topic (among the top 20 most important words and without changing the order), which are related to the assigned name.

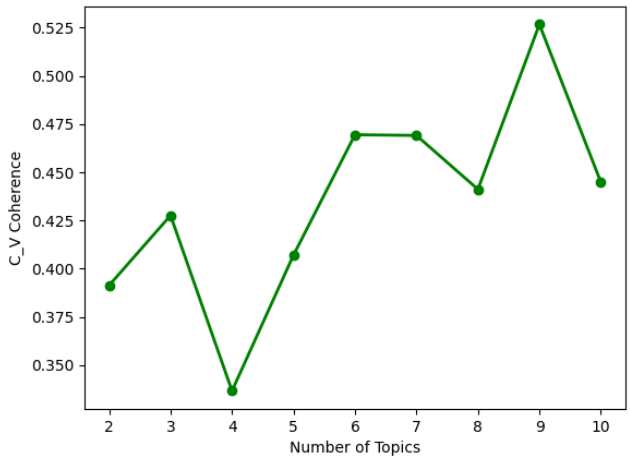


Figure 7.1: C_V coherence for different number of topics in the Machine Learning corpus

Table 7.1: Output of LDA on Machine Learning corpus

Topics	Assigned Name	Significant Words
$Topic_1$	Clustering	cluster local centroid mu step optima minimum means random initialization
$Topic_2$	Dimensionality Reduction (PCA)	approx pca representation dimensional map reduce compression projection vector point
$Topic_3$	Classification	classification contours feature class spam algorithms theta scaling threshold regression
$Topic_4$	Matrix Factorization (SVD)	matrix pca svd variance sigma covariance columns projection reduce diagonal
$Topic_5$	Neural Network	layer hidden neural network forward unit architecture activation propagation vector
$Topic_6$	Overview	network cluster neural similarity centroid regression svm feature logistic linear
$Topic_7$	Regression	theta features fit predict polynomial regularization vector regression line overfitting
$Topic_8$	Linear Algebra	inverse transpose algebra features pseudo matrices invertibility singular dependent decimals
$Topic_9$	Neural Network Classifier	propagation gradient descent cost layer units hidden network neural classifier

Text Mining. For the text mining skill, we calculated C_V coherence for different number of topics on text mining-related educational resources, as shown in Figure 7.2. Based on the result, 7 topics provide us with the best coherence value. Therefore, we set the parameter k of *LDA* to 7 and executed the analysis on our text mining corpus. The result is visible in Table 7.2.

Table 7.2: Output of LDA on Text Mining corpus

Topics	Assigned Name	Significant Words
$Topic_1$	Topic Modeling	lda plsa topic dirichlet parameters likelihood beta distribution alpha document
$Topic_2$	Sequence Models	prior string tag sequence markov hidden probabilities estimate position generate
$Topic_3$	Sentiment Analysis	features grams sentiment positive topic accuracy reviews negative idf tf
$Topic_4$	Matrix Factorization	matrix topic matrices squared diagonal svd factorization vectors approximation document
$Topic_5$	Text Classification	grams convolutional filter sentence corpus vec embeddings neural count modeling
$Topic_6$	Probabilistic Models	naive prior bayes given probability likelihood independent maximizes predicted significantly
$Topic_7$	Feature Extraction	sentence document frequency features phrase vec grammar parse grams term

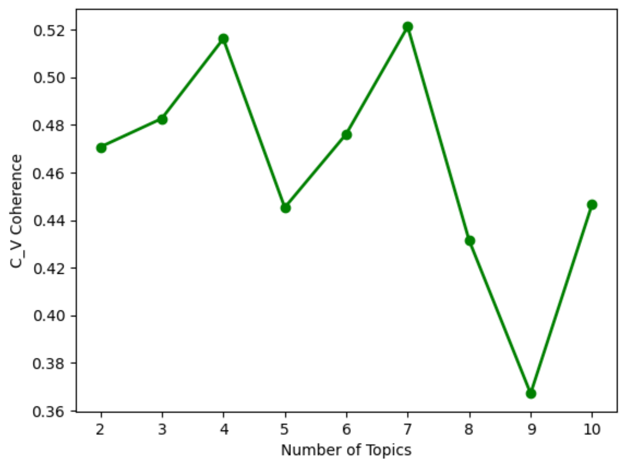


Figure 7.2: C_V coherence for different number of topics in Text Mining corpus

SQL Language. Figure 7.3 shows the output of calculating C_V coherence for different number of topics (from 2 to 10) in SQL language educational materials. As a result, *5* topics provide the best coherence in topics. Accordingly, we set the LDA parameter k to *5*. Table 7.3 illustrates the result of applying *LDA* on the SQL language corpus.

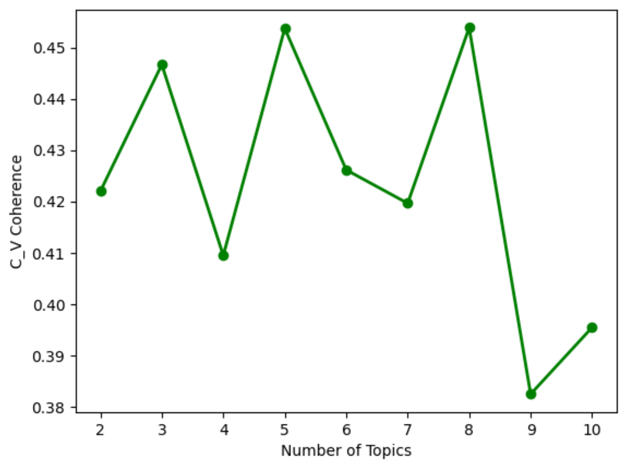


Figure 7.3: C_V coherence for different number of topics in SQL Language corpus

Table 7.3: Output of LDA on SQL Language corpus

Topics	Assigned Name	Significant Words
$Topic_1$	Basic Operation	union limit values retrieving conditions operator top statement greater select
$Topic_2$	Join Introduction	join cartesian cross inner filter criteria subqueries clause records product
$Topic_3$	Advanced Join	join left outer subset alias full table match relational right
$Topic_4$	Sort	sort descending ascending change join tables query multiple order record
$Topic_5$	Aggregate Functions	aggregate distinct count function average values max sum min avg

7.2.2 Topic Model Extraction Evaluation

To evaluate our topic models, we used our Youtube dataset in which topics were assigned to videos manually. This manual assignment was done by 3 data science experts with at least 2 years of teaching experience and 5 years of industrial experience in data science related areas. It should be mentioned that each participant allocated at least 2 minutes for analyzing each of the videos. Afterward, we applied our topic extraction models on each video transcript (e.g. apply our machine learning topic model on the related educational videos on machine learning). Finally, we compared the manually assigned topics (by experts) and the output of our topic extraction models. As a result, we were able to determine the quality of our topic extraction models in relation to manual, expert topic assignments. Table 7.4 illustrates the *F1-score* of each topic extraction model. Our model was able to extract covered topics of educational resources with an F1-score of **79%** on average.

Table 7.4: Accuracy of Topic Models in each target skill

Target Skill	F1-score
Machine Learning	81%
Text Mining	76%
SQL language	78%
Avg	**79%**

7.3 Validation

As the second step of validation, we integrated our model into an OER recommender system [4] and asked 8 data science experts to rate (between 1 to 5) the quality of topic extraction for the recommended OERs. A screenshot of the website regarding our topic extraction model is shown in Figure 7.4. For this step, we excluded the experts, who did the manual topic assignment in the previous step. The

8 experts involved in this step had at least 1 year of teaching experience and 3 years of industrial experience in data science related fields. In the end, more than 120 recommended OERs were rated. Table 7.5 shows the percentage for each of the rates. As you can see, experts' satisfaction was *3.66* out of *5* (**73.2%**), showing that our proposed topic extraction for OERs approach works well.

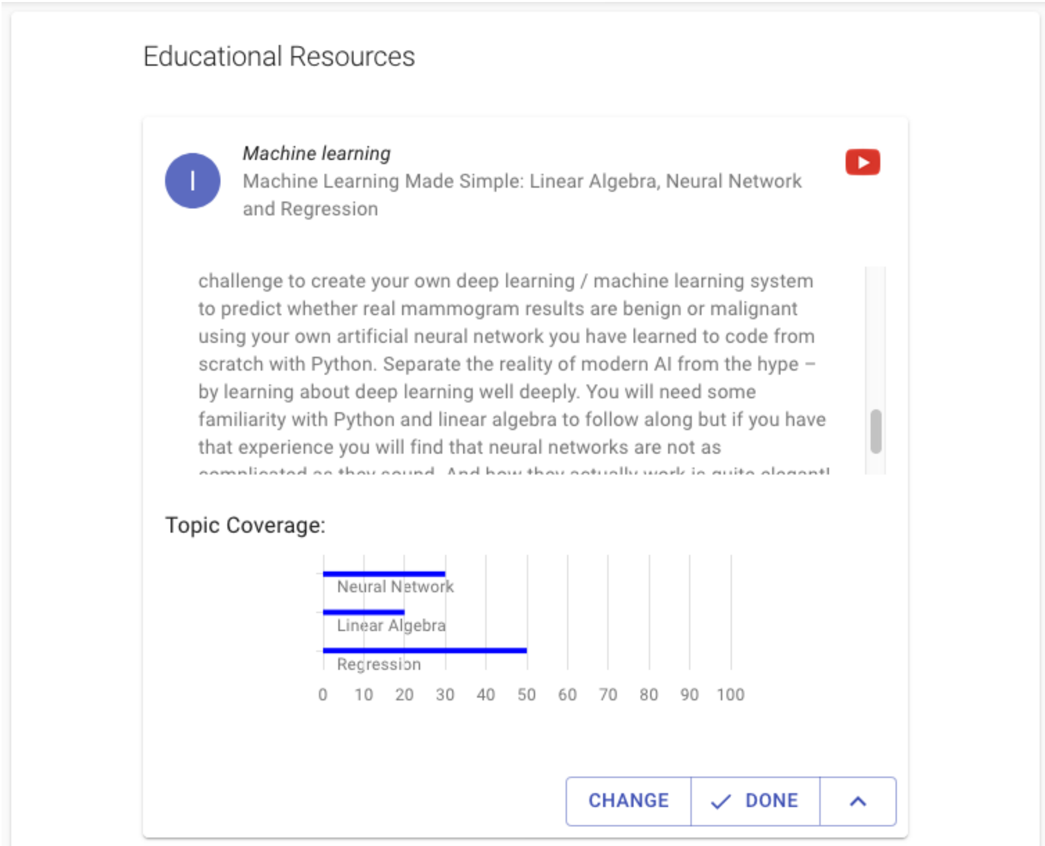


Figure 7.4: Screenshot of the website regarding our Topic Models

Table 7.5: Result of the Validation in the OER Recommender

Target Skill	Rate=1 (%)	Rate=2(%)	Rate=3(%)	Rate=4(%)	Rate=5(%)	Average
Machine Learning	6	9	21	33	31	3.74
Text Mining	6	6	31	29	28	3.67
SQL Language	8	6	30	32	24	3.58
Average	**≈7**	**≈7**	**≈27**	**≈31**	**≈28**	**3.66**

7.4 Outcome of Topic Analysis of Knowledge Areas and Educational Resources

This study is one of the steps towards 1) dynamic definition of topics that should be covered by particular knowledge areas in open educational content, and 2) extracting the topic distribution for a given OER, as one of the most important metadata, to help learners to build their own learning path. We collected 123 educational lectures from two repositories related to 3 data science related skills. After that, we applied *LDA* on the lectures' transcripts to extract the topic model for each skill. Finally, to evaluate the models, we used an educational Youtube dataset, and assigned covered topics with the help of 3 data science experts. Subsequently, we applied our topic extraction models, and compared the output of our model with the manually assigned topics. This exercise revealed that our models can extract topics with *F1-score* of **79%**. Moreover, as another validation step, we integrated these topic extraction models into an OER recommender system, and asked 8 data science experts to rate their satisfaction regarding the outcome of our model for the recommended OERs. The results showed that experts were satisfied *3.66* out 5 (**73.2%**) on average from more than 120 recommended OERs.

CHAPTER 8

Hybrid Human-AI Curriculum Development[1]

By investigating what we did in the previous steps, we concluded that although our models were promising and worth using, they are not 100% accurate. Therefore, expert opinions need to be used for having high-quality accurate curricula for learners. At the same time, encouraging experts to develop high-quality up-to-date curricula is difficult as it is an extremely time-consuming and complex process. Accordingly, we decided to use our developed intelligent models and build a system, which helps experts in creating curricula. Therefore, in this chapter, we propose an *Artificial Intelligence* (*AI*) and *Crowdsourcing* based approach to create and update curricula for informal learners. We show the design of this curriculum development system prototype, in which contributors receive AI-based recommendations to be able to define and update *high-level learning goals*, *skills*, and *learning topics* together with associated learning content. At the end of this chapter, our validation strategy for evaluating our developed system is also covered.

8.1 Curriculum Development Framework

We defined four main components for our framework: 1. *high-level learning goals* which consist of skills, 2. *skills* which consist of learning topics, 3. *learning topics* which consist of educational packages, and 4. *educational packages* which include one or more educational resources. To help the contributors manage their content, each of these components is enhanced with the following services: 1. Add service for defining a component for the first time, 2. Crowd management service for collecting users' suggestions on a component, and also automatically accepting or rejecting them based on crowd opinions, 3. Recommendation service for providing insights to contributors based on existing open

[1]This chapter has been published as follows: Tavakoli, M., Faraji, A., Molavi, M., Mol, S.T. and Kismihók, G., 2022. Hybrid Human-AI Curriculum Development for Personalised Informal Learning Environments. In LAK22: 12th International Learning Analytics and Knowledge Conference.

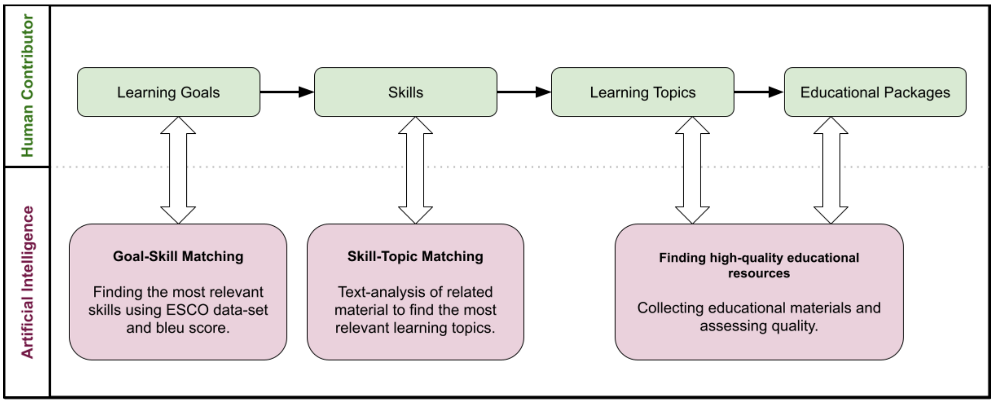


Figure 8.1: The conceptual model of our curriculum development framework

data (e.g. job vacancies, educational resources, and standard taxonomies) in order to help them add and update content. Figure 8.1 shows the conceptual model of our curriculum development framework.

At first, we explain the details of designing and implementing our system, and subsequently show a first use case on an existing personalized educational platform.

8.1.1 Managing High-level Learning Goals

For adding high-level learning goals (consisting of multiple skills), we collected *titles* and optional *descriptions* together with the following key contextual features to help learners [4]: *industry, company, city*, and *country*. However, contributors can decide not to contextualize learning objectives by keeping each of these fields as *General*. Afterward, our system recommends a list of goal-related skills based on the title of the high-level goal, in order to capture the necessary skills to master for the given learning goal. For these recommendations, we rely on the *ESCO* [2] dataset, which is a European multilingual classification system of skills, competencies, and occupations. ESCO is continuously updated by subject matter experts of the European Commission and includes 13,485 skills, linked to 2942 occupations. We match the title of the high-level goal with existing occupations in *ESCO* using the *Bleu score* [88] concept. We decided to use *Bleu score* as it allows us to capture the closest term (occupation in this case) in the *ESCO* dataset no matter in what language the title is. Finally, we recommend a list of skills linked to the closest occupation to the contributor. Contributors can either select skills from the list, or add new skills manually. After finalizing the skill list, the contributor can sort skills based on the order they need to be shown to learners. Figure 8.2 shows a screenshot about adding a high-level goal to our system.

After adding a high-level goal, users can view the page[3] of the newly added high-level goal, which includes:

[2]European Skills/Competences, qualifications, and Occupations: https://ec.europa.eu/esco/portal/home
[3]Figure: https://raw.githubusercontent.com/ali-faraji90/edoer/main/LAK22/Goal.png

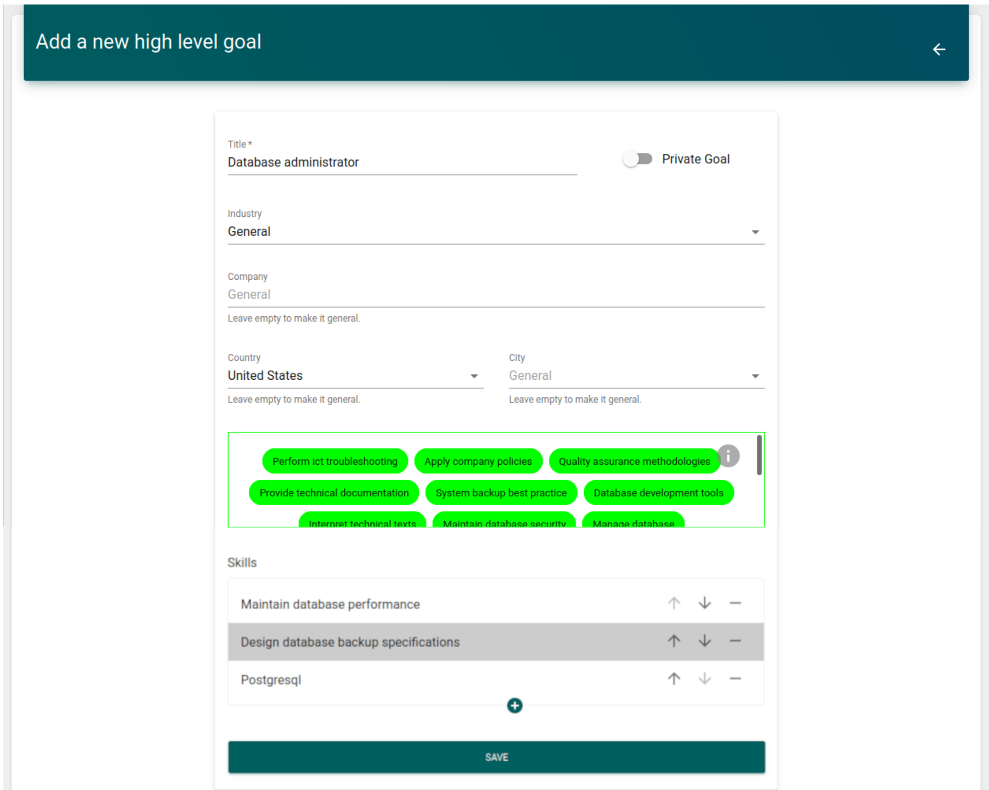


Figure 8.2: Screenshot of adding a high-level learning goal

- ***The list of skills associated with the goal.***

- ***Opinion of other users (crowd) regarding the importance of a particular skill in relation to this goal.*** At each skill, users can see the importance of the particular skill for that particular high-level goal, based on up and down-votes (see 8.1.5) of the crowd.

- ***The suggestion list.*** After defining a high-level goal for the first time, editing can be done by addressing the crowd's opinion. Therefore, all users can suggest: 1. adding a skill, 2. deleting a skill, and 3. reordering skills for each high-level goal. To provide insights for users and to keep each goal updated, our system recommends other, potentially related skills using the aforementioned algorithm. After adding a suggestion, other users can provide their opinion about a particular suggestion by up and down-voting. Our system captures these actions and suggests a decision whether to reject or accept the suggestion (see 8.1.5)

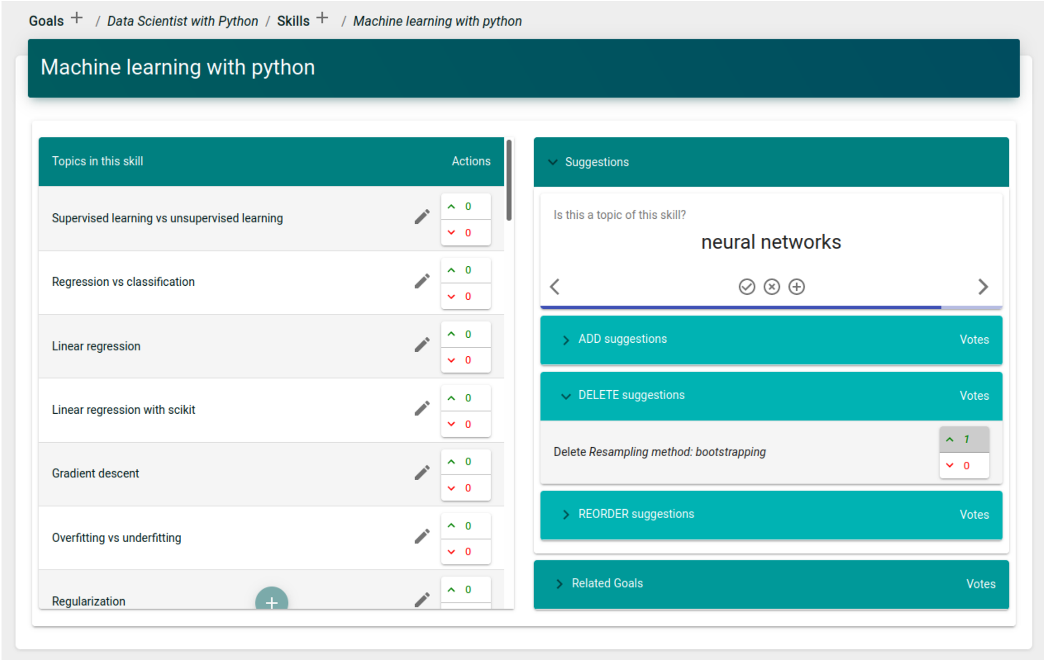


Figure 8.3: Screenshot from a skill page

8.1.2 Managing Skills

Defining Skills

In order to define a skill, contributors receive string auto-completions as they are typing the title of a skill. This function is based on existing skills in the *ESCO* standards. Auto-completion not only facilitates skill title definition, but also helps contributors to refrain from adding different titles for a single skill. After filling out the title and description fields, our system uses *Youtube* playlists in order to provide insight for contributors regarding existing *learning topics* on any new skill. The system automatically searches the skill title, collects related videos, and extracts the most important keywords by applying *TF-IDF* [89] algorithm on the collected video titles[4]. The contributor can either select from this AI-generated recommendation list or type in learning topics manually. Again, after finalizing the topic list, the contributor can sort the topics into an order for learners. As depicted in Figure 8.3 after defining the skill, a skill page will be created including: 1. the list of learning topics, 2. crowd opinion about the learning topic's importance (see 8.1.5), 3. the suggestion list including *system recommended learning topics* (see 8.1.2), *adding, deleting,* and *reordering* suggestions, which are monitored by our suggestion reviewing process (see 8.1.5), and 4. high-level learning goals associated with this skill.

[4]Figure: https://raw.githubusercontent.com/ali-faraji90/edoer/main/LAK22/Add-skill.png

Editing Skills

After adding a skill, similar to high-level goals, the crowd has the ability to edit its content. In order to contribute to the skill updating process, we provide some learning topic recommendations that other users can easily add as a suggestion (see Figure 8.3). These topic recommendations are generated as follows:

1. ***Collecting content for learning topics.*** We search each new learning topic title on *Youtube*[5], and collect at least 50 educational video transcripts labeled with that particular topic title. This exercise results in a two-column dataset including transcripts of educational resources in the first column, and 2. their associated learning topic in the second column.

2. ***Text pre-processing.*** On each collected transcript, we apply standard text pre-processing steps including *converting to lower case, removing special characters, stemming,* and *adding n-grams (bi/tri-grams)*.

3. ***Applying labelled topic modeling.*** We use *LLDA* (Labeled Latent Dirichlet Allocation) to extract topics from the transcripts. *LLDA* is a supervised version of *LDA* [68], which is a generative probabilistic topic model aiming at extracting the most important keywords for each topic label. We collect the ten most important keywords, and their probabilities for each topic and store them in the *recommendation list*

4. ***Collecting content for skills.*** The *recommendation list* we generated through topic modeling of learning topics is usually not exhaustive enough, as some skill-related topics may not be represented well on the list. Therefore, we also perform a search for the title of the given skill on *Youtube* playlists and collect at least 200 videos for each skill.

5. ***Applying* LDA *model on skill related learning resources.*** As for these resources we do not have any topic labels, we use an *LDA* model, an unsupervised machine learning method, which considers each transcript as a distribution of different topics, each topic as a distribution of different words, and aims to extract existing topics together with their distribution of words. To find the number of learning topics related to a skill, which is the input of the *LDA* model (as parameter k), we calculate C_V *Coherence* [86] for different number of topics (between 2 to 50), and set k with the topic amount that results in the highest coherence value. Again, for each newly extracted topic, we extract the ten most important words (and their probabilities) and add them to the previous words in our *recommendation list*.

6. ***Removing the existing topics.*** In order to prevent recommending learning topics that are already associated with a skill, we remove those, already associated, topics from our recommendation list.

[5] Using the *Pafy* library: https://pypi.org/project/pafy/

7. ***Sorting the recommendation list.*** To capture the most relevant topics quickly, we sort the words on the recommendation list according to their probability values in our topic models, and start recommending from the first item on the list.

8.1.3 Managing Learning Topics

Contributors can add learning topics by defining their titles and description first. Afterward, a topic page is created and contributors can start adding educational materials to a given topic. To facilitate adding an educational package to each topic, we collect a list of educational materials from *Youtube* and *Wikipedia*[6] based on the title of the topic. After that, each collected educational material passes through the following automatic quality control process:

1. ***Metadata based quality control.*** [90] showed a close relationship between the metadata quality and the content quality of educational resources. Accordingly, they built an openly available quality prediction model that detects the quality of educational materials based on their metadata. We adopted this approach to filter out low-quality educational materials. Furthermore, based on this quality prediction model, we also calculate a quality score, which is shown to users for each recommended educational resource.

2. ***Content based quality control.*** To predict the relevancy of a particular educational resource to a learning topic, and, at the same time, to filter out irrelevant educational resources, we use the probability values of each word in the *LLDA* model discussed in the previous section (see 8.1.2).

As recommended educational resources go through this automatic quality control process, contributors can directly associate those resources with learning topics. However, users can still provide suggestions, for instance decoupling an educational resource from a topic, when they think it is not an appropriate content (not relevant or it has low quality)[7]. Moreover, users can also have an overview of skills that require knowledge of that particular learning topic.

8.1.4 Managing Educational Packages

Contributors can also define educational packages including one or more educational resources. To do this, firstly, contributors need to name a title and an optional description for an educational package. Afterward, they can add one or more educational resources either by importing from a *URL*, or 2. uploading an educational resource. Subsequently, contributors can fine-tune and set properties (i.e. title, description, format type, estimated time needed to complete the resource, source of the content, includes an example/theory or not, level of details, and if it is a recording of a class-based instruction or not) for each of the resources in the package. It should be mentioned that in order to facilitate the

[6] Using Wikipedia library: `https://pypi.org/project/wikipedia/`
[7] Figure: `https://raw.githubusercontent.com/ali-faraji90/edoer/main/LAK22/Topic.png`

property setting process, we implemented a property extractor component that retrieves information from *Youtube* educational videos and *Wikipedia*. Finally, our system recommends related learning topics to contributors (using the LLDA model for the topics (see 8.1.2)) in order to help them set the learning topics that are covered by the new educational package.

8.1.5 Crowd's Opinion Management

In this part, we will explain our *voting system, how contributors receive points*, and the *process of suggestions' reviewing*.

Voting System. Users can up and down-vote skills, learning topics, and educational packages based on their perceived importance to their containing component. Therefore, with this feature, contributors are able to show how important 1. a skill for a high-level goal, 2. a learning topic for a skill, and 3. an educational package for a learning topic. Also, users can up/down-vote new suggestions (i.e. adding, deleting, and reordering suggestions) and give their achieved points in the context of the suggestion (see *Achieving Points* below) to help a suggestion be accepted or rejected. By using this mechanism of voting, the system puts more weight on reliable users in the suggestion reviewing process (see *Reviewing Process of Suggestions* below).

Achieving Points. Contributors can collect points on each skill and also each learning topic. They can collect a point if their defined skills and topics receive an up-vote or if they are added as a learning goal to others' profiles. Moreover, if the educational material they added receives up-votes from learners, they also receive a point for the target topic(s) of that educational material.

Reviewing Process of Suggestions. Each and every suggestion needs to receive a minimum number of points to be approved in the system (currently ten points, but it can be changed in our configuration files) within a predefined period (currently a week, however, it can also be changed in our configuration files), otherwise the suggestion will be automatically rejected. When a suggestion receives the minimum required points for approval, the system calculates the rate of positive received points to all received points. If the rate is greater than 75% (again customizable in our configuration), the suggestion is automatically accepted, otherwise rejected.

8.1.6 Use-case: Personalized, Goal-Driven Learning Recommendations

Our personalized learning system, *eDoer*[8], is a platform in which learners 1. set their learning goals, 2. receive skill lists related to their target goals, 3. select skills from the list, what they want to master, 4. generate their learning dashboard, which includes all selected skills and associated learning topics, and 5. receive personalized educational resources based on their preferences and behavior history on the platform[9]. We used this educational platform as the first use case for our AI-based

[8] www.edoer.eu

[9] You can find a screenshot of our learning dashboard here: https://raw.githubusercontent.com/ali-faraji90/edoer/main/Files/Curriculum.png

crowdsourcing system, to investigate on one hand how this technology aids contributors to define and maintain educational content, and on the other hand, how it empowers learners by providing up-to-date personalized curricula.

8.2 Validation

To evaluate the system, we decided to measure the accuracy of our recommendations during the authoring process in the domain of *Data Science*. Moreover, we designed a semi-structured subject matter expert interview protocol, to evaluate the authoring process with users. In this section, we will showcase the results of these efforts.

8.2.1 Recommendation Accuracy

To evaluate our recommendations, we asked three experts with at least five years of academic and ten years of industrial experience in the area of *Data Science* to manually (without using our platform) specify the following high-level goals: *Data Scientist*, *Data Analyst*, and *Business Analyst* together with related skills and learning topics. Subsequently, they used our system to define these jobs with our skill recommendations (see 8.1.1). The list of skills produced by the experts was then compared with the list of recommendations from our system. The results showed that our system, on average, had an F1-score of 89% when it comes to recommending relevant skills for high-level goals. When it comes to learning topic recommendation for data science related skills (see 8.1.2), our evaluation resulted in the following F1-scores: *Python programming*: 83%, *Machine learning*: 76%, *Statistics*: 79%, *Data visualization*: 75%, which meant 79% as a weighted average.

Finally, when recommending specific learning content for learning topics (see 8.1.3), experts examined their validity. Only if the recommended content was marked as high-quality and relevant to the topic, we considered it a valid recommendation. This evaluation revealed that our educational content recommendation method provided high-quality relevant materials in 93% of time[10].

8.2.2 Subject Matter Evaluation

In order to evaluate the proposed authoring process, we designed and executed interviews with 8 senior members (i.e. managers, professors, associate professors, and researchers) from different organizations: Participant_1 from *Ericsson* company in Sweden, Participant_2 from the *University of Amsterdam*, Participant_3 from *KU Leuven University*, Participant_4 from *TIB - Leibniz Information Centre for Science and Technology*, Participant_5 from the *University of Bonn*, Participant_6 from *Netherlands AI Coalition*, Participant_7 from *the American Psychological Association*, and Participant_8 from

[10] As different curricula are continuously developed in our system, to receive an up-to-date dataset of the existing components and recommendations, contact the authors.

Career and Life Planning (CALP). The structure of the interviews was as follows: 1. Introducing the system and its logic (~20 minutes), 2. using the authoring system and its features for about 20 minutes, and 3. going through a semi-structured interview with the assistance of a questionnaire[11] (~20 minutes).

Participant_1 and participant_2 mentioned that providing an environment in which learners can be informed about when and how to build their careers is the most important part of the system. Also, Participant_3 considered providing insights for authors regarding the relevant skills, learning topics, and educational materials as a key feature of our system. Participant_4, Participants_5 and Participant_7 emphasized that combining AI with crowdsourcing can improve the quality of each recommendation component, and therefore this is the most promising function of the system. However, they suggested to provide as much clarity as possible for users on how our *AI* and *crowd-management* components collaborate with each other. Participant_6, besides pointing at the usefulness of our system when it comes to matching jobs and their required knowledge dynamically, mentioned that integrating the ability for testing knowledge should be part of our future steps. Participant_8 believed the goals page and curriculum are the most important elements of our system and thought enriching these parts should have a priority in our future work.

Ultimately, *100%* of the interviewees agreed that "creating dynamic personalized curricula for learners" as our objective is extremely important and timely. Also, only *12.5%* (1 out of 8) was unsatisfied with the usability of our prototype system, which shows that most of the participants consider our prototype usable already.

8.3 Outcome of Hybrid Human-AI Curriculum Development

Throughout this chapter, we showcased a novel learning content authoring system, which helps contributors to build personalized curricula for learners. By providing intelligent recommendations, this system aids contributors to define 1. high-level learning goals consisting of skills, 2. skills built by learning topics, and 3. learning topics with related educational materials. We believe that such a system not only helps contributors to define and maintain their educational content, but also empowers learners through setting their own learning objectives, receiving personalized recommendations, and being up-to-date on desirable knowledge. Evaluating our recommendations in the context of data science showed that our system can provide 89% F1-accuracy in matching high-level goals and their skills, 79% F1-accuracy matching skills and their learning topics, and 93% precision in recommending high-quality, relevant educational materials. Moreover, we validated the main objective and usability of our system by interviewing eight subject matter experts in the area of education, which showed that they were satisfied with the objective and usability of our proposed system.

[11] The questionnaire is available here: `https://forms.gle/A17KHhWGoUH9WFuc8`

CHAPTER 9

An AI-based Open Recommender System for Goal-driven Personalized Education[1]

After developing the required components, we started building our final system, called *eDoer*, which combines all the components to offer scalable, open personalized education for learners. In this chapter, we discuss the complete system development cycle starting with a systematic user requirements gathering, followed by system design, implementation, and validation. Our recommender system 1) derives the skill (course) requirements for particular occupations (high-level goals) through an analysis of online job vacancy announcements; 2) decomposes skills into learning topics; 3) collects a variety of open online educational resources that address those topics; 4) checks the quality of those resources and topic relevance with three intelligent prediction models; 5) helps learners to set their learning goals towards their desired job-related skills; 6) recommends personalized learning pathways and learning content based on individual learning goals; and 7) provides assessment services for learners to monitor their progress towards their desired learning objectives. At the end of this chapter, we showcase the evaluation process of our system through a randomized experiment.

9.1 Method

Here, we illustrate our steps toward building our goal-driven personalized learning environment including 1. the requirement analysis, 2. labor market intelligence to match goals with their required skills, 3. decomposing target skills into educational topics, 4. collecting and curating educational resources for the educational topics, 5. building the personalized educational resource recommendation

[1] This chapter has been published as follows: Tavakoli, M., Faraji, A., Vrolijk, J., Molavi, M., Mol, S.T. and Kismihók, G., 2022. An AI-based open recommender system for personalized labor market driven education. Advanced Engineering Informatics, 52, p.101508.

part, and 6. implementing the learning dashboard which offers the developed components to learners.

9.1.1 Requirement Analysis

First, we collected relevant stakeholder requirements to further define our objectives and guide our investigation. For this exercise, we built an initial and bare-bones OER recommender prototype so as to be able to showcase our approach to key stakeholders. Through qualitative interviews, this prototype was evaluated by 23 subject matter experts (e.g. university instructors and Ph.D. students) with significant experience in both industry and learning/teaching [4, 12].

Based on their feedback, we designed a questionnaire[2] to capture the needs of those stakeholder groups that we expected to be potentially important beneficiaries of our learning recommender system. We identified the following stakeholder groups (personas) [3]:

- **Group1.** Recipients (e.g. Learners, Researchers, Students)
- **Group2.** Deliverers (e.g. Professors, Lecturers, Study Counselors)
- **Group3.** Facilitators (e.g. Managers, Educational Support Staffs)

We obtained 13 potential user requirements from the initial qualitative interviews, which we then presented to survey participants (see Table 9.1), asking the latter to rate those in terms of their importance and frequency of use. Since in this study we focus solely on the learner perspective, the following subsections showcase the most important outcomes and findings from **Group1** members.

Personal Information

Altogether 47 learning recipients (*Group 1*) from 10 countries completed our questionnaire and returned usable data. Of these **Group1** participants, 43.2% were female, 51.3% were male, and 5.5% did not provide any information on their gender. Of the participants, 12.8% had completed High-school or lower, 14.9% had a Bachelor, 36.2% had a Master, 34% had a Ph.D., and 2.1% had completed other educational degrees or qualifications.

Current Skill Progression Towards Learning Goals

Survey participants reported informing themselves about skill demands in the following ways: 86.5% during doing their everyday tasks, 62.2% through reading related papers or news, 54.1% by inquiring

[2]The questionnaire is available on:
`https://tib.eu/umfragen/index.php/survey/index/sid/977178/newtest/Y/lang/en`

[3]It should be mentioned that we allowed participants to answer our questionnaire from the perspective of multiple personae. This was important, as a single person can fulfill different roles in a learning process (e.g. a person can be a lecturer and manager at the same time)

with their supervisors, and 40.5% through job vacancy announcements of positions they apply to. Moreover, they mentioned courses (83.8%), educational videos (78.4%), books (72.9%), and Web pages/documents (64.9%) as dominant resources they used to develop themselves towards skills required by employers. Finally, with respect to open learning content for their self-development, participants bemoaned 1) the lack of personalization, 2) the identification/localization of high-quality learning content, and 3) the time-consuming search process, as the most pressing problems.

Importance and Frequency of Use of the Potential Requirements

Participants rated the importance (1: Not at all important - 5: Very important), and frequency (1: Never - 5: Daily) of usage for each potential user requirement. Once data collection was complete, we calculated the average of their ratings for each of the requirements and normalized the average rates using *Min-Max Normalisation* as 9.1 in which we replaced the *Value* with the *average rates*. Table 9.1 shows the potential user requirements, normalized average importance ratings, normalized average frequency ratings, and the composite rate (multiplication of the normalized importance and frequency rates) which have been sorted based on the composite rates.

$$Normalized\_value = \frac{Value - Minimum_{values}}{Maximum_{values} - Minimum_{values}} \tag{9.1}$$

Table 9.1: Average Importance and Frequency Ratings for Potential User Requirements

Requirement	Importance rate [0-1]	Frequency rate [0-1]	Composite rate [0-1]
*Req*1. Finding learning content about a problem I am working on at the moment	1.00	1.00	1.00
*Req*2. Identifying high-quality content which fulfills my learning needs	0.81	0.68	0.55
*Req*3. Knowing where to start learning when I need a new skill for my studies/job	0.75	0.38	0.29
*Req*4. Identifying which skills are required for my current/future job	0.70	0.36	0.25
*Req*5. Defining my own goals towards jobs I find attractive	0.53	0.29	0.15
*Req*6. Identifying which skills are required for my degree	0.40	0.27	0.11
*Req*7. Finding out how I can improve my skillset in order to qualify for my desired job	0.58	0.18	0.10
*Req*8. Monitoring my learning progress towards desired skills	0.23	0.24	0.06
*Req*9. Making sure that my learning objectives meet job requirements	0.40	0.11	0.04
*Req*10. Identifying which skills are the most important ones in terms of contributing to expected salary	0.05	0.07	0.004
*Req*11. Visualizing potential skill targets	0.05	0.05	0.003
*Req*12. Identifying which jobs I can fulfill with my skillset	0.15	0.00	0.00
*Req*13. Visualizing the structure of the content that I need to master to achieve my skill targets	0.00	0.04	0.00

Findings of our Requirement Analysis

By analyzing participants' ratings regarding these potential user requirements, we prioritized and constructed the following services for learners:

- ***Service_1*: Personalized Search.** *Req*1 and *Req*2 (Table 9.1), clearly received the highest ratings among all requirements. Therefore implementing an educational resources search service, which provides accurate and high-quality search results to address individual learning needs, became one of our top priorities. Clearly, the personalization and the content quality of the results of such a service are critical as demonstrated above, where learners pointed to the *lack of personalization* and *problems in identifying high-quality learning content* as two of the most important barriers to using open/free educational resources. Hence we focused on the context of the learners (e.g. job, skill-set, expertise level, language), and their learning preferences (e.g. their preferred format (e.g video or web pages)).

- ***Service_2*: Goal-driven Learning Content Recommendations.** According to *Req*2, *Req*3, and *Req*5, learners desire a service that helps them 1) explicate their learning objectives, 2) find suitable learning pathways that fit to their context (preferences), and 3) receive the most relevant and highest-quality learning resources needed to meet their learning objectives.

- ***Service_3*: Elucidating Job Skill Requirements.** Based on *Req*4 and *Req*6, the need can be observed to match jobs and the skills that are required to be effective in those jobs. This should be accompanied by visualization, which helps inform users about those skills they need to acquire. Based on this information one can set learning targets and obtain (and ultimately learn) relevant learning content.

- ***Service_4*: Learning Progress Monitoring.** Learners also expressed a strong interest in monitoring their progress toward their learning goals (*Req*8). Accordingly, we found it essential to provide an assessment service, which would help users to test the knowledge they set out to acquire. Additionally, we decided to provide further insights (through numbers, charts, etc.) about users' progress toward each of their learning goals.

9.1.2 Labor Market Intelligence

In order to match jobs to their skill requirements (*Req*4 and *Req*6), we deployed a Labor Market Intelligence (LMI) component to capture up-to-date skill requirements for jobs relevant to this study.

In an initial effort to demonstrate the applicability of our system, we decided to focus on *Data Science* related jobs. We did so because these jobs are both in high demand and particularly prone to change. We selected three associated jobs: *Data Scientist, Data Analyst,* and *Business Analyst*. Subsequently, we used a sample dataset of English job vacancies from *Monster.com*[4], which included 21,937 vacancies and their related skills.

Subsequently, we calculated the rate of occurrence for each of the skills in the target jobs and set the importance of the skills in each job based on this occurrence rate. We used this importance rate to sort the skills that learners need to learn. Based on this process, the following six skills were selected to

[4]The dataset is available on: https://www.kaggle.com/PromptCloudHQ/us-jobs-on-monstercom/version/1

represent our target jobs as they achieved the highest importance rates across our target jobs: *Python programming, R programming, Statistics, Machine learning, Data Visualisation, and Text mining*.

9.1.3 Educational Topic Detection for Selected Skills

In order to recommend open learning content for the selected skills (*Req*2 and *Req*3), we needed to decompose each skill into meaningful learning *Topic*s. Therefore, we extracted learning topics for these six skills by applying *Latent Dirichlet Allocation* (*LDA* [68]) to the transcripts of existing educational materials. Specifically, we used the method proposed by [17] to extract learning topics and determine the degree to which those topics were reflected in each educational resource. Finally, we asked three experts to prioritize each of the extracted topics with an eye on skill development. Table 9.2 shows the number of collected playlists (each of which comprises the educational resources per skill), the number of covered educational videos, and the final number of extracted topics for each skill. It should be mentioned that some of the topics were part of more than one skill (e.g. *Linear Regression* was a topic of both *Machine Learning* and *Statistics* skills)

Table 9.2: Collected Resources for each Skill

Skills	Numbr of Collected Playlists	Number of Covered Educational Videos	Number of Topics
Python programming	8	502	26
R programming	4	185	12
Statistics	9	621	27
Machine learning	9	472	35
Data visualizing	8	257	14
Text mining	6	194	18

9.1.4 Incorporation of Educational Content

In this section, we describe how relevant high-quality open educational resources were collected, filtered, and labeled (*Req*1 and *Req*2). We also depict how assessments were connected to the final set of educational resources included in our recommender (*Req*8).

Collection of Online Educational Resources

To collect open educational content for the six skills and their topics, we performed a search on *Google* and *Youtube*[5] using the concatenation of the skill and the topic (e.g. "Python programming

[5]Using Pafy python-youtube library: https://pypi.org/project/pafy/

Conditions") as the search keywords. We collected 3,228 educational resources[6] which includes 2,514 educational videos and 724 text-based resources (e.g. web pages, lecture notes, and book chapters). For each resource, we collected the following fields based on the available fields for online and open educational resources and the fields we needed to apply our automatic models[7]:

- *Source*. Records the original location of the content.
- *Format*. The format (e.g. Video, Web page, or Book chapter) of the content. This was set based on the source and file extension of the resources. For example, this field was set to *Video* for the resources from *Youtube*.
- *Title*. Records the title of the content.
- *Description*. Records the description of the content.
- *Transcript*. Records the transcription of the content. This field was set based on the transcript of the videos, and the content of the web pages, and book chapters.
- *Rating*. User ratings of the content. This field was calculated differently (e.g. based on 5 point scale rating or likes and dislikes) in the different sources. Therefore, we normalized the ratings for each of the resources.
- *Length*. This field shows the content length (in seconds only for videos).
- *View Count*. Total number of times that the educational content had been viewed by users.

Filtering based on Quality and Relevance

To provide high-quality educational content, which was one of the key outputs of our requirement analysis step (*Req2*), we applied the following filtering procedure on the collected OERs and other available educational resources:

- **Topic-based filtering.** In order to remove educational content that did not fit the search keywords detailed in the previous section, we used the output of our topic models that was described in section 9.1.3. Specifically, we extracted the target topic of each educational resource using our topic models, and removed those resources for which the extracted target topic did not match its search keywords. For instance, if a video was the result of the search keywords "Machine Learning Linear Regression", but our model detected its focus as "Support Vector Machine", we removed it from our resource list. This step resulted in the removal of a total of 1,116 resources (906 of which were video and 210 of which were textual resources)

[6] This is a new dataset and is different from the one we used for the topic detection step.
[7] It should be mentioned that some resources in our dataset did not include all the mentioned fields.

- **Metadata-based filtering.** Previously, [26, 90] showed that the metadata quality of OERs is indicative of their content quality. Based on this finding, we created a binary classifier to sort educational resources into two groups of *high-quality* and *low-quality*. By applying their machine learning model, educational resources with predicted low-quality content (a total of 727 resources of which 621 were video and 106 were textual resources) were removed from our educational content collection.

- **Quality-based filtering.** In our last filtering step, we checked whether OERs and other available educational resources fit the description of the target learning goal of the content (based on the *Wikipedia* page of the search keyword we used to collect the content), and the level of prior learners' satisfaction in terms of content ratings and view counts. This was accomplished through the quality prediction model proposed by [82]. This model leverages the similarity between the transcription of educational resources and the description of their target topics (from *Wikipedia*) in addition to their popularity features (e.g. rating and view count) to determine quality. To apply the model on our dataset, we rebuilt their proposed prediction model based on the features that existed in all of our collected resources (i.e. Transcript, Rating, and View Count) which led to 79.2% of the F1-score on their published dataset. As a result of this step, a total of 631 (547 video and 84 textual) resources were removed from our collection.

Through the application of the aforementioned filters, we distilled 764 high-quality (440 video and 324 textual) OERs and other available educational resources, covering all topics (9.1.3) in our six target skills (see Table 9.3). It should be noted that the number of educational resources for each topic ranged between 3 and 10, and that we had at least one video and one textual resource for each topic. Moreover, in our dataset, there were resources that addressed more than one topic (e.g. an educational video could cover both *Linear Regression* and *Gradient Descent*).

Table 9.3: Number of Resources which Passed Through our Filtering Steps

Skills	Number of Educational Resources	Avg Number of Resources per Topic
Python programming	124	4.77
R programming	49	4.08
Statistics	209	7.74
Machine learning	263	7.51
Data visualizing	100	7.14
Text mining	120	6.67

Educational Resource Labelling

To generate personalized recommendations for the learners, we analyzed and labeled all of the educational resources that were retained. Some features such as *Source, Format, Transcript, Rating, View Count* had already been extracted automatically (see 9.1.4). Additionally, for each skill, we asked two experts to review and label the resources (see below). As a result, the following features were collected for all filtered educational resources:

- *Length.* As we extracted the length of educational videos (in seconds), we asked experts to estimate how long it would take learners (in seconds) to scrutinize the text-based educational resources. Afterward, we grouped educational resources in such a way that we had groups with a similar number of resources, that we could describe to the learners easily. Therefore, we created 3 groups of *Short* $<$ *10 minutes* (included 308 resources), *10 minutes* $<$ *Medium* $<$ *20 minutes* (included 225 resources), and *Long* $>$ *20 minutes* (included 231 resources) resources.

- *Level of Detail.* This feature captures the level of detail in which specific content addresses a target topic[8]. Experts assigned the following labels to the resources: *Low Detail, Medium Detail, High Detail*.

- *Learning Strategy.* We defined three learning strategies of *Theory-based, Example-based,* and *Mixed* (which includes both theory and example) based on [91], and asked experts to label resources accordingly.

- *Is a Classroom-Based Instruction.* This field is a Boolean value that captures whether the resource has been recorded as a university class or not.

Table 9.4: Preference Features

Feature	Possible Values	Notes
Length	Short, Medium, or Long	Learner's preference about the length of educational resources
Detail	Low, Medium, or High	Learner's preference about the level of details in educational resources
Learning Strategy	Theory-only, Example-only, or Both	Learner's theoretical knowledge orientation
Classroom-based	Yes or No	Learner's preference about learning content originated from classrooms
Content Format	Video, Book, Web page, Slide	Learner's preference about learning content formats

Implementation of Learning Progress Monitoring

To produce well-defined and relevant assessments, three experts generated and carefully reviewed multiple-choice questions (test items) for each topic. In this process, a question was selected to be

[8]The topic can be a concept, formula, or an API

added to our test items, when all reviewers found it appropriate to assess the knowledge of learners in the topic(s) that the question targeted. This resulted in a repository of topic-based and skill-based test items. In our prototype, we implemented two different types of assessment, each of them are generated dynamically, according to the individual progress of each learner:

- A **progress assessment** is a test that only contains test items related to topics. This test validates the progress of a user, when they transit between consequent topics within a skill. Learners can only start a new topic if they pass the assigned progress test of the prerequisite topic(s). In case a learner passes a topic associated with a target skill, the topic is marked as completed in all corresponding (related) skills listed in our recommender. For instance, if a learner passes the topic "Linear Regression Concept" when studying for the "Machine Learning" skill, this topic will also be completed for the skill "Statistics", even though this skill is not among the skill targets of the learner. This method helps track individual development, by monitoring knowledge and skill proficiency levels across topics and skills, within and beyond individual learning objectives.

- A **skill assessment** can be interpreted as an assessment of skills (i.e. a topic aggregate), and can be used to provide feedback to learners about applying and combining acquired knowledge (topics) areas in relation to a specific skill. Therefore, these assessments include questions that cover all topics associated with a specific skill. Learners complete these assessments as soon as they have mastered the different components (topics) of a target skill.

Using progress and skill assessments, a learner can continuously evaluate their level of knowledge in a fine-, and coarse-grain manner.

9.1.5 Personalized Open Learning Content Recommendation

In this section, we demonstrate our proposed personalized recommendation system for learners to address *Req*1 and *Req*2.

Learner Profile

Based on the features we collected for educational resources (9.1.4), we also defined features for the preferences of each and every learner. These features are described in Table 9.4. Based on possible feature values, we created a long-, and a short-term 15-dimensional preference vectors for each learner which included the following features: *Length-Short, Length-Medium, Length-Long, Detail-Low, Detail-Medium, Detail-High, Strategy-Theory, Strategy-Example, Strategy-Both, Class-based, Non-class-based, Content-Video, Content-Book Chapter, Content-Web Page, Content-Slide*. Each feature value in a vector shows how much (a float value from 0 - the lowest, to 1 - the highest) a learner prefers receiving learning resources with that feature. The long-term vector is used as the basis for our learning content recommendation. Therefore, the complete history of each learner's feedback (5-scale ratings for the recommended educational contents) until the recent updating period is taken into account. The short-term vector shows learners' feedback in the recent updating period (last one month) and it affects the long-term vector at the end of each updating period; therefore, the short-term vector is emptied at the starting point of each updating period and updated after each feedback from the learner. The long-term vector helped us to capture the learners' preferences while using the recommended resources (it should be noted that the long-term vector is configured to place more weight on the recent ratings). We defined the *updating period* as a configurable period value (which could be set in our system), and set it to one month in this version of our system.

When a learner registers in our system for the first time, we ask questions regarding all preference features in order to populate the long-term preference vector. This is done by transforming the selected values into the corresponding values (float number between 0 and 1) in our preference vectors. For instance, when a learner prefers *Long* content, the *Length-Long* feature is set to 1, while the Length-Short and Length-Medium features are set to 0. As another example, if a learner, selects 3 on a 5-point rating scale rating regarding the video contents, the Content-Video feature is set to 0.5.

When the learners complete a learning content, we consider their feedback, which is a 5-scale rating, to update their short-term profile. For instance, assume that after recommending two pieces of learning content with a *High* level of detail to a learner, we receive the following feedback ratings: 1) 3 in a 5-scale rating (which means 0.5 out of 1 in our system), and 2) 5 out of 5 (which means 1 out of 1). As a consequence, the *Detail-High* feature of the short-term vector is set to 0.75 (which means 4 in a 5-scale rating) for the learner.

At the end of each updating period (which was set to one month), we update the long-term vector by calculating the average of the current long-term vector and the short-term vector. This updating procedure detects changes in long-term individual learning preferences and results in more relevant content suggestions. It should be mentioned that the values of the long-term vector can be also viewed and directly edited by learners through their dashboard, in their profile settings.

Recommendation Engine

To recommend learning content on a specific topic to a particular learner, first, we retrieved all the resources (the ones that passed our filtering process) which focused on the topic. Afterward, we created the same 15-dimensional vector (with the same features as the preference vector) for each retrieved learning resource, as we did for the learners (see section 9.1.5)[9]. Finally, we calculated the *Dot Product* [92] of the learner's long-term preference vector together with the created vectors of

[9] As an example, for a *Short* content, we set the Length-Short feature to 1, and the Length-Medium and Length-Long features to 0

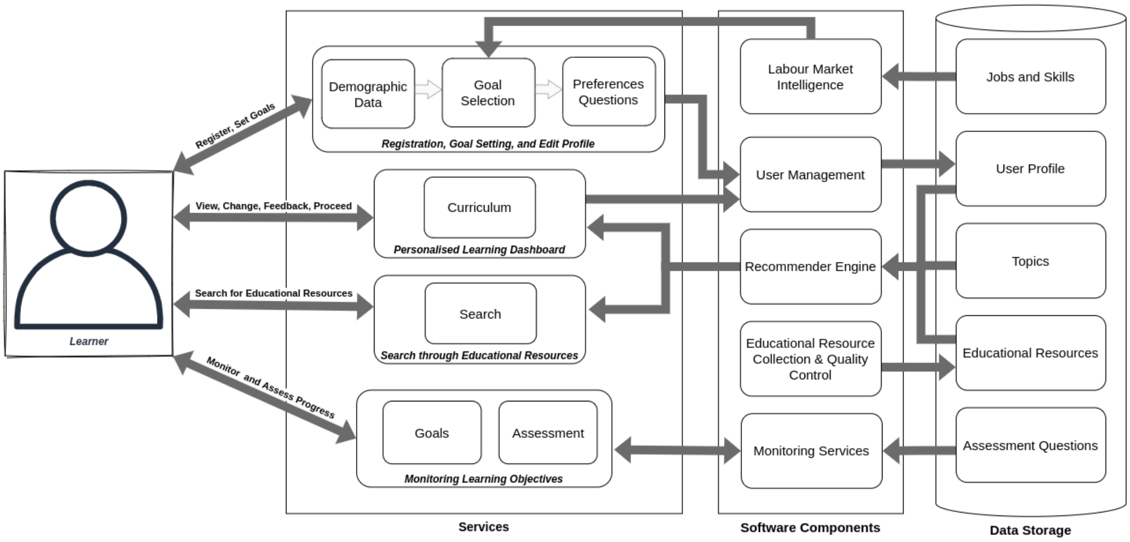


Figure 9.1: Interaction between different Parts of our Prototype Dashboard to Provide the Required Services

each retrieved learning content. As a result, our system recommends the content with the highest *Dot Product* result.

9.1.6 Learning Dashboard

In this section, we showcase our learning dashboard, called *eDoer*, that we implemented to provide our individualized learning services (9.1.1) to learners[10]. Figure 9.1 illustrates how the different technical components of our recommender prototype interact with one another (and with the learner) to create the learner's personal learning experience (9.1.1). For the User Interface (UI) we incorporated responsive web design and design guidelines [93, 94]. We provided learners with an interactive tutorial [95] at their first login, in order to familiarize them with the different functionalities of our learning dashboard.

Registration and Goal Setting. The registration path consists of three consecutive steps, each serving a different purpose: 1) In the first step we collect the necessary demographic information from new learners, including their name, email address, gender, and geographical location (country and city)[11]. 2) In the second step, learners search for and select a target job. Subsequently (as an implementation of *Service_3* depicted in 9.1.1), we show the required skills for the selected job by using our *labor Market Intelligence* (9.1.2), and ask learners to select those skills they want to master. In addition, users can search and select complementary skills (not connected to their target job) and

[10] https://github.com/ali-faraji90/edoer/blob/main/Files/Demo.mp4?raw=true
[11] https://github.com/ali-faraji90/edoer/blob/main/Files/RegistrationForm.png

add them to their target skills manually [12]. 3) The third (and last) step consists of setting learning preferences by answering a number of questions (see section 9.1.5), to further calibrate the learning content recommender algorithm for each particular learner[13].

Personalized Learning. To provide *Service_2* (see 9.1.1), a curriculum page was designed to structure and monitor the advancement of learners with respect to their target skills and related topics. Learners can visualize their personalized curriculum by selecting a skill. Once the skill is selected, the related list of topics is displayed, sorted by their priority (see 9.1.3)[14]. Each topic has a status, which shows whether the topic has been *passed*, is *in-progress*, or *forthcoming*. For each *in-progress* topic, one educational resource is recommended (displayed). Besides accessing (and learning) the content, the learner has the following options with respect to the recommended learning content:

- **Change**: If the learners are not satisfied with the content for some reason (e.g. it is not relevant, the instruction doesn't fit the preference, the format of the content is not preferred, low technical quality of the video/audio/text), they can replace the presented learning content, with another one addressing the same topic, at the same level. Thus, the recommendation engine records this *Change command* as an instance of feedback with a minimal value. At the same time, it updates the learners' short-term preference vector as described in section 9.1.5, and provides an alternative educational resource, on the basis of the updated vector.

- **Done**: When a learner completes a specific learning resource, they can indicate that with the *Done* button, and optionally rate the learning content on a 5-point rating scale. The learner's profile is automatically updated based on this rating, as described in section 9.1.5. Learners can also indicate whether they would like another learning content on the same topic, or whether they would like to try to progress to the next learning topic(s) related to a particular skill target by (successfully) taking a progress assessment (see 9.1.4).

Search Page. In order to address *Service_1* in section 9.1.1, we provided a straightforward and simple way for learners to search through all of the open and available learning resources that are accessible on our platform. Search results (a particular learning resource) can be added to the curriculum page, and they are displayed as extracurricular learning resources.

Monitoring Learning Objectives. To address *Service_4* (see 9.1.1), we implemented a Goal Page for learners to gauge their learning progress towards their skills targets. The page, therefore, provides detailed information on the number of completed learning topics for each skill[15]. It should be mentioned that changing the target job, or removing a skill from the skill targets may remove incomplete skill training curricula from the curriculum page. However, learners can view both the new

[12] https://github.com/ali-faraji90/edoer/blob/main/Files/GoalSelection.png
[13] https://github.com/ali-faraji90/edoer/blob/main/Files/Preferences.png
[14] https://github.com/ali-faraji90/edoer/blob/main/Files/Curriculum.png
[15] https://github.com/ali-faraji90/edoer/blob/main/Files/Goals.png

or updated target skills, and also those skills that have been removed or are incomplete. Learners can also reactivate incomplete skills by means of a simple click. Moreover, monitoring learning processes require attention to *Assessment*, as we discussed under *Service_4* in section 9.1.1. For this reason, we deployed an Assessment Page to structure and keep track of all skill assessments explained in section 9.1.4. On this page, learners can see and engage with comprehensive skill assessments for each of their target skills. Furthermore, to plot individual performance on skill assessments over time, a skill assessment history for each target skill is provided as a graph.

History Page. This page contains all learning resources that have previously been recommended to the learner. This gives learners the opportunity to review any of these learning resources at will. Content on the history page is also categorized based on learners' target skills and topics. Learners can also find information about their feedback regarding learning content, including a timestamp of completion.

Profile Page. This page provides access (read and edit) to all the data we collected during the registration process and beyond. This includes all demographic data, the target job, target skills, and learning preferences.

9.2 Validation

In an effort to demonstrate the effectiveness of our proposed online open educational resource recommender (*eDoer*), below we report on a randomized experiment carried out with the explicit aim of having real users interact with our prototype. The experiment was conducted to support the internal validity of our system, by answering the question of whether engagement with our system results in improved knowledge acquisition. In this part, we showcase the methods and results of this validation step.

9.2.1 Objective

As mentioned earlier, to more formally evaluate *eDoer* with a particular focus on evidencing the internal validity of our inferences pertaining to the effectiveness of *eDoer* in imparting knowledge to a sample of students, we set out to conduct an experiment in the context of learning about statistics. Specifically, we formulated and tested the following hypotheses, which were, by and large, premised on the fact that we specifically developed eDoer to address the most important requirements signaled by key stakeholders (see 9.1.1). Relying on an experimental design, enhanced our ability to rule out alternative explanations for any observed effects.

1. ***Hypothesis 1***. Using *eDoer*, as opposed to self-directed online search for open educational resources to learn about basic statistics, has a positive effect on knowledge of basic statistics.

2. ***Hypothesis 2***. Having *eDoer* provide personalized recommendations in terms of educational format (webpage, video, book, slide), length (short, medium, long), level of detail (low detail, medium detail, high detail), and content type (including example/theory or not), as opposed to having *eDoer* provide random content (from the quality controlled materials), has a positive effect on knowledge of basic statistics.

Although these hypotheses are limited in their breadth and coverage of the eDoer system, we feel they address the core functionalities/requirements that we wanted to evidence at this stage.

9.2.2 Procedure

For this experiment, we used the *Prolific* platform[16] which is a commercial service provider for connecting researchers with participants. In light of financial constraints associated with compensating respondents for their time (we paid each respondent 15.76 British pounds - approximately 21.74 US dollars - for their time and effort), we set out to collect high-quality learning data from a total of 150 participants. For this purpose, we decided to recruit a total number of 175 users as we predicted that we might need to remove some of the participants' data for different reasons (such as technical problems and/or missing data).

We selected "*Basic statistics for engineers*" as the target skill for this study and ran our topic extraction method on it which resulted in the following seven topics: *1. central tendency measures* (i.e. *mean, median, mode*), *2. variance and standard deviation, 3. covariance and correlation, 4. conditional probability and independent variables, 5. normal distribution, 6. linear regression*, and 7. *hypothesis testing, p-value, and confidence interval*. The reason that we selected this particular skill was to target a fundamental (engineering-related) skill while at the same time ensuring the availability of open educational resources for those people assigned to the control group (who would not be engaging with eDoer).

In order to take part in this study, the potential participants needed to complete the following steps:

1. **Step 1: pretest**. In the first step, all users participated in a pretest[17] on "*Basic statistics for engineers*" that assessed prior knowledge of the aforementioned seven topics. The test included seven questions (one question per topic) which were selected through a discussion between three experts. The experts were also asked to define the required time for each question in a way that if a participant knew a topic, he or she would have enough time to answer the question in the allotted time period.

   After completing the pretest, participants were randomly assigned to one of the following groups (to which they remained blind):

[16] https://www.prolific.co
[17] https://uvafeb.eu.qualtrics.com/jfe/form/SV_5AudD6pyhqWb5vU

- **Group 1**: Self-directed learning using online searches, but without any support from *eDoer*
- **Group 2**: Learning through *eDoer* without personalized recommendations
- **Group 3**: Learning through *eDoer* with personalized recommendations

2. **Step 2: Learning process**. In this step, the participants were granted 105 minutes (15 minutes per topic) and instructions (according to their assigned group) to study the aforementioned topics in order to be able to answer a new set of questions. The questions were on the same topics as the pretest and within the same level of difficulty. The instructions were as follows:
   - **Group 1**: In the learning process, the participants were presented with the 7 extracted topics for a finer grain searchability. They were free to engage with any type of educational content they could find (e.g. through online searches, reading books, and watching educational videos).
   - **Group 2 & 3**: These groups received simple instructions on 1. how to log in to *eDoer* using information from pre-registered new test-users, 2. fill the preference form on *eDoer*[18], and 3. adding the skill "*Basic statistics for engineers*" to their learning profile. Subsequently, they were directed to the curriculum page to start studying each of the topics for the target skill within the defined time period.
3. **Step 3: posttest**. After the learning process, all groups were directed to the posttest which included the same number of questions, on the same topics, and with the same level of difficulty level[19]. This set of questions was also differently timed in the same manner that we did for the pretest.
4. **Step 4: Feedback survey**. Finally, all participants filled out a short survey to provide us with feedback. *Group 1* received a survey[20] about the steps they took to learn the topics on their own. *Groups 2 & 3* received a survey[21] about their experiences using *eDoer*. Also, all groups were asked a question about their impression of the study in general.

Upon examining the data, we decided to remove 14 participants from our study as they had 1 (or less than 1) correct answer from all 14 questions. We did this to prevent the inclusion of respondents who were not seriously participating in our experiment. Also, we removed 5 participants' data because of the technical issues they faced during the study. In the end, *Group 1* consisted of *53* participants, *Group 2* of *50* participants, and *Group 3* of *53* participants.

[18] Although Group 2 were not receiving personalized material, they also filled out the preference form as they had not any information about which group they were assigned to

[19] https://uvafeb.eu.qualtrics.com/jfe/form/SV_4Sl8QGDg5AtECSq

[20] https://docs.google.com/forms/d/e/1FAIpQLSeEV5ekM6rAn_s0AscxTawgbVPm3eXjhwfF3Vjrqos_2HmnUg/viewform?usp=sf_link

[21] https://tib.eu/umfragen/index.php/887411?lang=en

9.2.3 Measures

We calculated both scores (i.e. pretest and posttest scores) for each individual participant as the number of correct answers divided by the total number of questions per test. Subsequently, we computed our first measure ***progress score*** by subtracting, for each participant, the pretest result from the posttest result.

Additionally, through *Step 4* (*Feedback survey*), we collected the participants' opinions on a 5-point scale (1: lowest to 5: highest) on the following items and converted their ratings into a number between *0.0* and *1.0* (i.e. 1 as 0.0, 2 as 0.25, 3 as 0.5, 4 as 0.75, 5 as 1.0):

- *Group 1*:
  - Availability of educational content
  - Quality of educational content
  - Satisfaction from the prolific experiment
- *Groups 2 & 3*:
  - Personalization of content
  - Quality of educational content
  - Satisfaction from the prolific experiment
  - Suggesting *eDoer* to other learners

Finally, to quantify the learner's overall satisfaction with *eDoer*'s recommendations, we decided to collect the ***Recommendations' Ratings*** (on a 5-point scale) for the recommended educational materials.

9.2.4 Analytical Procedures and Results

In the pretest, on average, *Group 1*, *Group 2*, and *Group 3* achieved average scores of *0.22*, *0.24*, and *0.20*, respectively. As expected, the pretest showed that most participants had no previous experience with *Statistics* before the experiment as their scores appear to reflect random responding. Also in the posttest, *Group 1*, *Group 2*, and *Group 3* achieved an average score of *0.34*, *0.42*, and *0.42*, respectively. Based on the pretest and posttest scores, we calculated our first measure as *progress score* which showed how each group improved their knowledge in the target skill. This measure was *0.12* for *Group 1*, *0.18* for *Group 2*, and *0.22* for *Group 3*. As you can see *Group 3*, which benefited from both *eDoer* and personalized recommendations, showed the most improvement. *Group 2* which benefited from *eDoer* but received random (non-personalized) recommendations also showed some degree of improvement. Finally, and as expected, *Group 1* which did not engage with *eDoer* had the lowest *progress score*.

To formally test our hypotheses, a one-way ANCOVA[22] was conducted. After adjustment for pretest scores, there was a statistically significant difference in posttest scores between the groups of learners, $F1,152 = 11.202, p < 0.001$. Further investigation through pairwise comparison of estimated means showed that there was a statistically significant difference $t152 = 2.31, p < .05$ between the posttest scores of the group receiving eDoers' non-personalized recommendation $M = 2.91, SD = 1.18$ and the group of self-directed learners $M = 2.38, SD = 1.16$. Furthermore, our findings also show a significant difference $t152 = 2.49, p < .05$ in test scores between self-directed learners and learners receiving eDoers' personalized recommendations $M = 2.98, SD = 1.27$. However, there was no significant difference between the posttest scores of the groups receiving non-personalized or personalized recommendations from eDoer $t152 = 0.137, p = .892$.

In support of hypothesis 1, our findings show that participants who used eDoer without personalization attained significantly higher scores on the posttest than participants who engaged in self-directed learning (i.e. those who did not use eDoer). Unsurprisingly, and further supporting hypothesis 1, participants who used eDoer with personalization also attained higher scores on the posttest than participants who engaged in self-directed learning. In contrast, no support was found for Hypothesis 2, in that there appeared to be no significant difference in the posttest scores between those receiving personalized recommendations and those receiving non-personalized recommendations, again after controlling for scores on the pretest. To account for capitalization on chance, we reran the pairwise comparisons of estimated means applying a more conservative Bonferroni correction. The results of these analyses indicated a significant difference in the expected direction between self-directed learners and learners receiving personalized recommendations $t152 = 2.49, p =< .05$ but the difference between self-directed learners and learners receiving non-personalized recommendations failed to reach statistical significance $t152 = 2.31, p.066$, even though it was in the expected direction. It should be noted, however, that the Bonferroni correction has been criticized for being overly strict.

Table 9.5 shows the results of the other measures incorporated in our study for each group. In eyeballing these data, it is noteworthy that ratings provided are most favorable for the personalized version of eDoer, followed by the non-personalized version, and finally the self-directed learning group. Moreover, the fact that 75% of the participants are willing to recommend *eDoer* to other learners, reflects their positive attitudes toward the *eDoer* platform.

[22] We also used the *Bayesian* analysis[96] to test both of our hypotheses. The reason that we also ran *Bayesian* hypothesis testing was to serve the interests of those who purport that Bayesian methods are superior[96]. However, the results did not change the conclusions we derived based on the traditional t-test.

Table 9.5: Results of the eDoer Evaluation Experiment

Measures	**Mean (out of 1)**			**Standard Deviation**		
	Group 1	*Group 2*	*Group 3*	*Group 1*	*Group 2*	*Group 3*
Progress-score	0.12	0.18	**0.22**	0.20	0.19	0.17
Availability of educational content	0.56	-	-	0.16	-	-
Quality of educational content	0.64	0.75	**0.82**	0.27	0.22	0.20
eDoer recommendations' rating	-	0.79	**0.87**	-	0.17	0.16
Satisfaction from the experiment	0.74	**0.90**	**0.90**	0.23	0.14	0.14
Suggesting *eDoer* to other learners	-	0.74	**0.76**	-	0.23	0.22

9.3 Outcome of Building AI-based Open Recommender System for Goal-driven Personalized Education

To remain employable and achieve required knowledge, learners continuously need to master skills and topics that are relevant for their desired goals (e.g., jobs in a dynamically changing labor market). We initiated the work reported in this chapter by conducting a requirement analysis to extract learners' need for such a learning environment. Based on the results of our analysis, we designed and implemented a system, called *eDoer*, that helps learners to set their learning goals and to receive a personalized learning path towards their goals. These learning paths contain high-quality educational materials, which have passed through our automatic quality control models (i.e. topic based, metadata based, and quality based prediction models). We evaluated our prototype system through an experiment in the context of a fundamental engineering skill (i.e. *Basic Statistics*). This validation showed tentative support for our first hypothesis, indicating that learners, who used our system, performed better on a post-test than learners engaging in self-directed learning (outside of eDoer). The findings for the learners who received non-personalized (i.e. randomly selected content) were less convincing. In our most conservative test, which was about the difference between the personalized and non-personalized groups in eDoer, our hypothesis failed to reach statistical significance.

CHAPTER 10

Horizontal Aspects of this Thesis

In this chapter, we cover the related research that we have done beside our research main story. In section 10.1, we explain the developed quality standards for OER. After that we illustrate our research on a personalized educational system which supports accessibility requirements in section 10.2. Finally, in section 10.3, we cover the proposed ontology on the area of labor market driven education.

10.1 Quality evaluation of open educational resources[1]

OpenCourseWare (OCW) is defined as free and open digital publication of educational and learning content [97]. OCW platforms organize education materials, known as Open Educational Resources (OER), in the form of online courses. These courses generally provide a learning plan and evaluation tools. Many OCW platforms exist (e.g., MIT[2]) with various OER representations, such as videos, audio and slides.

Finding high-quality OERs becomes increasingly cumbersome due to the growing amount of published resources [98]. However, selecting high-quality resources is crucial to ensure the quality of an online course. In this work, we propose evaluation metrics to assess the quality of OERs. The metrics are implemented within SlideWiki[3], a collaborative OCW platform focused on presentation slides. The implementation demonstrates how the metrics can be integrated within OER authoring tools. Although the presented metrics are evaluated on presentation slides, they can be applied to other OER representations as well.

[1] This chapter has been published as follows: Elias, M., Oelen, A., Tavakoli, M., Kismihok, G. and Auer, S., 2020, September. Quality evaluation of open educational resources. In European Conference on Technology Enhanced Learning (pp. 410-415). Springer, Cham.
[2] https://ocw.mit.edu
[3] https://slidewiki.org

This article addresses two research question: 1) *how to evaluate the quality of OER material?* and 2) *how to use this evaluation to guide OER authors and learners?* In order to define the quality metrics and to develop the implementation, we investigate related work to OER quality assessment. Accordingly, we select and extend the dimensions that are related to content representation, and we define a set of metrics for each dimension. Finally, we evaluate our work by conducting a questionnaire with OER expert users (i.e., instructors and PhD students) and by implementing a set of the metrics in an authoring tool.

The remainder of this paper is structured as follows: Section 10.1.1 discusses and analyzes the state-of-the-art of evaluation approaches used for assessing the quality of online educational systems and contents. Section 10.1.2 defines a list of quality metrics and describes each of them. Section 10.1.3 explains the implementation of the metrics and evaluates the results. Finally, Section 10.1.4 concludes this work.

10.1.1 Related Work

This section reviews the state-of-the-art of the quality evaluation approaches for OER repositories. We focused on approaches that address quality aspects related to the OER content and representation. We analysed dimensions found in the literature and categorized them based on the quality aspects: 1) feature quality (i.e., quality related to functionalities provided by the OER repository), 2) technological quality (i.e., quality related to the technology and implementation of the OER repository), and 3) content quality (i.e., quality related to the OER material and content representation). Table 10.1 shows the dimensions that were extracted and categorised as per our analysis.

Table 10.1: Summary of quality evaluation dimensions

References / Dimensions		[99]	[100]	[101]	[102]	[103]	[104]	[105]	[106]	[107]	[108]
Features quality	Availability	✓	✓	✓							✓
	Multiliguality	✓	✓	✓		✓					
	Reusability	✓		✓		✓		✓	✓		
	Provenance	✓						✓			✓
	Recency	✓						✓			✓
	Openness	✓	✓	✓	✓		✓	✓		✓	
Technological quality	Accessibility		✓		✓	✓	✓	✓	✓	✓	
	Alignment to standards		✓	✓	✓		✓		✓		✓
	Usability		✓		✓	✓		✓	✓	✓	
	Compatability		✓			✓		✓	✓		
Content quality	Structure		✓					✓	✓	✓	
	Accuracy		✓		✓			✓	✓	✓	
	Comprehensiveness		✓		✓		✓	✓			✓
	Discoverability	✓	✓	✓							
	Multimodality				✓	✓	✓	✓	✓	✓	✓
	Self-assessment	✓	✓		✓		✓				

From the analysis, we found that most of the evaluation approaches that were studied in Table 10.1, evaluate the dimensions and metrics either conceptually or by providing a checklist to experts or users. These checklists are either filled out manually or in the form of online surveys [109]. Automatic OER quality assessment and author quality guidance were not addressed. Since this study is focusing on evaluating the quality of *OER materials*, we focus on the dimensions defined in the content quality part from Table 10.1 and extended them in Table 10.2. We also use accessibility and compatibility from the technological aspect because they address OER content as well.

10.1.2 Proposed OER Quality Evaluation Metrics

The Open Education Consortium (OEC) defines OERs as materials that are composed of course planning, thematic content, and assessment tools [97]. Accordingly, we divided our evaluation approach of OERs into three components: content structure, learning content and self-assessment. Content structure defines the organization and navigation of the educational resource. Learning content refers to representation of the learning material. Self-assessment is related to the availability of questions to evaluate the learning process. Table 10.2 lists the dimensions and metrics to assess the quality of OER materials.

10.1.3 Implementation and Evaluation

For the implementation, a set of eight metrics (i.e., CS1.1, CS2.1, CS4.1, CS4.2, LC4.1, LC5.1, SA1.1, SA1.2) has been selected from Table 10.2. The set of quality metrics was selected based on relevance, appropriateness and technical viability within the SlideWiki platform.

An example of a quality report is available via SlideWiki[4]. The quality report is displayed on deck-level, and is visible for all users. There are several reasons for making the quality report public. Firstly, there is an extra incentive for OER creators to ensure that their presentation meets a certain quality standard. Secondly, learners can decide to use an OER based on its quality. And finally, due to the collaborative nature of SlideWiki, learners can help improving the slide deck based on the report. For each metric, the amount of detected issues is listed. In case no issues are found, the text “All good” is displayed. For metrics CS4.1 and CS4.2 a quality score is shown. Listed metrics can be expanded to view more details about a particular metric, including why adhering to this metric is important. In case an issue is detected, more information about this issue is displayed.

To evaluate our quality dimensions and metrics, we invited OER expert users (either university instructors or PhD students) and asked them about the importance (1: less important, 5: very important) of our metrics in each dimension with the help of a qualitative questionnaire[5]. Moreover, the participants provided opinions about the overall quality of existing OERs (as *Current Quality* column, 1: lowest quality, 5: highest quality), and overall usefulness (1: not useful, 5: very useful) of

[4] http://slidewiki.org/deck/90789/02-rdf-data-model/deck/90789

[5] https://forms.gle/2Y4bhzbEK3LTY5y78

Table 10.2: OER quality metrics

Dimension	Metrics	Description
Content Structure (CS)	CS1. Clearness of the taxonomies	CS1.1 Short and descriptive name (i.e., characters limit) CS1.2 Coherence with content title (i.e., consistent file name with the content title) CS1.3 Progress inference from title (i.e., consistent coding scheme)
	CS2. Easiness of navigation	CS2.1 Hierarchical design (i.e., well-organized structure) CS2.2 Depth of the taxonomy (i.e., less scrolling)
	CS3. Adaptability of the structure	CS3.1 Availability of adaptability mechanism (e.g. smaller chunks design)
	CS4. Discoverability of the content [26]	CS4.1 Availability of Standardized Metadata (i.e., sum of the normalized importance scores of metadata) CS4.2 Adherence to Standardized Metadata (i.e., including a Rating function)
Learning Content (LC)	LC1. Quality of text	LC1.1 Correctness of text spelling and grammar LC1.2 Comprehensiveness of text (i.e., using readability meters)
	LC2. Adaptability of content	LC2.1 Availability of various content formats (e.g., based content, web media, interactive media, video, audio) LC2.2 Availability of multiple content representation (e.g., multiple themes for learning slides) LC2.3 Consistency between the content types (i.e., synchronized maintenance and versioning management)
	LC3. Compatibility of content on multiple devices	LC3.1 The number of supported devices (e.g., mobile phone, tablet, laptop, assistive technologies) LC3.2 Availability of compatibility checking mechanisms (e.g., validating responsiveness of web pages)
	LC4. Accessibility of content representation	LC4.1 Compliance to guidelines of content representation (e.g., WCAG 2.1 guidelines) LC4.2 Availability of validation approach of content representation (e.g., validating that an image contains alternative description to support accessibility)
	LC5. Multilinguality of content	LC5.1 Availability of resources in more than one language (i.e., other than English) LC5.2 Existence of translation approach (i.e., automatic translation, expert-revised) LC5.3 Availability of synchronization of material translation
Self-assessment (SA)	SA1 Availability of self-assessment	SA1.1 Existence of self-assessment content SA1.2 Availability of answers SA1.3 Average number of question covering the content (i.e., number of questions per each learning object) SA1.4 Existence of question generation approach (e.g., automatic generation or author entry)
	SA2. Variety of self-assessment questions	SA2.1 Available type of questions (e.g., multiple choice, close text, sorting). SA2.2 Average number of question per assessment type

our metrics in each dimension. We collected the feedback of ten participants who had experience with OERs as author (2 participants), learner (5 participants), and teacher (5 participants). The evaluation results of each dimension and metric are: 1) *Content Structure* is considered useful by 100% of the participants, 2) *Learning Content* is considered useful by 60% of the participants, and 3) *Self-assessment* is considered useful by 80% of the participants. Regarding the usefulness and coverage of the proposed dimensions and metrics, 70% of the participants find our dimensions and metrics useful and 50% of the participants agreed that the proposed dimensions and metrics cover the important metrics needed for evaluating the quality of OER materials, while 30% of the participants provided a neutral response.

10.1.4 Conclusion

This paper proposes quality evaluation metrics for OERs to help learners and teachers to find high-quality OERs and guide OER repositories to improve their content. Two research questions were addressed in this article. To answer the first question "how to evaluate the quality of OERs", we established and distributed quality evaluation metrics covering three aspects of OER quality assessment: content structure, learning content and self-assessment. For the second question "how to use the evaluation metrics to guide authors and learners of OERs", we selected seven of these metrics and implemented them in SlideWiki. Quality reports are publicly visible for all users in order to help learners find high quality content, and encourage authors to improve their materials. We evaluated our metrics by collecting feedback from OER users and creators via a questionnaire. As future work, we will continue implementing the metrics in SlideWiki and study the effect of the evaluation reports on the learners' selection of OERs, and authors' ways of creating and editing OERs.

10.2 An OER Recommender System Supporting Accessibility Requirements[6]

Open Educational Resources (OERs) are free and open-licensed educational materials that are ideally composed of course planning, thematic content, and assessment tools [97]. OERs are typically provided by OpenCourseWare systems (e.g., MIT[7]) in various formats (e.g., videos, audio, slides). Since self-learning is oftentimes the most typical way to acquire new skills or update existing skills to match the rapidly changing requirements of the labor market [4], OERs can potentially provide open access materials that can be used by a wide range of learners over the web. Consequently, there is a need to identify high-quality OERs that address learners' needs and preferences in a wide range of contexts. These learners include people with disabilities who have diverse needs, depending on the type and severity of their disabilities. As per the WHO statistics, one billion of the world population has some form of disability, and it is expected to double by 2050 [110]. At the European level, about *60%* of citizens with disabilities are employed (employment rate of persons without disabilities is *82%*), and *22.5%* of the youth with disabilities abandon education systems early (only *11%* of youth without disabilities) [111]. The lack of access to education, vocational rehabilitation, and training is among the most important reasons of low rates of employments [112].

[6] This chapter has been published as follows:: Elias, M., Tavakoli, M., Lohmann, S., Kismihok, G. and Auer, S., 2020, October. An OER Recommender System Supporting Accessibility Requirements. In The 22nd International ACM SIGACCESS Conference on Computers and Accessibility (pp. 1-4).

[7] https://ocw.mit.edu/

Our research seeks to address the following questions: 1) how to represent profiles for learners with accessibility needs, and 2) how to retrieve high quality OERs with respect to these learners' preferences. In order to answer these questions, the following steps were carried out: 1) We formally represented the accessibility requirements of OCW by using the concepts of the AccessibleOCW ontology [113]. Afterwards, 2) we reused the OER recommender system [82] to implement our approach and retrieve high-quality OERs that are relevant to the accessibility needs of learners as defined in the AccessibleOCW ontology and the learner profile. Finally, 3) we evaluated the accessibility of the results by means of manual and automatic testing, and also by getting feedback from experts.

10.2.1 Related work

Accessibility and Design for All refers to the creation of products, environments, programs and services that can be used by all people, to the greatest extent possible, without the need for adaptation or specialized design [114]. In general, accessibility requirements are defined by the web accessibility guidelines, such as W3C Web Content Accessibility Guidelines (WCAG) 2.1 [115], W3C Cognitive and Learning Disabilities Accessibility Task Force (Cognitive A11Y TF) [116], IMS AccessforAll [117], and Easy-to-Read [118], to lead the development of accessible systems. Inclusive OCWs, therefore, should address these accessibility requirements [119].

Although there is a large amount of OCW platforms and OER repositories (e.g., MERLOT collection hosts over 40,000 openly available resources from over 250 providers[8]), accessibility is still not widely addressed by OERs [120]. According to a systematic review focusing on recommender systems in e-learning [121], from 108 papers that were studied, only one has considered accessibility in its approach. Therefore, there is a need to help learners define their preferences, and retrieve OERs matching their needs (e.g., blind users might prefer textual over video resources).

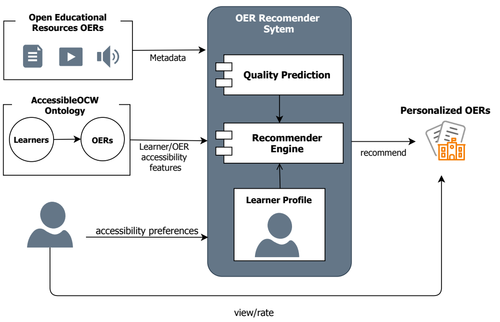
In this paper, we are reusing our AccessibleOCW ontology to represents the accessibility needs and features of OCW systems [113] and the open education recommender system [4, 82] to implement our approach. The AccessibleOCW ontology reuses and extends the *User* concept from the ACCESSIBLE ontology [122] to represent users with disabilities along with the accessibility specifications of e-learning systems as defined by the IMS Global AccessForAll (IMS AfA) [117]. The open education recommender [4, 82] is built to help learners to self develop towards skills based on their personalised needs, OER properties, and skill descriptions (from Wikipedia).

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

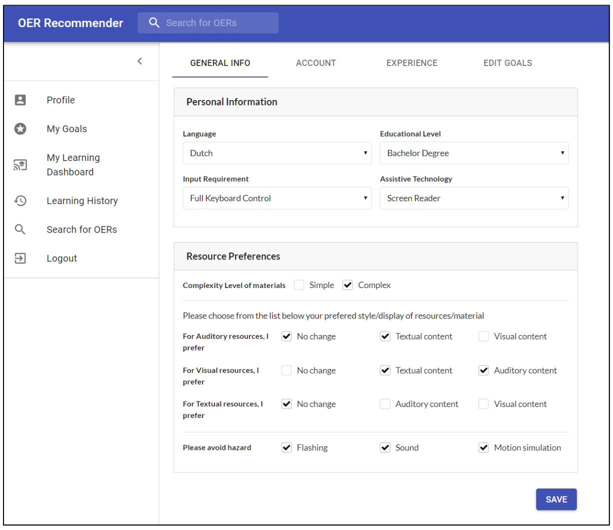
10.2.2 Accessible OERs Recommendation Approach

Our recommender system[9] uses the knowledge of the AccessibleOCW ontology[10] to describe the accessibility preferences of the learner profiles and educational resources; a learner profile with low version is described in Section 10.2.3. For this, learners are asked to create a profile and the recommender engine uses the learners' profile, OER quality prediction, and learners' ratings to recommend the best matching OER, as illustrated in Figure 10.1, and explained as follows.

Learner Profile. During the registration, a new user is asked to optionally enter: 1) Personal information (i.e., name, gender, date of birth) and current occupation, 2) Accessibility preferences, and 3) Target skills and their current level. To avoid disability disclosure, the learners have the option to initialise their own accessibility preferences for each accessibility field. For the fields that are not filled by learners, we use the list of the users with similar disabilities and initialise the accessibility fields based on their preferences. It should be mentioned that in case that we do not find similar users, we set the accessibility preferences as defined in the AccessibleOCW ontology.



(a) Accessible OER recommendation architecture



(b) Learner Profile - Accessibility Preferences

Figure 10.1: OER Recommender System supporting accessibility requirements

Quality Prediction. To predict the quality of OERs, we used the approach [26] that creates a scoring model for OER metadata, and a prediction model of OERs quality based on their metadata. The study showed that there is a tight relationship between OER metadata quality and OER quality control processes, in such a way that the more an OER passes quality control processes, the higher is the probability of containing high-quality metadata. Accordingly, the model predicts whether an OER passed the quality control process or not based on its metadata. Therefore, we applied this prediction

9 https://edoer.eu
10 https://vocol.iais.fraunhofer.de/accessibilityOnto/visualization

model on the collected OERs and removed the ones that were indicated as *Without Quality Control*.

Recommendation Engine. In order to include accessibility in our OER recommender engine, we create a 28-dimensional vector of X (according to the available accessibility list[11]) for each OER regarding their accessibilities. For this, when an OER has a specific accessibility, we set its corresponding value in the list to 1, and otherwise set the value to 0. Respectively, for each learner, we define a 28-dimensional vector P as a preference vector based on his/her accessibilities preferences that contains a float weight (between 0 and 1) for each parameter in X. The goal is to find the best weights (P vector) for each learner based on their rating satisfaction. Therefore, we use *Gradient Descend* to optimize the preference vector (P) based on users' ratings by minimizing the following loss function:

$$LossFunction = {}_{o=recommended\_OERs} |P \cdot X_o - Y_o| \tag{10.1}$$

where X_o is the 28-dimensional vector of an OER o and Y_o is the satisfaction rating (between 0 and 1) of the learner for that particular OER o. Finally, to recommend an OER to a learner u for a particular skill s, our system checks the available OERs according to the learner's occupation and the level that learner u has in skill s, and calculates cosine similarity for them to recommend the OER with the closest X vector to the user preference vector (P).

10.2.3 Learner Profile Example

Figure 10.2 represents an example of a low vision learner profile as described in the ontology, with an ontology instance (i.e., Turtle syntax), and the matching accessibility metadata of SkillsCommons.

10.2.4 Evaluation

We evaluated our recommender system through two use cases (recommended by experts): *Use case 1:* English Language educational resources that are relevant to visually impaired users, and *Use case 2:* Business educational resources that are relevant to cognitive impaired users (i.e., intellectual and neurodevelopmental disabilities).

For each use case, we went through our education dataset and filtered the OERs according to the quality prediction model and the accessibility preferences which are required by each learner profile of the use case, as defined in our previous work [113]. Afterwards, we evaluated the accessibility of the OER search results manually (e.g., NVDA tool[12] was used to simulate the activities of visually impaired

[11] To create our educational resources dataset, we used the APIs of *SkillsCommons* `https://www.skillscommons.org/`. The accessibility metadata of OER is composed of any of the following 28 accessibility features: *color, contrast, complextImageText, decorativeImages, imageAltText, hyperlinkActive, interactiveMarkup, interactivePromptText, keyboardInteractive, languageMarkup, languageMarkupAlt, multimediaAccessiblePlayer, multimediaTextTrack, multimediaTranscript, noFlickering, readingLayoutCompatible, readingLayoutPageNumbers, readingLayoutPageNumbersAlt, readingOrder, stemMarkup, stemNotationMarkup, structuralMarkupLists, structuralMarkupReaders, structuralMarkupText, tableMarkup, textAccess, textAdjustable, textAdjustmentCompatible*.

[12] `https://www.nvaccess.org/`

Learner Profiles	Low Vision		
General preferences			
Input Requirements	fullKeyboardControl		
Hazard Type Avoidance	flashing, sound, motionSimulation		
Interoperability to Assistive Technology	TRUE		
Educational Complexity Of Adaptation	simplified or enriched		
Educational Level Of Adaptation	string		
Language Of Adaptation	string		
Digital resource pereferences			
Original Access Mode	Required Access Mode	Required Adaptation Type	Required Adaptation Detail
auditory	textual, auditory	LongDescription, transcript	verbatim
colour	visual, textual	highContrast, alternativeText	enhanced
itemSize	visual	[Zooming]	enhanced
orientation	visual, textual, auditory	alternativeText, longDescription	recorded, enhanced, synthesized
position	visual, textual, auditory	alternativeText, longDescription	recorded, enhanced, synthesized
textOnImage	visual, textual, auditory	highContrast, alternativeText or longDescription	recorded, enhanced, synthesized
textual	visual	[Zooming], [CSS]	as is
visual	visual, textual, auditory	highContrast, alternativeText, transcript, e-book, audioDescription	enhanced, synthesized, realtime, recorded

AccessibleOCW ontology - Low Vision learner profile

```
## http://purl.org/accessible_ocw#Learner_lowVision
:Learner_lowVision rdf:type owl:NamedIndividual , :Learner
;
 :hasHazardAvoidance :flashing ;
 :hasReqAccessMode :auditory_auditory ,
                   :auditory_textual ,
                   :colour_textual ,
                   :orientation_textual ,
                   :position_textual ,
                   :textOnImage_textual ,
                   :textual_textual ,
                   :visual_auditory ,
                   :visual_textual ;
 :hasLanguageOfAdaptation "English"^^xsd:string ;
 :hasLanguageOfInterface "English"^^xsd:string ;
  GenericOntology:User_has_Disability :Colour_Blindness .
```

SkillsCommons accessibility metadata matching low vision profile:

[contrast, color, complextImageText, decorativeImages, imageAltText, hyperlinkActive, keyboardInteractive, multimediaAccessiblePlayer, multimediaTranscript, tableMarkup, stemMarkup, noFlickering, textAdjustable, textAdjustmentCompatible]

Figure 10.2: A low vision learner profile
The image contains an example of a low vision learner profile with all the preferences, and the ontology representation code (i.e, Turtle) and the matching accessibility metadata from SkillsCommons

users) and using automatic accessibility checking approaches (e.g., Visual ARIA bookmarklet[13]). We focused on testing the most important accessibility feature for each use case (e.g., *Use case 1*: color and contrast, headings and order, images description, and *Use case 2*: readability test, Easy-to-Read test, text adjustment, availability of visual content). In general, most of the resulting OERs passed these accessibility tests except for some checks (e.g., *Use case 1*: images that did not have alternative description, and heading order that failed in PDF format files, and *Use case 2*: the Easy-to-read test). A detailed analysis of the results can be found at https://bit.ly/30PY04C.

Finally, we selected a sample of OERs that were not retrieved by our recommender and tested their accessibility; we found that some of these OERs are not accessible because they contain scanned PDF files which are not accessible by screen readers.

Moreover, we asked three experts (for visually impaired users) and two experts (for cognitively impaired users) to rate (between 0 to 5) the quality of recommended OERs in terms of accessibility features for each of the use cases. At the end, we received more than 100 ratings regarding the recommended OERs. Table 10.3 shows the percentage of the rates in each use case. As can be seen, experts rated with a score of *3.41* out of *5* on average, which shows that our recommender system works well in satisfying these users' needs.

[13] https://whatsock.com/training/matrices/visual-aria.htm

Table 10.3: Results of the validation by experts

Use Cases	Rate=0 (%)	Rate=1 (%)	Rate=2(%)	Rate=3(%)	Rate=4(%)	Rate=5(%)	Average Rate
Use Case 1 (English Language)	0	6	7	21	33	33	3.8
Use Case 2 (Business)	2	14	19	26	24	15	3.01
Average	**≈1**	**≈10**	**≈13**	**≈23**	**≈29**	**≈24**	**3.41**

10.2.5 Conclusion and Future work

In this paper, we presented an OER recommender system that recommends OERs considering the learner's occupation, skills and accessibility preferences. Moreover, we used OER metadata, a quality prediction model, and user ratings to retrieve high quality OERs relevant to the learner's profile. Finally, to evaluate our approach, we validated the accessibility by manual and automatic checks and by collecting feedback from experts (i.e., average ratings (*3.41* out of *5*). As future work, we plan to continue adding OER repositories and validating the accessibility of OER content (of various types, such as videos, slides, or images), using accessibility guidelines and available APIs. Moreover, extracting more learner preferences (e.g. length and type of educational resources) and improving the personalisation of our recommender system are among the most important next steps.

10.3 EduCOR: An Educational and Career-Oriented Recommendation Ontology[14]

In recent years, digital education is increasingly relying on Educational Resources (ERs) and Open Educational Resources (OER). These ERs are available in many different formats, such as videos, slide decks, audio recordings from lectures, digital textbooks, or simple web pages. Furthermore, ERs and OERs usually come with low-quality metadata [90], and they are isolated from other, content-wise similar ERs. That is one of the crucial reasons for lacking high-quality services, such as recommendation and search services, based on OERs [26]. Therefore, it is not surprising that the Semantic Web (SW) community shows increased interest in organising and classifying ERs, and enhancing the metadata in publicly available ER and OER [123, 124]. Although many schemata and vocabularies were suggested in the past for the educational domain, only a few of them are still available online and can accommodate particularities of OERs, and related personalised recommendation systems' features. Furthermore, recent works revealed the increased interest in educational Knowledge Graphs [125, 126], which, however, often lack an underlying ontology or schema [127]. Commercial products seem to follow a similar direction, as they usually do not use or do not publish their underlying

[14]This chapter has been published as follows: Ilkou, E., Abu-Rasheed, H., Tavakoli, M., Hakimov, S., Kismihók, G., Auer, S. and Nejdl, W., 2021, October. EduCOR: An Educational and Career-Oriented Recommendation Ontology. In International Semantic Web Conference (pp. 546-562). Springer, Cham.

knowledge schema[15]. Additionally, surveys in e-learning have shown that an ontology helps to achieve personalised recommendation systems [128, 129]. Moreover, there is an increased interest on the education side to enrich current tools with Artificial Intelligence to achieve Smart Education. In this line, ontologies offer a wide variety of benefits for Smart Tutoring Systems [130]. In addition, the SW has a significant focus on question answering and (learning) recommendation systems. The latter is evolving rapidly to offer interoperability, explainability, and user privacy while providing personalised learning recommendations [23, 131].

On the broader community side, there is strong evidence of the everyday usage of online learning. Societies put enormous effort into the digital transformation of education, such as the Digital Educational Plan of the European Union[16], on matching work and relevant skills, and on executing skill development in online learning platforms [132]. These online learning platforms are used daily by millions of learners, especially during the COVID-19 pandemic, when education has been pushed towards online environments worldwide. Consequently, a need for lifelong learning tools emerged that could assist people in career changes, (re)skilling, or (re)entering the labor market after a period of unemployment. This trend is visible in the last decade through an increased public interest in online learning supportive platforms, such Coursera [133] for lifelong learning, or Khan Academy[17] for school education. These platforms usually contain ERs in video format and assessments to validate learners' knowledge, yet they also indicate new challenges by shifting learning towards personalised recommendations.

However, this personalisation agenda of education requires novel ways to model learning processes, especially in complex learning environments. This is especially challenging when the ingredients of the learning process are originated from the angles of education (learning content and instruction), the labor market (learning context), and individual needs of learners (learning objectives). Ontologies engineered by the SW community can play a crucial role here. While there are plenty of works available, both as e-learning and occupational ontologies, no model is available currently to connect these two domains.

Therefore, following both SW and broader community interest, we developed the Education and Career-Oriented Recommendation Ontology, the EduCOR ontology. This syntactic formalism describes ERs, skills, and the user profile in rich metadata. It creates the bridge between the demanding and constantly changing needs of the labor market and the educational domain. EduCOR provides both the basis of an educational Knowledge Graph, and serves as a potential framework for personalised, OER recommendation systems. To the best of our knowledge, the EduCOR ontology is breaking new ground on modelling ERs for a personalised recommendation system based on the learner's learning path and user profile. Moreover, EduCOR fills an essential gap in connecting personalised learning

[15] An example is the Mathspace `https://mathspace.co`, a math education platform that offers personalised learning based on a Knowledge Graph. However, its knowledge schema is not publicly available.

[16] `https://ec.europa.eu/education/education-in-the-eu/digital-education-action-plan_en`

[17] `https://www.khanacademy.org`

recommendation systems, educational data and skills with the labor market, making it a vital schema for future applications.

10.3.1 EduCOR Ontology

The EduCOR ontology is proposed to organise different domain ERs and OERs under a common ontology, link to the labor market, and offer personalised recommendation systems in the e-learning domain. A general cross-domain educational ontology should serve different purposes. Given this multidisciplinary interest and diversity of applications, there is a need for semantic representation under a unified framework that can accommodate associations between entities and attributes. We performed a requirement analysis for e-learning platforms to host personalised recommendations by reviewing the literature and an existing e-learning system. As a result, we identified the key components around which we constructed our ontology.

Ontology Composition

Our ontology introduces the necessary classes and properties to construct an e-learning environment that supports personalised recommendations. Before developing our ontology, we examined state-of-the-art related works, open standards, and best practices.

Since our goal was to create a general ontology, we limited our conceptual work to high-level, fundamental constructs. Consequently, we examined a series of open standards related to educational content, and we critically choose those that offer a wide coverage over the narrower focused ones. Thus, we adopted the widely used IEEE LOM Standard[18] and LRMI Standard[19]. Furthermore, we reuse parts from the Curriculum Course Syllabus Ontology (CCSO) [134] and schema.org[20]. Furthermore, our ontology is aligned with FAIR principles [135]. Our data are assigned globally unique and persistent identifiers, and they are described with rich metadata, which is accessible and retrievable as it is demonstrated in the ontology page[21]. We use OWL for the ontology representation, and we reuse vocabularies that follow FAIR principles and include references to them. We describe the scope of our data and have them published under the licence CC0 1.0 Universell (CC0 1.0) Public Domain Dedication[22], and it has the canonical citation: "E. Ilkou et al: EduCOR: An Educational and Career-Oriented Recommendation Ontology. April 2021. https://github.com/tibonto/educor".

Before finalising our design, we had an expert evaluation phase, where we received feedback from domain and ontology experts. The ontology also offers classes as plug-in points, where other ontologies can be mapped for more specific utilisation. Such an example is the 'Learning Preference'

[18] https://standards.ieee.org/standard/1484_12_1-2020.html
[19] https://www.dublincore.org/specifications/dublin-core/dces
[20] https://schema.org/
[21] http://ontology.tib.eu/educor
[22] https://creativecommons.org/publicdomain/zero/1.0/deed.de

that could host a thorough analysis as it is presented by CiloglugilI18. In Figure 10.3, we present a conceptual overview of the classes in EduCOR ontology with connections to a domain ontology and job ontology. A comprehensive presentation of each class's object and data properties can be found on the ontology page.

Patterns

EduCOR consists of independent modules that can be combined to create the complete schema of the ontology. We also refer to the modules as patterns. Based on our requirement analysis, we identified the key components of a personalised learning recommendation system. Taking these components as the central theme of each module presentation, we created the additional patterns, respectively. The patterns EduCOR identifies are the following: *Educational Resource*, *Knowledge Topic*, *Skill*, *Learning Path*, *Test*, *Recommendation*, *User Profile*. Each pattern stands alone and can be added to another ontology, used as a single pattern separated from the EduCOR ontology, if an application does or does not need it accordingly. In Figure 10.3, the classes of each pattern are represented in different colours.

In the *Educational Resource* pattern in Figure 10.3 pattern (A), the 'Educational Resource' class represents the learning material or learning object. It can have multiple types that are covered by the 'Multimedia Data' class. The 'Education Resource' also has a 'Quality Indicator', reflecting any quality measure required by the hosting content repository. Learners' different access requirements are covered through the 'Accessibility' class, which represents the access rights and methods of the learning material.

Each 'Educational Resource' refers to a specific 'Knowledge Topic' in *Knowledge Topic* pattern (D). Knowledge Topics represent specific themes in a particular domain of knowledge, such as the "Quadratic Equations" in the "Mathematics" domain. A 'Knowledge Topic' has a 'Theory' and an 'Exercise' content, which the learner experiences through a specific 'Methodology'. The 'Exercise' class is connected to both the *Knowledge Topic* and *Test* patterns.

In *Test* pattern (C), the 'Test' class represents the learning assessment procedure. It is composed of one or more 'Exercises', which in turn have questions and corresponding answers. A 'Test' can be composed of exercises that belong to many knowledge topics, skills, and domains.

Knowledge Topics are the requirements of achieving a target 'Skill'. The 'Skill' class, in *Skill* pattern (B), is the link between knowledge topics and the labor market job ontology.

Mastering a targeted 'Skill' and 'Knowledge Topic' can happen through their unique 'Learning Outcome'. Such 'Learning Outcome' results from the recommended 'Learning Path', in *Learning Path* pattern (E). The 'Learning Path' represents the sequence of knowledge topics needed to reach a user-defined 'Learning Goal' through the intermediate 'Learning Outcomes' of each 'Knowledge Topic' in the recommended 'Learning Path'.

The 'Recommendation' class, in *Recommendation* pattern (F), is designed to cover a range of

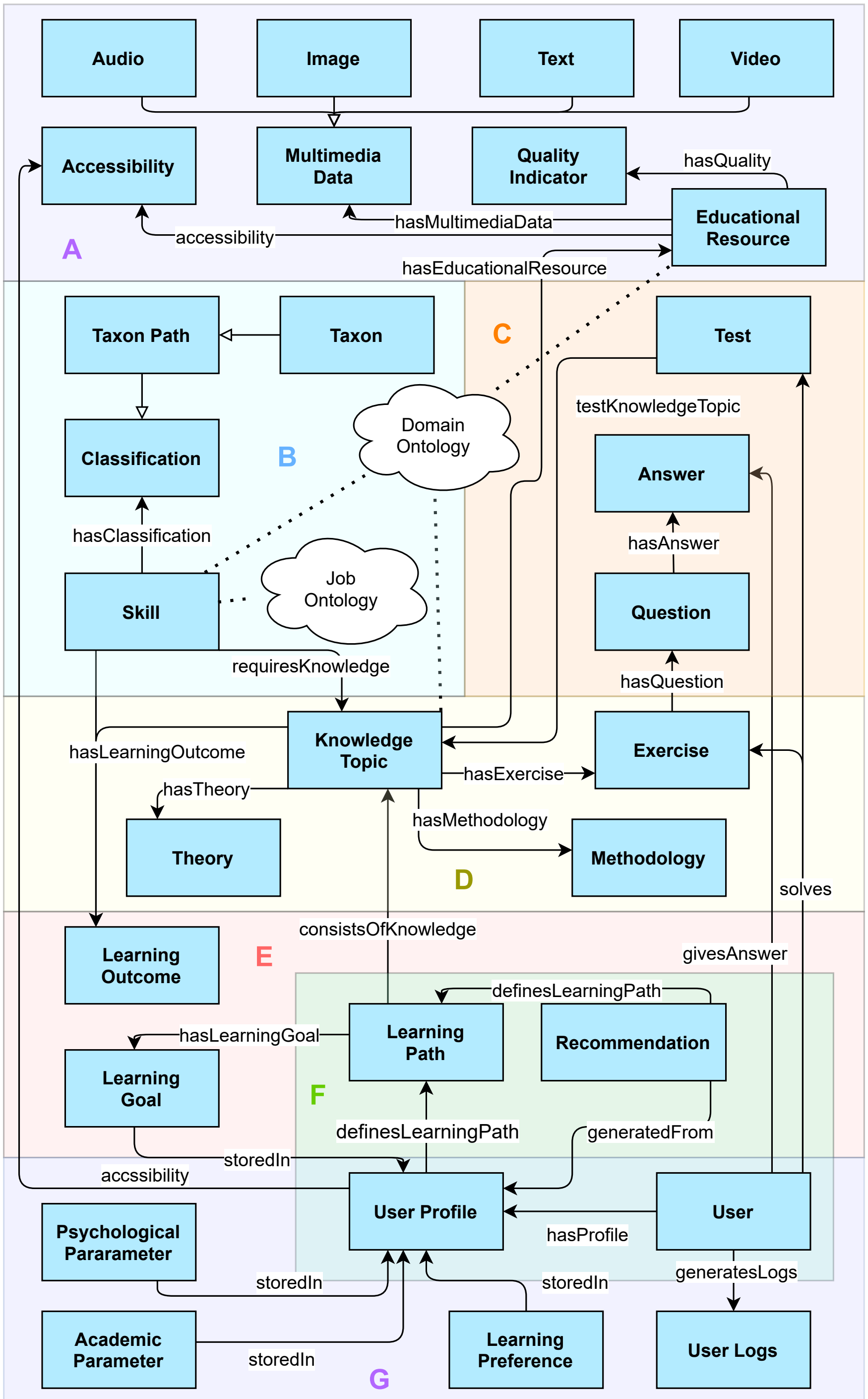


Figure 10.3: An overview of the EduCOR ontology classes. Each pattern of the ontology is highlighted individually, **A**: Educational resource pattern, **B**: Skill pattern, **C**: Knowledge topic pattern, **D**: Test pattern, **E**: Learning path pattern, **F**: Recommendation Pattern, **G**: User profile pattern.

recommended item-types based on the use-case requirements. A 'Recommendation' is directly generated from the 'User Profile', in pattern (G), which is the means of modelling the 'User' in the proposed ontology.

We design the *User Profile* to cover the interest, intention and behavioural aspects defined in [136]. Those are represented by the classes 'Learning Preference', 'Learning Goal', 'Academic Parameter', and 'Psychological Parameter'. The 'Academic Parameter' captures the learner's performance, such as test scores, while the 'Psychological Parameter' reflects the state-of-mind of the learner, such as being tired. This focus on the psychological state is due to its influence on the overall learning process and performance. The 'User Profile' is also linked to the 'Accessibility' class. The latter could describe user accessibility, content access rights, and user privacy issues.

10.3.2 Use Case Scenario

We describe a general and a specific use case. In a general use case, an OER repository owner could utilise the EduCOR ontology to model the learning materials in their repository. The repository serves learners through a standard search and information retrieval functionality. In the future, it could be possible to integrate an automatic decision-support system with minimum to zero adjustments of the repository structure.

We also used our ontology in specific use case, in the development of eDoer[23] platform, an open learning recommender system prototype, focusing on Data Science related jobs [4, 12, 82]. Since eDoer aims to empower learners through open, personalised learning and curriculum recommendations based on labor market information and OERs, the following components have been deployed using the EduCOR ontology: 1) we used the *Skill* pattern to bridge between jobs and their required qualities, 2) we applied the *Knowledge Topic* pattern to decompose each skill into relevant learning components, 3) the *Learning Path* pattern was used to create a path for learners which includes a sequence of knowledge topics towards their learning goals (i.e. target job or skills), 4) to store the required learning resources into our system, we applied the *Educational Resource* pattern , 5) in the process of building a personalised learning content recommender engine, we benefited from *Recommendation* and *User Profile* patterns to offer the most relevant learning items (i.e. knowledge topics and learning materials) to learners based on their learning goals, learning preferences, and their current knowledge level, and 6) the *Test* pattern was used to offer assessment services in order to help learners to monitor their progress towards their learning goals.

Therefore, on the eDoer platform, learners can set their target job, and the system will provide them with a list of skills they need to master for that particular job. Learners are offered to select one or more of those skills and set them as learning objectives. Moreover, learners can search through other existing skills and add different learning goals. They can also set their learning preferences,

[23] http://edoer.eu

such as the type of learning materials and the length of content, which results in personalised learning content recommendations. The generated learning path includes the target skills and the necessary knowledge topics covered for each skill. Subsequently, users receive OERs for each knowledge topic, which can be viewed, rated, and changed. Based on the users' feedback (i.e. ratings) on each of the recommendations, eDoer updates the users' preferences to capture any changes in user preferences. Moreover, there are various assessments available both on skill and knowledge topic levels that provide means to monitor the learning process[24]. Up to now, we evaluated eDoer in the context of a *Business Analytics* course at the *University of Amsterdam*. This evaluation revealed that 24 students out of 97, who worked with our system voluntarily, achieved higher course grades than those that did not.

10.3.3 Evaluation

Several evaluation methods have been introduced in the literature on ontology development. A recent survey [137] classified evaluation methods under five main approaches: 1) Gold-standard based, 2) Corpus-based or data-driven, 3) Task-based of Application-based, 4) Criteria-based, and 5) Evaluation by humans.

To ensure objectivity when evaluating EduCOR, we decided to use inductive methods following [138, 139] to select the most relevant evaluation criteria for our proposed ontology. Therefore, based on [140, 141], we focus on coverage and adaptability as key performance indicators (KPIs) of the EduCOR ontology. In the context of ER representation for learning-material repositories, the coverage is defined as the ability to describe learning materials by classes. Adaptability is defined as the potential to represent multiple repositories homogeneously. Based on these two KPIs, we conduct the gold-standard and task-based evaluation approaches. The gold-standard valuation is meant to compare EduCOR directly to other repository schemata, while the task-based evaluation is meant to validate its performance in real-world use cases. We also evaluated the proposed ontology design with experts in the ontology development domain to validate its structure and classes qualitatively.

Gold standard-based evaluation

To measure EduCOR's coverage and adaptability towards other existing ontologies, we selected three well-established repositories for ER resources, namely Merlot[25], SkilsCommons[26], and OERCommons[27]. We chose these repositories due to their richness in metadata that describes ERs and OERs. This, in turn, enabled extracting a comprehensive schema that can be used for the evaluation. Since those repositories' APIs are not open, we conducted a thorough analysis of repositories' schemas based on the information on their websites, user guides, and the use of hosted materials and resources.

[24] You can watch a demo of eDoer here: https://youtu.be/5PRcUgNa7tA
[25] https://www.merlot.org
[26] https://www.skillscommons.org
[27] https://www.oercommons.org

Table 10.4: Recall values of EduCOR as calculated for each gold schema

	OER-Commons	SkillsCommons	Merlot
EduCOR ontology	0.833	0.857	0.875

We extracted the overall class representations of the three schemas. Ultimately, these schemas are accepted as gold standards, against which the EduCOR ontology is compared. The comparison is conducted through four steps: 1) the extraction of the gold standard repositories, 2) analyzing class names and their meanings, 3) mapping EduCOR classes to the underlying schema of each repository, and 4) calculating the coverage score for each gold standard repository. Repository schemata and the four steps of comparison are elaborated in detail on EduCOR's resource page.

The mapping process refers to identifying classes in gold standard repositories that are also represented in EduCOR. Since mapping is dependent on the clear definition of a schema's own vocabulary, it may lead to a subjective evaluation. Therefore, we conducted this mapping as a multi-fold process, in which four different developers assessed the meaning of the classes in the proposed ontology and the compared schema. Once the mapping process was conducted, we sought a tangible representation of the coverage and adaptability metrics. To accomplish this task, we followed the work of [140] to calculate the recall based on the definition from the information-retrieval domain to represent the coverage of EduCOR. In this adaption, we defined the true positive value as the number of classes covered by EduCOR and existing in the gold schema. In contrast, the false negative value was defined by the number of classes in the schema that EduCOR did not cover. The calculated recall values are given in Table 10.4. They indicate the ability of the EduCOR ontology to represent data in the selected repositories with a coverage level of more than 83%. Suppose a class is not directly mapped to EduCOR. In that case, repository owners can either represent it with a different (but similar) class or datatype property from EduCOR or add it explicitly to their own schema. In other words, false negative values of the recall do not hinder adopting EduCOR as a comprehensive foundation of an ER or OER repository.

To evaluate adaptability, we refer to the definition as mentioned earlier of this measure in the context of ER repositories. Here we qualitatively assess the ability of EduCOR to represent three different repositories, which have distinct differences in focus when representing the ERs and OERs. Examples of those differences include the emphasis of Merlot on user roles, the links in SkillsCommons between ERs and industrial occupations, and the focus on educational and evaluation standards in OER-Commons. Despite those differences, our proposed ontology homogeneously represented them all, with high recall values. Moreover, EduCOR ontology provides other repositories with additional features in learning material representation, user modelling and learning recommendations. This can be seen from linking ERs and OERs to the labor market through the 'Skill' class, the inclusion of 'Psychological Parameter' in the user profile, and through the 'Recommendation' and 'Learning Path' classes that enable a personalized learning experience.

Task-based evaluation

In this step of the ontology evaluation, we defined specific tasks and evaluated EduCOR's ability to fulfil them. For the task-based evaluation, we followed the approach of chari2020explanation, where competency questions are defined to reflect the main contributions of EduCOR, based on a sample use case that is expected to be executed by a potential user of the ontology. Such a use case is described as a general use case in Section 10.3.2. This use case was designed to manifest the contributions of EduCOR in representing ERs and OERs from multiple repositories and enabling user-centric, job market-oriented learning recommendations.

From the previous use case, we define three main tasks that EduCOR should fulfill:

1. Adaptable representation of OERs from multiple sources.
2. Consideration of labor market skills in the learning path.
3. User-centred design, considering learner's academic and psychological needs within the user profile.

To evaluate EduCOR's ability of performing these tasks, the following set of questions were designed:

- Q1: How to retrieve OERs from multiple sources for a learning goal?
- Q2: How can a personalized OER difficulty be chosen for the user?
- Q3: How to provide an OER to a user with a specific access mode?
- Q4: How to retrieve required OERs for a certain job skill?
- Q5: What is required to generate a personalized learning path?
- Q6: How to personalize a learning recommendation based on a user's psychological state?

The first question Q1 reflects the adaptability metric in the evaluation of the ontology. Questions Q2 and Q3 focus on the personalisation of the retrieved material towards specific user needs, such as the difficulty levels and accessibility modes. Those questions represent the richness in data-type properties, which scaffolds the personalisation of retrieved or recommended ERs and OERs. Q4 evaluates links that the ontology draws between the ERs and the labor market needs. This allows the ER repository developer to support the users with career-oriented recommendations. Q5 and Q6 evaluate the user-centricity of the ontology. They assess the representation of the user's academic and psychological parameters in a recommendation or the retrieval of ERs. These parameters are important as they reflect the user's status, mentally and academically, which allows the recommendations to be more tailored towards their actual needs from the ERs. These competency questions are directed to the EduCOR ontology through SPARQL queries, where their answers are retrieved from any available data associated with the ontology. A sample SPARQL query is provided in Listing 1.1. The full description of queries and their answers are accessible on the documentation web page.

Listing 10.1: SPARQL query to answer the competency question Q2

```
PREFIX ec: <https://github.com/tibonto/educor#>
PREFIX dc: <http://purl.org/dcx/lrmi-vocabs/alignmentType/>

SELECT *
WHERE {
    ?test           ec:testKnowledgeTopic   ?knowResource.
    ?knowResource   ec:difficulty           ?difficulty.
    ?user           ec:solves               ?test.
    ?user           ec:hasProfile           ?userProfile.
    ?acadParam      ec:storedIn             ?userProfile.
    ?acadParam      dc:educationalLevel     ?currentLevel.
}
```

10.3.4 Related Work

Ontology development for the educational domain is not a new task. Many ontologies have been developed in the last years related to education systems and learning materials [129]. However, we find a series of issues that dated published ontologies have, such as maintainability, online availability, metadata, and their quality[28]. The biggest challenge is that most of the relevant works are not publicly available anymore. Another critical factor to consider is that the main interest in educational domain ontologies comes from educators and non-technical personnel. Therefore, the majority of these ontologies focus on educational perspectives rather than rich metadata.

In the plethora of educational and e-learning ontologies, we find the majority of ontologies in the domain of application or task-specific. Only a small minority were developed to describe the learning domain and learner data [142]. This creates a challenge in adopting such ontologies to general settings and applications. Such an example could be the recent work in ontology-based curriculum mapping by ZouriF21, which is focused on creating a core ontology for curricula and courses in higher education institutions. Such an ontology raises significant challenges when trying to fit in a general purpose e-learning environment as they cannot be mapped accurately to another domain. General domain educational ontologies are closely related to our goal; hence, we focus our analysis there.

koutsomitropoulos2018learning create an ontology-based on the IEEE LOM standard and SKOS for OER repositories. They propose an enhancement of the ER's metadata, and they link to thesauri dataset. However, they offer no personalised content capabilities. Recently ChimalakondaN20 suggested "an

[28] An example is the Medical Educational Resource Aggregator https://bioportal.bioontology.org/ontologies/MERA

ontology based modelling framework for design of educational technologies". Similar to their model, we include context and domain-specific ontology to our design and add the "GoalsOntology" as 'Learning Goal' in our system. However, in contrast to their framework, our design offers personalised recommendation features.

Another related domain in the literature are personalised recommendation systems. bulathwela2020truelearn propose an OER recommendation system based on learner background knowledge and content but without an underlying ontology. However, recent reviews show the growing significance of personalisation and recommendation systems in e-learning models, and ontologies are proven to be useful in this respect [128]. jando2017personalized show that most techniques use such an ontology to accomplish personalisation, such as the work in [143, 144]. A review by TarusNM18 presents the state-of-the-art for "ontology-based recommenders in e-learning". It points out the gained popularity of e-learning resource-recommendations and "their ability to personalise learner profiles based on the learner's characteristics, such as background knowledge, learning style, learning paths and knowledge level". It is noticeable from the state-of-the-art that despite the variety of ontology-based recommender systems in the last years, only the most recent works have developed the ontology in OWL or RDF and offer metadata descriptions. Moreover, the vast majority of publications use an ontology as a tool that provides information to a recommendation algorithm rather than integrating recommendation requirements in the ontology itself. We address this issue in EduCOR by integrating a recommendation class with the overall representation of ERs and user profile.

In terms of connecting the labor market representation with an educational ontology, one of the most related approaches is the "Ontology-based personalised course recommendation framework" by ibrahim2018ontology, which uses a course, a student and a job ontology to recommend courses and jobs. Inspired by their design, we divided the student ontology into *User Profile* and *Skill* patterns, offering personalisation capabilities, such as the 'Learning Preference' class.

User modelling plays an essential role in ontology-based recommendations [128] since the information about the user is vital to personalise the recommendation itself. EkeNSN19 present a comprehensive review on user modelling and argue that ontologies are the best solution to unify the user profile representation. GaoLW10 categorise user modelling approaches under three main classes: behavioral modelling, interest modelling and intention modelling. They show that personalisation is based on these three pillars. User profiling and content modelling are both considered inputs to a filtering algorithm, such as a recommendation system, to generate a personalised output. The content of user profiles has also been witnessing increased attention in recent years. This is also influenced by the ability to transfer the user profiles among multiple applications and domains [150]. In the educational domain, not only the academic parameters are essential in generating personalised recommendations, but also the psychological parameters, as pointed out by Fatahi19. This importance is shown in their adaptive e-learning environment study, where they showed enhanced student performance when receiving personalised recommendations. Students in their study also showed more

Table 10.5: Table comparison of the related work compared to EduCOR

Paper	FAIR	Evaluation	Data availability	Personalisation	Reuse of vocabularies
[145]	No	Yes	Yes	Goals (Learning goals)	No
[143]	No	No	No	Learning preferences, Learning style, Learner characteristics, Knowledge level, Learning activities	W3C recommendation ontology
[146]	No	Yes	No	Education information, Job related skills	No
[144]	No	No	No	Learning Style, Learning pathways	IEEE LOM
[147]	No	Yes	No	Datatype properties	IEEE LOM, thesauri, SKOS
[148]	No	Yes	No	Accessibility, Activities, Health conditions	No
[149]	No	No	No	Learning pathways	No
Ours	Yes	Yes	Yes	Learning Goal, Learning pathways, Accessibility, Learning preferences, Psychological parameter, Academic parameter, Recommendation, Datatype properties	IEEE LOM, CCSO, DCMI, SKOS, schema.org

attraction to the personalised system, since it “can understand their emotional state better”. Further, the authors in skillen2014ontological developed an ontological representation of users, putting a focus on their psychological health conditions alongside their learning-related preferences and activities. We found these previous approaches necessary in the educational field. Therefore, we expanded and complemented this set of ontological user profiling works by proposing a hybrid representation in EduCOR. As a result, in our *User Profile* pattern, static and dynamic parameters represent the learner’s both academic and psychological aspects.

Table 10.5 shows a summary of the comparison between EduCOR and those mentioned above related educational ontologies. From this summary, one can notice that EduCOR exceeds state of the art. It is aligned with the FAIR data principles and provides richer personalisation features, both in classes and datatype properties, compared to related ontologies. Furthermore, EduCOR extends these works by embedding the ‘Recommendation’ and ‘Skill’ classes in a unified representation, offering stronger links between the ERs and personalised recommendations.

10.3.5 Discussion and Future Steps

EduCOR is a publicly available, findable, registered[29], and lightweight ontology that can host ERs and OERs, personalised recommendation system features, and user profiles. It is created to address the gap between the educational domain, the labor market, and personalised learning. EduCOR can be used as a whole or as parts via the patterns introduced in Section 10.3.1. It is a semantically enhanced ontology that is adaptable. Therefore, EduCOR can be used in different educational domains, such as Computer Science, to support online learning platforms and personalised education systems. EduCOR is enriched with the necessary vocabulary and rich metadata to be general enough to be used in different settings. We leverage and maintain compatibility with existing educational repositories related to Massive Open Online Courses (MOOC) and OERs, as shown in Section 10.3.3. Moreover, we expand on them to include personalised representational primitives needed for modelling the components of a recommendation system.

However, EduCOR does not provide data specific to an application domain, and expert intervention may be necessary to seamlessly align the domain-specific ontology to the EduCOR ontology. Also, EduCOR does not offer automatic mapping of courses and curricula to its ontology. Although, this can happen by identifying courses, or chapters' learning objectives, and classifying them in skill categories with corresponding knowledge topics. An automatic alignment system for domain and task-specific ontologies mapping to EduCOR ontology is also part of future efforts.

We have implemented the basic ER and OER components that are necessary to link with the labor market and offer personalised learning. However, some aspects of OERs, and the recommendation system might need more thorough analysis. We foresee EduCOR extensions to include further analysis of some classes. The quality indicators could extend to summarize the resource multimedia and metadata quality with user's feedback ratings. Another extension could be the analysis of learning preferences, which could further link to special education coverage. Also, the accessibility analysis could expand to offer additional representations in our system, by covering user accessibility, preferences, and content access rights. In this line, we could additionally focus on the user's privacy, which at the moment boils down to each developer's implementation plan to decide how to implement This work will additionally aim to assist in the user privacy and profile restrictions alignment with our ontology.

In future work, we plan to publish an Open Educational Knowledge Graph, connecting educational resources with the labor market while offering personalised recommendation features by combining ERs from multiple sources. Upon identifying the appropriate content and repositories, we wish to gather the requirements and publish the Knowledge Graph based on the EduCOR ontology. Therefore, we foresee a sustainability plan for the following years as we plan to use the EduCOR ontology as the basis of our future work. We are committed to its maintenance and extensibility to address future

[29]You can find EduCOR's presentation at http://ontology.tib.eu/educor and on our GitHub page at https://github.com/tibonto/educor

challenges and meet future requirements.

10.3.6 Conclusion

We have built an open-source, free access ontology to model educational resources, personalised learning recommendations, user profiles, and labor market skills. We argued that this interdisciplinary attempt is vital both for the SW, educators and the broader community. Our requirement analysis came from reviewing the literature and an existing e-learning system that revealed the key components of a perspective system around which we built our ontology. We presented our design and ontological components, which adopt open community standards and FAIR data principles. We evaluated EduCOR with gold-standard and task-based approaches and showed that the EduCOR ontology achieves high coverage of multiple OER repositories. Through a carefully crafted set of competency questions, we evaluated the capabilities of EduCOR in assisting the system designers in e-learning based recommendation systems to determine the necessary elements for their design. We believe our ontology can be a beneficial tool for system designers as they implement personalised features in their recommendation system. We are committed to continuing this line of work towards supporting future requirements that would extend our ontology.

CHAPTER **11**

Conclusion and Future Work

In this thesis, we argued that in the broad domain of informal learning, there is a pressing need for intelligent learning recommender systems that exploit the hidden potential of OERs. To meet this need, we showcased the development of a system, eDoer, with two major interlinked components to provide open personalized education: 1) a human-AI curriculum and educational content curation mechanism, to create a quality-assured knowledge base for informal learning, and 2) a personalized learning dashboard, which supports learners in building their own curricula on the basis of the structure and the learning content in the knowledge base, and adapted to learners' own learning objectives and contexts.

11.1 Summary of the Findings

In order to implement our system, at first we started with developing a prototype based on our idea to gather the preliminary key learner requirements for such a personalized learning environment (see *RQ1*) which showed that our idea of providing personalized education toward learners' goals (e.g., labor market driven goals) was worth investigating. Accordingly, we decomposed our proposed idea into components and started working on them separately. The first component was about investigating the effects of property analysis of educational content in building personalized education (see *RQ2*), which showed that having information (e.g., quality level and metadata) about educational resources are key for offering such personalized educational services. Afterward, we conducted research on automatic, metadata-based quality prediction of educational resources, as it is essential for recommending high-quality content to learners due to the huge number of resources published on a daily basis (see *RQ3*). The outcome of this study showed that we were able to build a machine-learning model that predicts the quality of educational resources according to their metadata. During the metadata analysis of educational resources, we revealed that extracting the target topics covered by educational content is extremely important for building learning pathways for individual learners. Therefore, we developed a text mining approach, which extracts the existing learning topics in educational resources (see *RQ4*).

Based on the components developed through *RQ2-RQ4*, in order to make our solution dynamic and scalable, we built a hybrid Human-AI curriculum development system that empowers educational service providers to create, validate and maintain up-to-date curricula (See *RQ5*). Using this curriculum development system and a systematic requirement analysis, we implemented our open, personalized educational dashboard in which learners can 1) set their learning goals, 2) receive a personalized path toward achieving their goals, and 3) monitor and assess their learning process in order to reflect on their progress (See *RQ6*). We also evaluated our prototype through an experiment in the context of a fundamental engineering skill (i.e. *Basic Statistics*). This validation indicated that learners benefited from receiving recommendations from *eDoer*, and particularly so when such recommendations were personalized, evidenced by higher scores on the posttest, as compared with self-directed learning (outside eDoer).

The hypothesized findings for the difference between learners, who received personalized content as opposed to those who received non-personalized content (i.e. randomly selected content) were less convincing, in that our most conservative test of this hypothesis, failed to reach statistical significance. Having said that, we should remind ourselves that personalization is a feature of our tool, and that based on the findings for hypothesis 1 (see 9.2), we may conclude that it made a difference to students' learning, despite the effect pertaining to the difference between the personalized and non-personalized group not reaching statistical significance. When it comes to the lack of support for hypothesis 2 (see 9.2), one explanation is that both the personalization and the non-personalization group received quality content, and that in some instances members of the non-personalization group may in fact have received personalized content by chance (according to the limited number of educational resources that were offered for each topic). This would mean that those members contaminated what ought to have been an all non-personalization group with some degree of personalization, therewith reducing the effect size.

11.2 Limitations and Future Work

The initial results of our work are promising in that they seem to indicate that engagement with eDoer, particularly when it offers personalized recommendations pertaining to statistics, appears to contribute to knowledge acquisition. Nevertheless, and as with all research, clearly there are a number of limitations that need to be acknowledged. First off, the sample size of our requirements gathering was quite limited, in that learners in different contexts, at different levels, and of different ages, and from different cultures may have different requirements that we have yet to learn about. Furthermore, people with (learning) disabilities also have needs that are not addressed by the current rendition of the system. A related challenge we faced in the requirements gathering process was how to reconcile free text input (in which we could qualitatively identify all the different requirements that learners felt needed to be addressed) with the ranking of these same requirements (with which we could determine which requirements were most important). Future work must be carried out to identify and address these needs, particularly if eDoer is to contribute to meeting the United Nations Sustainable Development Goal of providing inclusive and equitable quality education and promoting lifelong learning opportunities for all, as was suggested in the introduction.

Despite the positive validation results, there are also several issues that are noteworthy with regard to our experiment. Our validation comprised a limited sample of learners, studying but a single topic for a very limited amount of time. It remains to be seen whether results will be equally promising when eDoer is deployed in different contexts (for instance with unpaid learners, refugees, and/or those seeking to qualify themselves for a new occupation), in other cultures, with other learning content, and for a longer duration. To illustrate our point about duration, when we examined well-known courses on basic statistics from *Standford University*[1], University of Amsterdam[2], and *Khan Academy*[3], for instance, we determined that their students spend an average of 10 hours (600 minutes) to master the aforementioned topics on basic statistics. Given that the current study established a treatment effect for what constituted but a very limited 'dosage' of training, strengthens us in the belief that stronger effects can be booked with trainings of greater duration and depth. Clearly, however, future research is needed to further develop and evidence this tool, with different samples, different topics, and training of greater durations.

In addition to training duration, one may also wonder about the longer term retention of that which was learned, in that our posttest was administered quite soon after the training. Future research will need to examine the extent to which that which was learned is retained over time. Here too, however, we feel that retention is only likely to improve with trainings of greater duration.

Based on the feedback and the lessons we learned during the development process, we also conclude that more work needs to be done on the personalization and scalability components of our prototype. Specifically, to personalize the learning experience, we collected several initial personal features from learners (i.e. length, level of detail, learning strategy, and content format (9.1.5)). However, this still needs to be extended to describe the learners' context in a fine-grained manner. Therefore we see value in capturing more preference features in the future, such as language preferences, preferred authors, location, or sensory information on learners' cognitive and mental state (e.g. tiredness, well-being).

Moreover, currently, we use long-term and short-term vectors to plot learner preferences. At the moment, it puts more emphasis on their recent feedback about learning content they studied (9.1.5). In the user profile, however, learners can edit their long-term vector (the basis for recommendations (see 9.1.5)) directly, which overwrites their preference score, computed by our model, based on actual learner feedback and behavior (see 9.1.5). Therefore, we will need to fine-tune this scoring algorithm

[1] https://www.coursera.org/learn/stanford-statistics
[2] https://www.coursera.org/learn/basic-statistics
[3] https://www.youtube.com/watch?v=uhxtUt_-GyM.&list=PL1328115D3D8A2566&ab_channel=KhanAcademy

by, for instance, providing an option for learners to decide about the balance between their long-term and short-term vectors.

APPENDIX A

List of Publications

1. **An AI-based open recommender system for personalized labor market driven education.** Tavakoli, M., Faraji, A., Vrolijk, J., Molavi, M., Mol, S.T. and Kismihók, G., 2022. Advanced Engineering Informatics Journal.

2. **Hybrid Human-AI Curriculum Development for Personalised Informal Learning Environments.** Tavakoli, M., Faraji, A., Molavi, M., Mol, S.T. and Kismihók, G., 2022. In LAK22: 12th International Learning Analytics and Knowledge Conference.

3. **Metadata analysis of open educational resources.** Tavakoli, M., Elias, M., Kismihók, G. and Auer, S., 2021. In LAK21: 11th International Learning Analytics and Knowledge Conference.

4. **Extracting Topics from Open Educational Resources.** Molavi, M., Tavakoli, M. and Kismihók, G., 2020. Extracting Topics from Open Educational Resources. In European Conference on Technology Enhanced Learning (EC-TEL).

5. **Quality prediction of open educational resources a metadata-based approach.** Tavakoli, M., Elias, M., Kismihok, G. and Auer, S., 2020. In 2020 IEEE 20th international conference on advanced learning technologies (ICALT).

6. **A recommender system for open educational videos based on skill requirements.** Tavakoli, M., Hakimov, S., Ewerth, R. and Kismihok, G., 2020. In 2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT).

7. **OER recommendations to support career development.** Tavakoli, M., Faraji, A., Mol, S.T. and Kismihók, G., 2020. In 2020 IEEE Frontiers in Education Conference (FIE).

8. **Labour market information driven, personalized, OER recommendation system for lifelong learners.** Tavakoli, M., Mol, S.T. and Kismihók, G., 2020. In the 12th Computer Supported Education Conference (CSEDU).

9. **Improving the Quality of Posts in the Stack Overflow.** Tavakoli, MohammadReza, Abbas Heydarnoori, submitted to 27th IEEE International Conference on Software Analysis, Evolution and Reengineering. IEEE, 2019.

10. **Requirement Analysis Towards Building a Personalized OER Recommender, based on Labour Market Information** Tavakoli, Mohammadreza, Gabor Kismihok, Stefan Mol. In the Proceedings of Learning and Student Analytics Conference (LSAC), LSAC, 2019.

11. **Customer Segmentation and Strategy Development based on RFM analysis and Data Mining techniques: A case study** Tavakoli, MohammadReza, MohammadReza Molavi, Vahid Masoumi, Majid Mobini. In the Proceedings of International Conference on E-Business Engineering (ICEBE), IEEE, 2018.

12. **Improving the quality of code snippets in stack overflow.** Tavakoli, MohammadReza, Abbas Heydarnoori, and Mohammad Ghafari. Proceedings of the 31st Annual ACM Symposium on Applied Computing. ACM, 2016.

Bibliography

[1] N. Sclater, "Open educational resources: Motivations, logistics and sustainability", *Content management for e-learning*, Springer, 2011 179 (cit. on p. 1).

[2] V. B. Kobayashi, S. T. Mol, H. A. Berkers, G. Kismihók and D. N. Den Hartog, *Text mining in organizational research*, Organizational research methods **21** (2018) 733 (cit. on p. 1).

[3] M. Tavakoli et al., *An AI-based open recommender system for personalized labor market driven education*, Advanced Engineering Informatics **52** (2022) 101508 (cit. on p. 2).

[4] M. Tavakoli, G. Kismihok and S. T. Mol, "Labour Market Information Driven, Personalized, OER Recommendation System for Lifelong Learners", SciTePress, 2020 (cit. on pp. 2, 47, 51, 56, 66, 87, 88, 97).

[5] F. Wang, Z. Jiang, X. Li and G. Li, *Cognitive factors of the transfer of empirical engineering knowledge: A behavioral and fNIRS study*, Advanced Engineering Informatics **47** (2021) 101207 (cit. on p. 2).

[6] E. Colombo, F. Mercorio and M. Mezzanzanica, *Applying machine learning tools on web vacancies for labour market and skill analysis*, Terminator or the Jetsons? The Economics and Policy Implications of Artificial Intelligence (2018) (cit. on pp. 2, 14).

[7] V. Castello et al., "Promoting dynamic skills matching: challenges and evidences from the smart project", *INTED2014 Proceedings*, Citeseer, 2014 2430 (cit. on pp. 2, 14).

[8] X. Li, Z. Jiang, Y. Guan, G. Li and F. Wang, *Fostering the transfer of empirical engineering knowledge under technological paradigm shift: An experimental study in conceptual design*, Advanced Engineering Informatics **41** (2019) 100927 (cit. on p. 2).

[9] I. Wowczko, "Skills and vacancy analysis with data mining techniques", *Informatics*, vol. 2, 4, Multidisciplinary Digital Publishing Institute, 2015 31 (cit. on pp. 2, 14).

[10] V. B. Kobayashi, S. Mol and G. Kismihok, *Labour market driven learning analytics*, Journal of Learning Analytics **1** (2014) 207 (cit. on p. 2).

[11] M. M. McGill, "Defining the expectation gap: a comparison of industry needs and existing game development curriculum", *Proceedings of the 4th International Conference on Foundations of Digital Games*, ACM, 2009 129 (cit. on p. 2).

[12] M. Tavakoli, A. Faraji, S. T. Mol and G. Kismihók, *OER Recommendations to Support Career Development*, IEEE Frontiers in Education (FIE) (2020) (cit. on pp. 2, 66, 97).

[13] B. B. Lockee, *Online education in the post-COVID era*, Nature Electronics **4** (2021) 5 (cit. on p. 2).

[14] L. Zhang, J. D. Basham and S. Yang, *Understanding the implementation of personalized learning: A research synthesis*, Educational Research Review (2020) 100339 (cit. on p. 2).

[15] M. d. C. Saraiva et al., *Relationships among educational materials through the extraction of implicit topics: Relacionamentos entre materiais didáticos através da extração de tópicos implıcitos*, (2019) (cit. on pp. 2, 3, 17).

[16] M. de Carvalho Saraiva and C. B. Medeiros, "Finding out topics in educational materials using their components", *2017 IEEE Frontiers in Education Conference (FIE)*, IEEE, 2017 1 (cit. on pp. 2, 3, 17).

[17] M. Molavi, M. Tavakoli and G. Kismihók, "Extracting Topics from Open Educational Resources", *European Conference on Technology Enhanced Learning*, Springer, 2020 455 (cit. on pp. 2, 3, 18, 69).

[18] A. Kanwar and S. Mishra, *Global Trends in OER: What is the Future?*, (2018) (cit. on p. 3).

[19] K.-H. Ha et al., "A novel approach towards skill-based search and services of Open Educational Resources", *Research Conference on Metadata and Semantic Research*, Springer, 2011 312 (cit. on p. 3).

[20] G. Sun, T. Cui, D. Xu, J. Shen and S. Chen, "A heuristic approach for new-item cold start problem in recommendation of micro open education resources", *International conference on intelligent tutoring systems*, Springer, 2018 212 (cit. on pp. 3, 15).

[21] A. Ruiz-Iniesta, G. Jimenez-Diaz and M. Gomez-Albarran, *A semantically enriched context-aware OER recommendation strategy and its application to a computer science OER repository*, IEEE Transactions on Education **57** (2014) 255 (cit. on pp. 3, 15).

[22] J. Chicaiza, N. Piedra, J. Lopez-Vargas and E. Tovar-Caro, "A user profile definition in context of recommendation of open educational resources. An approach based on linked open vocabularies", *IEEE Frontiers in Education Conference*, IEEE, 2015 1 (cit. on pp. 3, 15).

[23] J. Chicaiza, N. Piedra, J. Lopez-Vargas and E. Tovar-Caro, "Recommendation of open educational resources. An approach based on linked open data", *Global Engineering Education Conference*, IEEE, 2017 1316 (cit. on pp. 3, 15, 93).

[24] S. Wan and Z. Niu, *An e-learning recommendation approach based on the self-organization of learning resource*, Knowledge-Based Systems **160** (2018) 71 (cit. on pp. 3, 15).

[25] G. Sun et al., *Towards massive data and sparse data in adaptive micro open educational resource recommendation: a study on semantic knowledge base construction and cold start problem*, Sustainability **9** (2017) 898 (cit. on pp. 3, 15).

[26] M. Tavakoli, M. Elias, G. Kismihok and S. Auer, "Quality Prediction of Open Educational Resources - A Metadata-based Approach", IEEE, 2020 (cit. on pp. 3, 17, 18, 37, 38, 41, 42, 71, 86, 89, 92).

[27] J. Djumalieva and C. Sleeman, *An open and data-driven taxonomy of skills extracted from online job adverts*, Developing Skills in a Changing World of Work: Concepts, Measurement and Data Applied in Regional and Local Labour Market Monitoring Across Europe (2018) 425 (cit. on pp. 3, 14).

[28] J. Wang, J. Xiang and K. Uchino, "Topic-specific recommendation for open education resources", *International Conference on Web-Based Learning*, Springer, 2015 71 (cit. on pp. 3, 17).

[29] Wikipedia, *Educational technology*, https://en.wikipedia.org/wiki/Educational_technology, 2022 (cit. on p. 7).

[30] weforum, *These 3 charts show the global growth in online learning*, https://www.weforum.org/agenda/2022/01/online-learning-courses-reskill-skills-gap/, 2022 (cit. on p. 8).

[31] Wikipedia, *Open Education*, https://en.wikipedia.org/wiki/Open_education, 2022 (cit. on p. 8).

[32] Wikipedia, *Curriculum Development*, https://en.wikipedia.org/wiki/Curriculum_development, 2022 (cit. on p. 9).

[33] Wikipedia, *Recommender System*, https://en.wikipedia.org/wiki/Recommender_system, 2022 (cit. on p. 9).

[34] Wikipedia, *Goal Setting*, https://en.wikipedia.org/wiki/Goal_setting, 2022 (cit. on p. 10).

[35] Wikipedia, *Artificial Intelligence*, https://en.wikipedia.org/wiki/Artificial_intelligence, 2022 (cit. on p. 11).

[36] Wikipedia, *Crowdsourcing*, https://en.wikipedia.org/wiki/Crowdsourcing, 2022 (cit. on p. 11).

[37] E. M. Sibarani, S. Scerri, C. Morales, S. Auer and D. Collarana, "Ontology-guided job market demand analysis: a cross-sectional study for the data science field", *Proceedings of the 13th International Conference on Semantic Systems*, ACM, 2017 25 (cit. on p. 14).

[38] M. Khobreh et al., *An ontology-based approach for the semantic representation of job knowledge*, IEEE Transactions on Emerging Topics in Computing **4** (2015) 462 (cit. on p. 14).

[39] M. Hepp, *Possible ontologies: How reality constrains the development of relevant ontologies*, IEEE Internet Computing **11** (2007) 90 (cit. on p. 14).

[40] F. Colace et al., "Towards Labour Market Intelligence through Topic Modelling", *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019 (cit. on p. 14).

[41] I. Karakatsanis et al., *Data mining approach to monitoring the requirements of the job market: A case study*, Information Systems **65** (2017) 1 (cit. on p. 14).

[42] R. Boselli, M. Cesarini, F. Mercorio and M. Mezzanzanica, *Classifying online job advertisements through machine learning*, Future Generation Computer Systems **86** (2018) 319 (cit. on p. 14).

[43] R. Boselli et al., *WoLMIS: a labor market intelligence system for classifying web job vacancies*, Journal of Intelligent Information Systems **51** (2018) 477 (cit. on p. 14).

[44] A. Verma, K. M. Yurov, P. L. Lane and Y. V. Yurova, *An investigation of skill requirements for business and data analytics positions: A content analysis of job advertisements*, Journal of Education for Business **94** (2019) 243 (cit. on p. 14).

[45] A. Gardiner, C. Aasheim, P. Rutner and S. Williams, *Skill requirements in big data: A content analysis of job advertisements*, Journal of Computer Information Systems **58** (2018) 374 (cit. on p. 14).

[46] M. M. Maer-Matei, C. Mocanu, A.-M. Zamfir and T. M. Georgescu, *Skill Needs for Early Career Researchers—A Text Mining Approach*, Sustainability **11** (2019) 2789 (cit. on p. 14).

[47] X. N. Lam, T. Vu, T. D. Le and A. D. Duong, "Addressing cold-start problem in recommendation systems", *Proceedings of the 2nd international conference on Ubiquitous information management and communication*, ACM, 2008 208 (cit. on p. 15).

[48] J. Lopez-Vargas, N. Piedra, J. Chicaiza and E. Tovar, "Recommendation of OERs shared in social media based-on social networks analysis approach", *IEEE Frontiers in Education Conference*, IEEE, 2014 1 (cit. on p. 15).

[49] J. Duffin, B. Muramatsu and S. Henson Johnson, *OER Recommender: A recommendation system for open educational resources and the National Science Digital Library*, White paper funded by the Andrew W. Mellon Foundation for the Folksemantic. org project (2007) (cit. on p. 15).

[50] T. R. Bruce and D. I. Hillmann, "The continuum of metadata quality: defining, expressing, exploiting", *Metadata in Practice*, ALA editions, 2004 (cit. on p. 16).

[51] X. Ochoa and E. Duval, *Quality Metrics for Learning Object Metadata*, World Conference on Educational Multimedia, Hypermedia and Telecommunications (2006) (cit. on p. 16).

[52] M. Elias, A. Oelen, M. Tavakoli, G. Kismihok and S. Auer, "Quality Evaluation of Open Educational Resources", *Proceedings of the 15th European Conference on Technology-Enhanced Learning (EC-TEL 2020)*, Springer, 2020 (cit. on p. 16).

[53] X. Ochoa and E. Duval, *Automatic evaluation of metadata quality in digital repositories*, International journal on digital libraries **10** (2009) 67 (cit. on p. 16).

[54] A. R. Pelaez and P. P. Alarcon, "Metadata quality assessment metrics into OCW repositories", *Proceedings of the 2017 9th International Conference on Education Technology and Computers*, ACM, 2017 253 (cit. on p. 16).

[55] N. Palavitsinis, N. Manouselis and S. Sanchez-Alonso, *Metadata quality in learning object repositories: a case study*, The Electronic Library (2014) (cit. on p. 16).

[56] M. E. Phillips, O. L. Zavalina and H. Tarver, "Using metadata record graphs to understand digital library metadata", *International Conference on Dublin Core and Metadata Applications*, 2020 49 (cit. on p. 16).

[57] A. Romero-Pelaez, V. Segarra-Faggioni, N. Piedra and E. Tovar, "A Proposal of Quality Assessment of OER Based on Emergent Technology", *2019 IEEE Global Engineering Education Conference (EDUCON)*, IEEE, 2019 1114 (cit. on p. 16).

[58] M. Margaritopoulos, T. Margaritopoulos, I. Mavridis and A. Manitsaris, *Quantifying and measuring metadata completeness*, Journal of the American Society for Information Science and Technology **63** (2012) 724 (cit. on p. 16).

[59] A. Romero-Pelaez, V. Segarra-Faggioni and P. P. Alarcon, "Exploring the provenance and accuracy as metadata quality metrics in assessment resources of OCW repositories", *Proceedings of the 10th International Conference on Education Technology and Computers*, ACM, 2018 292 (cit. on p. 16).

[60] D. Gavrilis et al., "Measuring quality in metadata repositories", *International Conference on Theory and Practice of Digital Libraries*, Springer, 2015 56 (cit. on pp. 16, 17).

[61] B. P. Nunes, R. Kawase, B. Fetahu, M. A. Casanova and G. H. B. de Campos, "Educational forums at a glance: Topic extraction and selection", *International Conference on Web Information Systems Engineering*, Springer, 2014 351 (cit. on p. 17).

[62] A. R. Fabbri et al., *Tutorialbank: A manually-collected corpus for prerequisite chains, survey extraction and resource recommendation*, arXiv preprint arXiv:1805.04617 (2018) (cit. on p. 17).

[63] A. Garcıa-Floriano et al., *Social web content enhancement in a distance learning environment: intelligent metadata generation for resources*, International Review of Research in Open and Distributed Learning **18** (2017) 161 (cit. on p. 17).

[64] M. Xie, C. Wu and Y. Zhang, *A New Intelligent Topic Extraction Model on Web.*, JCP **6** (2011) 466 (cit. on p. 17).

[65] M. Somasundaram, P. Latha and S. S. Pandian,
Curriculum Design Using Artificial Intelligence (AI) Back Propagation Method,
Procedia Computer Science **172** (2020) 134 (cit. on p. 18).

[66] M. K. Pattanshetti, S. Jasola, V. Gupta and A. Rajput,
The open corpus challenge in eLearning,
Knowledge Management & E-Learning: An International Journal **10** (2018) 67 (cit. on p. 18).

[67] M. K. Pattanshetti, S. Jasola, A. Rajput and V. Pant,
"Proposed eLearning Framework using Open Corpus Web Resources",
2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), IEEE, 2021 1 (cit. on p. 18).

[68] H. Jelodar et al.,
Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey,
Multimedia Tools and Applications **78** (2019) 15169 (cit. on pp. 18, 48, 59, 69).

[69] Y. Jiang, D. Schlagwein and B. Benatallah,
"A Review on Crowdsourcing for Education: State of the Art of Literature and Practice.",
PACIS, 2018 180 (cit. on p. 18).

[70] A. Cross, M. Bayyapunedi, D. Ravindran, E. Cutrell and W. Thies,
"VidWiki: Enabling the crowd to improve the legibility of online educational videos",
Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing, 2014 1167 (cit. on p. 18).

[71] V. Pandey et al., "Gut instinct: Creating scientific theories with online learners",
Proceedings of the 2017 CHI conference on human factors in computing systems, 2017 6825 (cit. on p. 18).

[72] D. S. Weld et al., "Personalized online education—a crowdsourcing challenge",
Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence, 2012 (cit. on p. 18).

[73] A. Farasat, A. Nikolaev, S. Miller and R. Gopalsamy,
"Crowdlearning: Towards collaborative problem-posing at scale",
Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale, 2017 221 (cit. on p. 18).

[74] O. Stewart, D. Lubensky and J. M. Huerta,
"Crowdsourcing participation inequality: a SCOUT model for the enterprise domain",
Proceedings of the ACM SIGKDD Workshop on Human Computation, 2010 30 (cit. on p. 19).

[75] K. Yang and H. Qi, *The Nonlinear Impact of Task Rewards and Duration on Solvers' Participation Behavior: A Study on Online Crowdsourcing Platform*, Journal of Theoretical and Applied Electronic Commerce Research **16** (2021) 709 (cit. on p. 19).

[76] A. Joulin, E. Grave, P. Bojanowski and T. Mikolov, *Bag of tricks for efficient text classification*, arXiv preprint arXiv:1607.01759 (2016) (cit. on p. 22).

[77] M. Elias, S. Lohmann and S. Auer, "Ontology-based representation of learner profiles for accessible opencourseware systems", *International Conference on Knowledge Engineering and the Semantic Web*, Springer, 2017 279 (cit. on p. 24).

[78] N. Piedra, J. Chicaiza, J. López-Vargas and E. T. Caro, *Seeking Open Educational Resources to Compose Massive Open Online Courses in Engineering Education An Approach based on Linked Open Data.*, J. UCS **21** (2015) 679 (cit. on p. 24).

[79] J. Atenas and L. Havemann, *Quality assurance in the open: an evaluation of OER repositories*, INNOQUAL: The International Journal for Innovation and Quality in Learning **1** (2013) 22 (cit. on p. 24).

[80] M. Elias, A. James, S. Lohmann, S. Auer and M. Wald, "Towards an open authoring tool for accessible slide presentations", *International Conference on Computers Helping People with Special Needs*, Springer, 2018 172 (cit. on p. 24).

[81] J. Pennington, R. Socher and C. D. Manning, "GloVe: Global Vectors for Word Representation", *Empirical Methods in Natural Language Processing (EMNLP)*, 2014 1532, URL: http://www.aclweb.org/anthology/D14-1162 (cit. on p. 32).

[82] M. Tavakoli, S. Hakimov, R. Ewerth and G. Kismihok, "A Recommender System For Open Educational Videos Based On Skill Requirements", IEEE, 2020 (cit. on pp. 38, 48, 71, 88, 97).

[83] T. U. of Arizona, *Hot Topics in Health Care*, https://opa.uahs.arizona.edu/outreach/speakers-bureau-topics, 2020 (cit. on p. 38).

[84] Wikipedia, *Standard score/Z-score*, https://en.wikipedia.org/wiki/Standard_score, 2020 (cit. on p. 41).

[85] A.-N. Moldovan, I. Ghergulescu and C. H. Muntean, *VQAMap: A novel mechanism for mapping objective video quality metrics to subjective MOS scale*, IEEE Transactions on Broadcasting **62** (2016) 610 (cit. on p. 43).

[86] M. Röder, A. Both and A. Hinneburg, "Exploring the space of topic coherence measures", *Proceedings of the eighth ACM international conference on Web search and data mining*, 2015 399 (cit. on pp. 48, 59).

[87] D. M. Blei, A. Y. Ng and M. I. Jordan, *Latent dirichlet allocation*, Journal of Machine Learning Research **3** (2003) 993 (cit. on p. 48).

[88] K. Papineni, S. Roukos, T. Ward and W.-J. Zhu, "Bleu: a method for automatic evaluation of machine translation", *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, 2002 311 (cit. on p. 56).

[89] J. Ramos et al., "Using tf-idf to determine word relevance in document queries", *Proceedings of the first instructional conference on machine learning*, vol. 242, 1, Citeseer, 2003 29 (cit. on p. 58).

[90] M. Tavakoli, M. Elias, G. Kismihók and S. Auer, "Metadata Analysis of Open Educational Resources", *LAK21: 11th International Learning Analytics and Knowledge Conference*, 2021 626 (cit. on pp. 60, 71, 92).

[91] T. F. Hawk and A. J. Shah, *Using learning style instruments to enhance student learning*, Decision Sciences Journal of Innovative Education **5** (2007) 1 (cit. on p. 72).

[92] Wikipedia, *Dot Product*, `https://en.wikipedia.org/wiki/Dot_product`, 2020 (cit. on p. 74).

[93] Mozilla, *Responsive Design*, `https://developer.mozilla.org/en-US/docs/Learn/CSS/CSS_layout/Responsive_Design#responsive_design`, 2020 (cit. on p. 75).

[94] Google, *Material Design*, `https://material.io/`, 2020 (cit. on p. 75).

[95] Introjs, *Introduce users to your product*, `https://introjs.com/`, 2020 (cit. on p. 75).

[96] J. K. Kruschke, *Bayesian estimation supersedes the t test.*, Journal of Experimental Psychology: General **142** (2013) 573 (cit. on p. 81).

[97] OEC, *What is Open Courseware?*, `https://www.oeconsortium.org/faq/what-is-open-courseware/` (cit. on pp. 83, 85, 87).

[98] S. Palvia et al., *Online Education: Worldwide Status, Challenges, Trends, and Implications*, Journal of Global Information Technology Management **21** (2018) 233 (cit. on p. 83).

[99] S. Vahdati et al., "OpenCourseWare observatory: does the quality of OpenCourseWare live up to its promise?", *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, 2015 73 (cit. on p. 84).

[100] G. Moise et al., "MASECO: A multi-agent system for evaluation and classification of OERs and OCW based on quality criteria", *E-Learning Paradigms and Applications*, Springer, 2014 185 (cit. on p. 84).

[101] J. Atenas and L. Havemann, *Questions of quality in repositories of open educational resources: a literature review*, Research in Learning Technology **22** (2014) (cit. on p. 84).

[102] H. Leary et al., *Developing and using a guide to assess learning resource quality in educational digital libraries*, Digital libraries-methods and applications (2011) 181 (cit. on p. 84).

[103] E. Kurilovas et al., *Methodology for Evaluating Quality and Reusability of Learning Objects*, Electronic Journal of e-Learning **9** (2011) 39 (cit. on p. 84).

[104] A. Achieve, *Rubrics for evaluating open education resource (OER) objects*, Washington, DC: Achieve, Inc. Retrieved January **9** (2011) 2013 (cit. on p. 84).

[105] M. Pérez-Mateo et al., *Learner generated content: Quality criteria in online collaborative learning*, EU Journal of Open, Distance and E-Learning **14** (2011) (cit. on p. 84).

[106] J. Nesbit and J. Li, "Web-based tools for learning object evaluation", *International conference on education and information systems: Technologies and Applications*, 2004 21 (cit. on p. 84).

[107] M. Haughey and B. Muirhead, *Evaluating learning objects for schools.*, E-Journal of Instructional Science and Technology **8** (2005) (cit. on p. 84).

[108] M. Custard and T. Sumner, *Using machine learning to support quality judgments*, D-Lib Magazine **11** (2005) 1082 (cit. on p. 84).

[109] M. Yuan and M. Recker, *Not all rubrics are equal: A review of rubrics for evaluating the quality of open educational resources*, International Review of Research in Open and Distributed Learning **16** (2015) 16 (cit. on p. 85).

[110] World Health Organization, *Disability*, https://www.who.int/health-topics/disability#tab=tab_1, 2020 (cit. on p. 87).

[111] I. L. Organization, *Making the future of work inclusive of people with disabilities*, https://www.ilo.org/wcmsp5/groups/public/---ed_emp/---ifp_skills/documents/publication/wcms_729457.pdf, 2019 (cit. on p. 87).

[112] World Health Organization, *World report on disability. Chapter 8: Work and employment.* https://www.who.int/disabilities/world_report/2011/report/en/, 2011 (cit. on p. 87).

[113] M. Elias, S. Lohmann and S. Auer, *Ontology-Based Representation for Accessible OpenCourseWare Systems*, Information **9** (2018) 302, URL: https://www.mdpi.com/2078-2489/9/12/302/pdf (cit. on pp. 88, 90).

[114] R. Mace, G. Hardie and J. Place, *Accessible Environments: Toward Universal Design*, Van Nostrand Reinhold, New York, 1991 (cit. on p. 88).

[115] *W3C Web Content Accessibility Guidelines*, https://www.w3.org/TR/WCAG21/ (cit. on p. 88).

[116] W3C, *Cognitive and Learning Disabilities Accessibility Task Force*, https://www.w3.org/WAI/PF/cognitive-a11y-tf/, 2017 (cit. on p. 88).

[117] IMS Global Learning Consortium, *IMS Access For All*, https://www.imsglobal.org/activity/accessibility, 2012 (cit. on p. 88).

[118] I. Europe, *Easy-to-Read*, https://www.inclusion-europe.eu/easy-to-read/, 2019 (cit. on p. 88).

[119] M. Elias et al., "Accessibility and Personalization in OpenCourseWare - An Inclusive Development Approach", *Proceedings of the 20th IEEE International Conference on Advanced Learning Technologies (ICALT)*, IEEE, 2020 (cit. on p. 88).

[120] X. Zhang et al., *Accessibility within Open Educational Resources and Practices for Disabled Learners : A Systematic Literature Review*, **2** (2020) 1 (cit. on p. 88).

[121] G. George and A. M. Lal, *Review of ontology-based recommender systems in e-learning*, Computers & Education **142** (2019) 103642 (cit. on p. 88).

[122] K. Grammati-Eirini and R. Lopes, *Deliverable 4.1 - A set of formalisms and taxonomies for accessibility assessment procedures and their inherent meta models*, tech. rep., ACCESSIBLE (Grant Agreement No. 224145), 2009 (cit. on p. 88).

[123] C. G. Durán and C. M. Ramırez, *Integration of Open Educational Resources Using Semantic Platform*, IEEE Access (2021) (cit. on p. 92).

[124] D. Koutsomitropoulos, A. Andriopoulos and S. Likothanassis, *Semantic classification and indexing of open educational resources with word embeddings and ontologies*, Cybern Inf Technol **20** (2020) 95 (cit. on p. 92).

[125] F. Dang, J. Tang and S. Li, "Mooc-kg: A mooc knowledge graph for cross-platform online learning resources", *2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, IEEE, 2019 1 (cit. on p. 92).

[126] E. Ilkou and B. Signer, "A Technology-enhanced Smart Learning Environment based on the Combination of Knowledge Graphs and Learning Paths", *Proceedings of the 12th International Conference on Computer Supported Education, CSEDU 2020, Prague, Czech Republic, May 2-4, 2020, Volume 2*, SCITEPRESS, 2020 461 (cit. on p. 92).

[127] P. Chen, Y. Lu, V. W. Zheng, X. Chen and B. Yang, *KnowEdu: A System to Construct Knowledge Graph for Education*, IEEE Access **6** (2018) 31553 (cit. on p. 92).

[128] G. George and A. M. Lal, *Review of ontology-based recommender systems in e-learning*, Comput. Educ. **142** (2019) (cit. on pp. 93, 102).

[129] J. K. Tarus, Z. Niu and G. Mustafa, *Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning*, Artif. Intell. Rev. **50** (2018) 21 (cit. on pp. 93, 101).

[130] A. M. Salem and A. Y. Nikitaeva, "Knowledge Engineering Paradigms for Smart Education and Learning Systems", *42nd International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2019, Opatija, Croatia, May 20-24, 2019*, IEEE, 2019 1571 (cit. on p. 93).

[131] J. Barria-Pineda, K. Akhuseyinoglu and P. Brusilovsky, "Explaining need-based educational recommendations using interactive open learner models", *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, 2019 273 (cit. on p. 93).

[132] H. Davies, V. Lehdonvirta, A. Margaryan, J. Albert and L. Larke, *Developing and matching skills in the online platform economy: Findings on new forms of digital work and learning from Cedefop's CrowdLearn study*, (2020) (cit. on p. 93).

[133] Coursera, *Coursera | Build Skills with Online Courses from Top Institutions*, 2012, URL: https://www.coursera.org/ (visited on 08/02/2021) (cit. on p. 93).

[134] E. Katis, H. Kondylakis, G. Agathangelos and K. Vassilakis, "Developing an Ontology for Curriculum and Syllabus", *The Semantic Web: ESWC 2018 Satellite Events - ESWC 2018 Satellite Events, Heraklion, Crete, Greece, June 3-7, 2018, Revised Selected Papers*, vol. 11155, Lecture Notes in Computer Science, Springer, 2018 55 (cit. on p. 94).

[135] M. D. Wilkinson et al., *The FAIR Guiding Principles for scientific data management and stewardship*, Scientific data **3** (2016) 1 (cit. on p. 94).

[136] M. Gao, K. Liu and Z. Wu, *Personalisation in web computing and informatics: Theories, techniques, applications, and future research*, Inf. Syst. Frontiers **12** (2010) 607 (cit. on p. 97).

[137] T. Ivanova and M. Popov, "Ontology Evaluation and Multilingualism", *Proceedings of the 21st International Conference on Computer Systems and Technologies '20*, CompSysTech '20, Ruse, Bulgaria: Association for Computing Machinery, 2020 215, ISBN: 9781450377683 (cit. on p. 98).

[138] S. Chari et al., "Explanation Ontology: A Model of Explanations for User-Centered AI", *The Semantic Web - ISWC 2020 - 19th International Semantic Web Conference, Athens, Greece, November 2-6, 2020, Proceedings, Part II*, vol. 12507, Lecture Notes in Computer Science, Springer, 2020 228 (cit. on p. 98).

[139] N. Li, E. Motta and M. d'Aquin, *Ontology summarization: an analysis and an evaluation*, CEUR Workshop Proceedings **666** (2010) (cit. on p. 98).

[140] C. Brewster, H. Alani, S. Dasmahapatra and Y. Wilks, *Data Driven Ontology Evaluation*, (2004) (cit. on pp. 98, 99).

[141] A. Degbelo, "A Snapshot of Ontology Evaluation Criteria and Strategies", *Proceedings of the 13th International Conference on Semantic Systems, SEMANTICS 2017, Amsterdam, The Netherlands, September 11-14, 2017*, ACM, 2017 1 (cit. on p. 98).

[142] K. Stancin, P. Poscic and D. Jaksic, *Ontologies in education - state of the art*, Educ. Inf. Technol. **25** (2020) 5301 (cit. on p. 101).

[143] M. Harrathi, N. Touzani and R. Braham, "A Hybrid Knowlegde-Based Approach for Recommending Massive Learning Activities", *14th IEEE/ACS International Conference on Computer Systems and Applications, AICCSA 2017, Hammamet, Tunisia, October 30 - Nov. 3, 2017*, IEEE Computer Society, 2017 49 (cit. on pp. 102, 103).

[144] J. Jeevamol and V. Renumol, *An ontology-based hybrid e-learning content recommender system for alleviating the cold-start problem*, Education and Information Technologies (2021) 1 (cit. on pp. 102, 103).

[145] S. Chimalakonda and K. V. Nori, *An ontology based modeling framework for design of educational technologies*, Smart Learn. Environ. **7** (2020) 28 (cit. on p. 103).

[146] M. E. Ibrahim, Y. Yang, D. L. Ndzi, G. Yang and M. Al-Maliki, *Ontology-based personalized course recommendation framework*, IEEE Access **7** (2018) 5180 (cit. on p. 103).

[147] D. A. Koutsomitropoulos and G. D. Solomou, *A learning object ontology repository to support annotation and discovery of educational resources using semantic thesauri*, IFLA journal **44** (2018) 4 (cit. on p. 103).

[148] K. Skillen et al., *Ontological user modelling and semantic rule-based reasoning for personalisation of Help-On-Demand services in pervasive environments*, Future Gener. Comput. Syst. **34** (2014) 97 (cit. on p. 103).

[149] M. Zouri and A. Ferworn, "An Ontology-Based Approach for Curriculum Mapping in Higher Education", *11th IEEE Annual Computing and Communication Workshop and Conference, CCWC 2021, Las Vegas, NV, USA, January 27-30, 2021*, IEEE, 2021 141 (cit. on p. 103).

[150] C. I. Eke, A. A. Norman, L. Shuib and H. F. Nweke, *A Survey of User Profiling: State-of-the-Art, Challenges, and Solutions*, IEEE Access **7** (2019) 144907 (cit. on p. 102).

List of Figures

List of Tables

Mohammadreza Tavakoli
https://www.linkedin.com/in/mohamadreza-tavakoli-4b62b714b/

Basic Information

Place of Birth: Iran
Date of Birth: 03.Dec.1989

Mobile: +4915163043128
Address: Clausthaler Weg 43, Hannover 30419, Germany
E-mail: reza.tavakoli@tib.eu

Education

Sharif University of Technology, Tehran, Iran

M.Sc. Computer Engineering School

- September 2014 - September 2016
- Major in Computer Engineering, Minor in Software Engineering
- GPA: 18.24/$_{20.00}$
- M.Sc. Thesis Title: Improving Answers in the Stack Overflow Q&A Website according to the User Behaviors
- M.Sc. Thesis Description: The goal of the thesis is finding deficient questions in Stack Overflow according to the User Behaviors. Therefore, in this project a model has been created using Machine Learning methods in order to find low quality questions and try to improve them by assigning them to the experts of the questions fields.
- M.Sc. Thesis Supervisors: Dr. A. Heydarnoori
- Sharif University is the First-ranked Technical University in Iran

Shiraz University, Shiraz, Iran

B.Sc. Computer Science and Engineering Department

- September 2008 - September 2013
- Major in Computer Engineering, Minor in Software Engineering
- GPA: 16.00/$_{20.00}$
- B.Sc. Project Title: "Proposing an Intelligent Question Bank for students"
- B.Sc. Project Description: The goal of the project is providing an intelligence Question Bank in order to save questions with various level of hardness and try to create exam according to the selected difficulty level.
- Supervisor: Dr. M. Fakhrahmad

National Organization for Development of Exceptional Talents High School, Shiraz, Iran

High School

- September 2004 - September 2008
- Diploma in Mathematics and Physics
- GPA: 18.89/$_{20.00}$

Research Interests

- Learning Analytics & Open Education
- Machine Learning
- Software Development

Publications and Conference Presentations

- **An AI-based open recommender system for personalized labor market driven education.** Tavakoli, M., Faraji, A., Vrolijk, J., Molavi, M., Mol, S.T. and Kismihók, G., 2022. Advanced Engineering Informatics Journal.
- **Hybrid Human-AI Curriculum Development for Personalised Informal Learning Environments.** Tavakoli, M., Faraji, A., Molavi, M., Mol, S.T. and Kismihók, G., 2022. In LAK22: 12th International Learning Analytics and Knowledge Conference.
- **Metadata analysis of open educational resources.** Tavakoli, M., Elias, M., Kismihók, G. and Auer, S., 2021. In LAK21: 11th International Learning Analytics and Knowledge Conference.
- **Extracting Topics from Open Educational Resources.** Molavi, M., Tavakoli, M. and Kismihók, G., 2020. Extracting Topics from Open Educational Resources. In European Conference on Technology Enhanced Learning (EC-TEL).
- **Quality prediction of open educational resources a metadata-based approach.** Tavakoli, M., Elias, M., Kismihok, G. and Auer, S., 2020. In 2020 IEEE 20th international conference on advanced learning technologies (ICALT).

- **A recommender system for open educational videos based on skill requirements.** Tavakoli, M., Hakimov, S., Ewerth, R. and Kismihok, G., 2020. In 2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT).
- **OER recommendations to support career development.** Tavakoli, M., Faraji, A., Mol, S.T. and Kismihók, G., 2020. In 2020 IEEE Frontiers in Education Conference (FIE).
- **Labour market information driven, personalized, OER recommendation system for lifelong learners.** Tavakoli, M., Mol, S.T. and Kismihók, G., 2020. In the 12th Computer Supported Education Conference (CSEDU).
- **Improving the Quality of Posts in the Stack Overflow.** Tavakoli, MohammadReza, Abbas Heydarnoori, submitted to 27th IEEE International Conference on Software Analysis, Evolution and Reengineering. IEEE, 2019.
- **Requirement Analysis Towards Building a Personalized OER Recommender, based on Labour Market Information** Tavakoli, Mohammadreza, Gabor Kismihok, Stefan Mol. In the Proceedings of Learning and Student Analytics Conference (LSAC), LSAC, 2019.
- **Customer Segmentation and Strategy Development based on RFM analysis and Data Mining techniques: A case study** Tavakoli, MohammadReza, MohammadReza Molavi, Vahid Masoumi, Majid Mobini. In the Proceedings of International Conference on E-Business Engineering (ICEBE), IEEE, 2018.
- **Improving the quality of code snippets in stack overflow.** Tavakoli, MohammadReza, Abbas Heydarnoori, and Mohammad Ghafari. Proceedings of the 31st Annual ACM Symposium on Applied Computing. ACM, 2016.

Computer Skills

- **Competent**: Python, SQL, React Native, Git
- **Familiar**: Elasticserach, R, Java, VB, Scheme, React, RapidMiner, Kibana, Logstash, MongoDB, Docker, Maven, X86 & Arm Assembly
- **Type Settings**: Microsoft Office, LaTeX
- **Operating Systems**: Microsoft Windows, Linux (Ubuntu, MINT)

Awards & Honors

- **Iranian Nationwide University Entrance**
  - Ranked 1st (among around 20,000 participants) of Nationwide University Entrance for M.Sc. in Computer Engineering in fields of Software Engineering and Algorithms, 2013.
  - Ranked 24th (among around 15,000 participants) of Nationwide University Entrance for M.Sc. in Computer Engineering in field of Artificial Intelligence, 2013.
- **Scientific Contests**
  - 3rd Place in Alibaba challenge for Ticket Request Prediction (2019).
  - 1st Place in Fanavard Data Mining Challenge (the biggest Data Mining Competition in Iran). The competition was about Price Prediction (2018).
  - 3rd Place in Tap30 challenge for Online Taxi Demand Prediction (2018).

Working Experiences

- **Researcher at L3S/TIB**, Hannover, Germany 2019 - Present
  We are working on Open Educational Resources (OERs) recommendation systems in order to build an intelligent system which connect labour market and education
- **Project Manager at Kian Intelligent Asset Management**, Tehran, Iran 2018 - 2019
  Kian is a group of Iran-based regulated financial institutions offering investment banking, asset management and securities services. My responsibility is designing an Intelligent Asset Management product in order to recommend best choices for each user using Machine Learning Methods and Data Mining Techniques.
- **Product Owner (Leader of Technical Aspects) of Customer Segmentation Project and Customer Club Project at DigiKala Company**, Tehran, Iran 2017 - 2018
  Digikala is an Online Retailer and the Biggest E-commerce in the Middle East. I was the leader of the Customer Segmentation and Customer Club, two of the significant projects of Digikala. My team had six members and my responsibility was managing the technical team in order to achieve the objectives of the projects. Accordingly, my team and I tried to cluster the customers in order to be used by other team such as Marketing and Customer Services. Also, Customer Segmentation is used for appropriate recommendation to customers that is extremely important for E-Commerce and thus, we have implemented a widget in order to recommend products to the customers according to their history and also various E-Commerce strategies such as upselling, down selling, cross selling, deep selling. Another important point is that since the product catalog and user base of Digikala increase dramatically, we tried to use new products and new users in our process in order to handle cold start problem. Moreover, since our customer club has around

3 million users and around 20,000 *online* users, we have provided our services in a cloud base structure.

- **Data Scientist of R&D team at DigiKala Company**, Tehran, Iran 2016 - 2017
  I was a Data Scientist at the R&D Team of the Company. My Responsibility was working on Digikala's Recommender System in order to resolve its problems and improve the effectiveness of the recommendations.
- **Researcher and Developer in Pegah Data Miners of Sharif**, 2013 - 2016
  Backtory which is the first BAAS (Back end As A Service) in Iran has been developed in this Company, which was founded by some of graduated students of Sharif University of Technology. Backtory is a product such as AWS (Amazon Web Services) and OpenStack. My responsibility was developing cloud-code and authentication services. Moreover, we participated in developing Object-Storage product in Backtory
- **Research & Development team at Power Research Institute**, 2012 - 2013
  Power Research Institute is responsible for all research about power plants management in Iran in order to buy power from different region and distribute power between all users. I provided an appropriate visualization for the functionality and other properties of each power plants to help managers made decision based on the data.

Other Projects

- Design and Implementation of an Intelligent Question Bank. (Using C# 2011)
- Design and Implementation of a Website and Android Application for Statistical Center Of Iran. (Using HTML, CSS, PHP, and Java 2011)
- Design and Implementation of an application for Managing School Process such Grading, Visualizing the Grades, etc. (Using Visual Basic 2010)
- Design and Implementation of an Accounting application. (Using Visual Basic 2010)

Teaching Experiences

- Tutor of students who want to participate in Nationwide University Entrance M.Sc. of Computer Engineering (2013 - 2014)
  - Algorithm Design
  - Artificial Intelligence
  - Statistics
  - Database
  - Operating Systems
  - Programming Languages
  - Computer Architecture
- C# tutoring for undergraduate students (Shiraz University 2012)
- Tutor of High School Computers for Students Who Wants to Participate in Computer Olympiad (Noshad High Schools, Shiraz. 2011 - 2012)

Teaching Assistance Experiences

Software Evolution
- **2016: Dr. Abbas Heydarnoory, Sharif University of Technology**

Design Patterns
- **2016: Sobhan Forooghi, Sharif University of Technology**

Computer Architecture
- **2012: Dr. Farshad Khunjush, Shiraz University**

Digital Design
- **2011: Dr. Farshad Khunjush, Shiraz University**

Assembly Language
- **2010-2011: Dr. Sattar Hashemi, Shiraz University**

Related Courses

- **Artificial Intelligence 19.68/**$_{20.00}$
- **Operation Research 18.5/**$_{20.00}$
- **Social and Economic Networks 19/**$_{20.00}$
- **Software Evolution 17/**$_{20.00}$
- **Software Architecture 18.2/**$_{20.00}$
- **Ambient Intelligence 18.3/**$_{20.00}$
- **Design Patterns 17.4/**$_{20.00}$
- **Principles of Programming 17.4/**$_{20.00}$

Self Study

- **Machine Learning (Stanford University)**
- **Data Science Specialization (Johns Hopkins University)**
- **Machine Learning Specialization (University of Washington)**
- **Foundations of marketing analytics (ESSEC Business School)**
- **Data Warehousing for Business Intelligence Specialization (University of Colorado)**

Languages

- **Persian: Mother tongue**
- **English: Fluent, TOEFL (February 2018): 99/**$_{120}$

Extra Activities

- **Author in Setak Psychology and Sociology Magazine, Shiraz University. (2010 - 2011)**
- **Playing Football, Chess & Volleyball.**
- **Exploring Social & Cultural Issues.**