IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 7, 2021, accepted December 23, 2021, date of publication January 4, 2022, date of current version January 27, 2022.

Digital Object Identifier 10.1109/ACCESS.2021.3139537

Automatic Adaptation of Open Educational Resources: An Approach From a Multilevel Methodology Based on Students' Preferences, Educational Special Needs, Artificial Intelligence and Accessibility Metadata

PAOLA INGAVÉLEZ-GUERRA[1], (Member, IEEE), VLADIMIR E. ROBLES-BYKBAEV[1], (Senior Member, IEEE), ANGEL PÉREZ-MUÑOZ[1], JOSÉ HILERA-GONZÁLEZ[2], AND SALVADOR OTÓN-TORTOSA[2]

[1]GI-IATa, UNESCO Chair on Support Technologies for Educational Inclusion, Universidad Politécnica Salesiana, Cuenca 010102, Ecuador

[2]Department of Computer Science, University of Alcalá, 28805 Alcala, Spain

Corresponding author: Paola Ingavélez-Guerra (pcingavelez@ups.edu.ec)

This work was supported in part by the Erasmus+ Programmer of the European Union through the Project EduTech under Grant 609785-EPP-1-2019-1-ES-EPPKA2-CBHE-JP.

ABSTRACT The need for adaptive e-learning environments that respond to learning variability is now a fundamental requirement in education, as it helps to ensure that students learn and pass their courses within a set time frame. Although guidelines, techniques and methods have been established in recent years to contribute to the development of accessible and adaptable e-learning environments that promote digital inclusion, their implementation is challenging due to the lack of knowledge of an adequate way to do it and because it is considered more of a technological competence for scholars in the area. In this context, automated support for adapting material that responds to the correct use of accessibility metadata not only provides a way to improve the description of adapted educational resources, but also facilitates their search according to the needs and preferences of students, particularly those with disabilities. In this article, we carry out a multilevel methodological proposal for the automatic adaptation of open educational resources, in order to provide a tool that contributes to the accessibility and correct use of their metadata in e-learning environments. A research is conducted with students with disabilities to establish their real needs and preferences, highlighting the need to strengthen the adequate description and coherent alternative text in images, the correct subtitling in videos and the conversion of audio to text, data that are relevant to our proposal. The research conducted aims to contribute with an automated support tool in the generation of accessible educational resources that are correctly labeled for search and reuse. This research also aims to support researchers in artificial intelligence applications to address challenges and opportunities in the field of virtual education, in addition to providing an overview that could help those who generate educational resources and maintain their interest in making them accessible.

INDEX TERMS Adaptive systems, distance learning, accessibility, metadata, artificial intelligence.

I. INTRODUCTION

The development of technology and its application in education constitutes the constant study of changing and varied innovations. It is necessary to establish a follow-up that supports the whole process, both pedagogical and technological. Currently, virtual learning environments are the best way to offer a complex series of opportunities and tasks to educational institutions in the pursuit of the teaching-learning process, and even more so with the pandemic

The associate editor coordinating the review of this manuscript and approving it for publication was Amin Zehtabian.

and confinement situation the world is facing. The virtual environment is currently considered the most widely used tool in education, as it allows the distribution of digital educational resources (text, audio, videos, simulators, etc.) that facilitate communication, both in real time and according to the time availability of the user - student. The International Council for Open and Distance Education points out that a total of 414 million students will be enrolled in higher education worldwide in 2030 [1], so the constant study of interaction, learning and requirement for adaptation in e-learning is indispensable.

Accessibility and adaptability in digital educational resources, constantly demands the pursuit for updated research that responds to trends on the variability of student learning and their diversity. Students with disabilities are a valuable source of information in the adaptation requirements for the development of educational material that responds to universal design.

In the case of disability, it is important to consider the worldwide figures and their trend, since around 1 billion inhabitants, or 15% of the world's population, have some type of disability, and its incidence is higher in developing countries [2].

The 2030 agenda for sustainable development [3] determines the commitment to ensure equal access to all levels of education including persons with disabilities, considering inclusive and effective educational environments. 125 countries worldwide have signed and ratified the convention on the rights of persons with disabilities [4], so they face the challenge of providing quality education for all, making viable and strengthening the inclusion approach, addressing the high rates of exclusion, discrimination and educational inequality. The creation of conditions for the development of education for all, which guarantees access to information with equity, implies transformations and adaptations in educational resources, involving active and participatory evaluation processes that validate the efforts made.

The World Education Forum 2015 [5] states that "Information and communication technologies (ICT) must be harnessed to strengthen education systems, knowledge dissemination, access to information, effective and quality learning, and more efficient delivery of services".

It is necessary the generation of tools that support the automatic adaptation of digital educational resources, according to the needs and preferences of students. The correct adoption of standards that consider accessibility, not only strengthens quality characteristics, but also provides communication and timely search of desired educational resources, generating satisfaction and fidelity.

This paper presents the analysis of a tool created for adaptation using artificial intelligence techniques and is organized as follows: Section II presents the background and related work so far. Section III provides details of the proposed architecture for automatic adaptation of open educational resources. Section IV presents the analysis of results. Section V the limitations encountered, and Section VI concludes with the discussion of findings and recommendations.

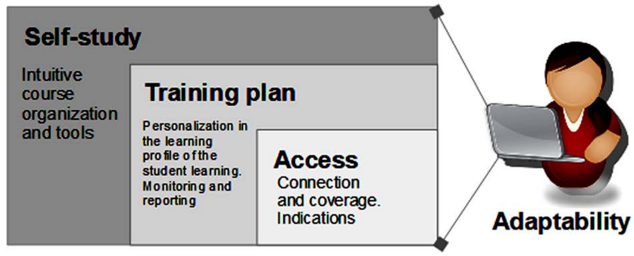


FIGURE 1. Levels of adaptability.

II. BACKGROUND

The existing relationship between the different digital educational resources that make up a virtual environment and their interaction with the user, demands the establishment of characteristics that allow to analyze the accessibility and adaptability in each of them.

A. ACCESSIBILITY AND ADAPTABILITY

Adaptability and accessibility are two terms that converge when it comes to meeting the diversity of human beings (adaptability), seeking to provide flexibility in their environment (accessibility), so as to accommodate the needs of each user and their preferences.

The ISO/IEC 24751-2:2008 standard, titled "Information technology – Individualized adaptability and accessibility in e-learning, education and training", defines accessibility as "Usability of a product, service, environment or facility by individuals with the widest possible range of abilities", and adaptability as "the ability of a digital resource or delivery system to adjust the presentation, control methods, structure, access mode and user support in its presentation" [6].

The measurement of adaptability in e-learning, according to [7], can be done with indicators defined for three levels, where levels 2 (Training plan) and 3 (adaptability to self-learning) require greater emphasis on diagnostic evaluation, and continues to seek greater effectiveness and efficiency even in the post-training process. (see Figure 1).

Reference [8] considers that accessibility is not only framed in technology and its interaction, it also requires feedback from the design of learning experiences for all, considering not only technology and pedagogy, but also ethics.

B. OER, LO AND METADATA

In 1994 Hodgins defined the concept of learning object and received acceptance for the premise of ease of reuse [9]. Technological progress and the use of digital resources in the mediation of learning, makes its concept evolve constantly. The definition of [10] as "…digital entity, self-contained and reusable, with a clear educational purpose, constituted by at least three editable internal components…", and the constant coincidence in the characteristics of identification, recovery,

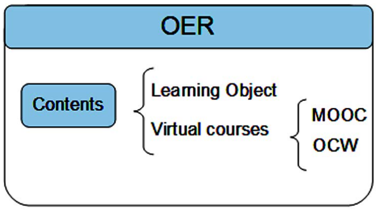


FIGURE 2. Relationship between OER, LO and metadata.

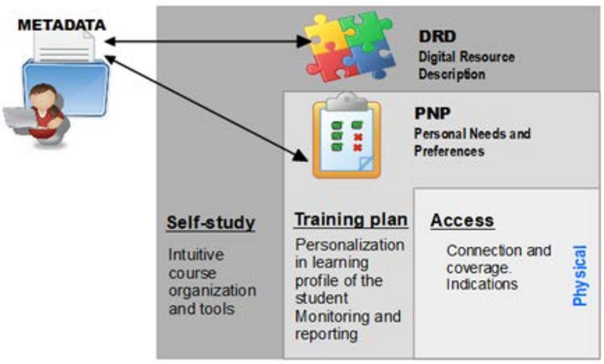


FIGURE 3. Accessibility metadata according to AfA.

detectability, reusability and interoperability; allows delimiting it but at the same time understanding the variability and cultural evolution of its practice in virtual learning environments. It is in this path of evolution that legal aspects and reuse licenses are established, which gives rise to OERs (Open Educational Resource). The term "open" involves an active participation in five activities determined by [11] as the 5Rs: retain, reuse, revise, remix and redistribute.

Based on the contributions of [12], [13] the existing relationships between OER, LO and virtual courses are expressed in Figure 2.

For the description of the accessibility characteristics of the contents published in learning objects, it is necessary to use information description mechanisms based on metadata, which would facilitate the information of a digital resource and its possible requirement based on preferences and needs of the student.

Reference [14] is considered as a reference in metadata frequently used. The accessibility metadata defined by Schema.org are based on IMS AfA v3.0 [15], which responds to the ISO/IEC 24751 standard on individualized adaptability and accessibility in e-learning, education and training [6]. In the case of students with disabilities, the information of a resource is relevant because it facilitates their interaction. Accessibility metadata allows describing the accessibility characteristics of the resource (DRD) [16], as well as the user's preferences and needs (PNP) [17], as shown in Figure 3:

C. ARTIFICIAL INTELLIGENCE: MULTILEVEL

By virtue of the different user requirements, the process of adapting the learning object must be flexible enough, in order to meet the user's diverse requirements. Therefore, our proposal is supported by previous work described in [18]–[20], where alternatives focused on content adaptation at different levels of granularity using techniques based on uncertain reasoning, reuse of learning fragments/objects to create new ones, and multilevel clustering techniques are proposed.

However, our model employs an approach based on artificial intelligence techniques that allow combining different levels of granularity in order to meet users' adaptation requirements. As shown in Figure 4, fundamentally auditory, textual and visual information needs to be adapted according to the users' needs. For this purpose, we propose a level of abstraction of adaptation of each information element, using different artificial intelligence techniques described in the following sections. For example, a user may require that a video that does not have subtitles has them and that a summary can also be presented using texts adapted to easy reading. To do this, the system must first perform automatic speech recognition in order to extract the corresponding textual information. Then the operations and results are combined at two different levels of abstraction, the first one converting the plain text to a text that is easy to read, and from that text the required summary is obtained.

Similarly, for the image-description process, the system extracts all the images as well as the near texts. In this line, the images are analyzed considering four levels that rely on nine CNN (Convolutional Neural Network) for image classification, two CNN + RNN (Recurrent Neural Network) for photos and equations description, and one LSTM NN (Long Short-Term Memory Neural Network) from Tesseract OCR library for text recognition. For the CNN + RNN the system does not need the texts that are near to the images, whereas for the CNN we have implemented an NLP (Natural Language Processing) process that helps describe the images using the near texts.

With this, the user can have as many combinations as he wants, and much more flexible adaptation processes are feasible.

D. RELATED WORK

Some projects have been developed to favor accessibility and adaptability in virtual environments. The shared experiences EU4ALL [9], ESVIAL [21], TILE, AEGIS, ACCESSIBLE [22] OBBA in Brazil [23], to mention a few, point to research and implementation efforts to favor educational inclusion. In parallel, the evaluation of digital learning resources, generates proposals for models and standards to be applied, for which, the accessibility criterion is considered relevant but still does not achieve a consensus of information. Standards such as ISO 9241-11 [6], ISO/IEC 19796-3 [24], ISO/IEC 24751-3 [16], establish guidelines that are related to accessibility; however, the applicability and diffusion is still limited. Methodological proposals focused on the quality of virtual educational resources are based on ISO standards, establishing guidelines for applying ICTs in teaching. Authors such as [21], [25] identify the lack of

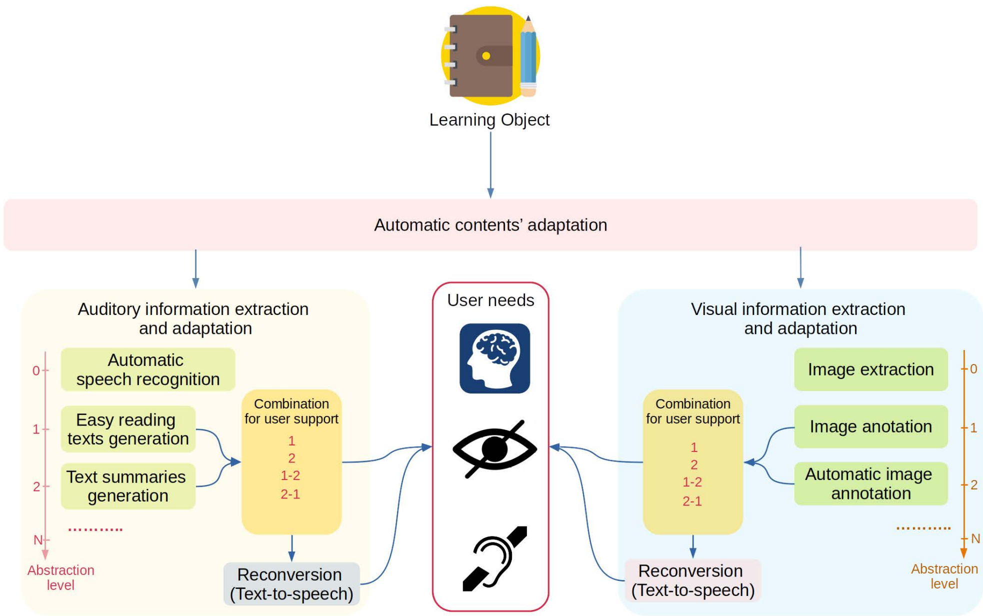


FIGURE 4. Multilevel approach to adaptation of learning objects considering elements of artificial intelligence.

an accessibility methodology with a holistic and adaptive approach.

The collaborative approach from which OERs are born leads to the joint pursuit of pedagogical and technological challenges to achieve an improved reconstruction with quality. [26]–[28]. It is necessary to evidence defined metrics that endorse methodologies [21] and reference international guidelines or instructions related to design for all [29].

The incorporation of intelligent systems could contribute in the evaluation of accessible resources and in the feedback of profiles and personalization from the user experience [30], [31].

The efficient publication of accessibility information would facilitate optimal navigation and search of resources according to student needs and preferences [32], [33].

III. STATISTICAL ANALYSIS OF THE NEEDS OF USERS WITH DISABILITIES

As a previous step to the development of the proposal and the different tools described in this article, it was considered essential to know the needs of real users with different types of disabilities. For this purpose, a pilot study was conducted with 47 volunteers (20 women and 27 men), aged between 17 and 63 years (mean = 27.3, SD = 11.18) who interacted with virtual platforms in both undergraduate and graduate training environments. In order to gain insight into the perceptions of users' needs and requirements, a survey was developed and organized into two sections: one to collect demographic and disability data (9 questions) and another to determine the technological tools they use and the preferences/needs and difficulties they face with respect to the accessibility of virtual educational environments and their resources (36 questions).

TABLE 1. The number of volunteers grouped by gender and disability.

Volunteers(grouped by gender)			Disability
Female	Male	Total	
2	4	6	Deafness (1 woman and 2 men have around 50% of hearing loss)
2	4	6	Deafness but has cochlear implant/audiophones (can hear)
5	7	12	Physical disability
4	4	8	Intellectual/developmental disability
3	1	4	Mental health problems or psycho-social disorders (eg depression, bipolar disorder, schizophrenia, etc.)
3	5	8	Visual impairment (mild, moderate, severe, blindness)
1	2	3	Other disabilities (cerebral palsy or acquired brain damage, diseases of organ systems, etc.)
	Total	47	

After the explanation process provided to the participants, the survey was conducted with the support and guidance of a group of experts in the area of educational inclusion, health, labor inclusion and computer science, in case they had difficulties in understanding the questions or could not interact with the virtual platform. The survey was validated with Cronbach's Alpha test, obtaining a value of 0.94. Table 1 describes the types of disability that these people have, grouped according to their gender. As can be seen, the group of volunteers was intended to represent cases of people who commonly access or have had previous experience in the management of online learning environments or in the use of virtual tools.

In this regard, 5 of the women who participated in this initial study had postgraduate studies, 2 had a bachelor's

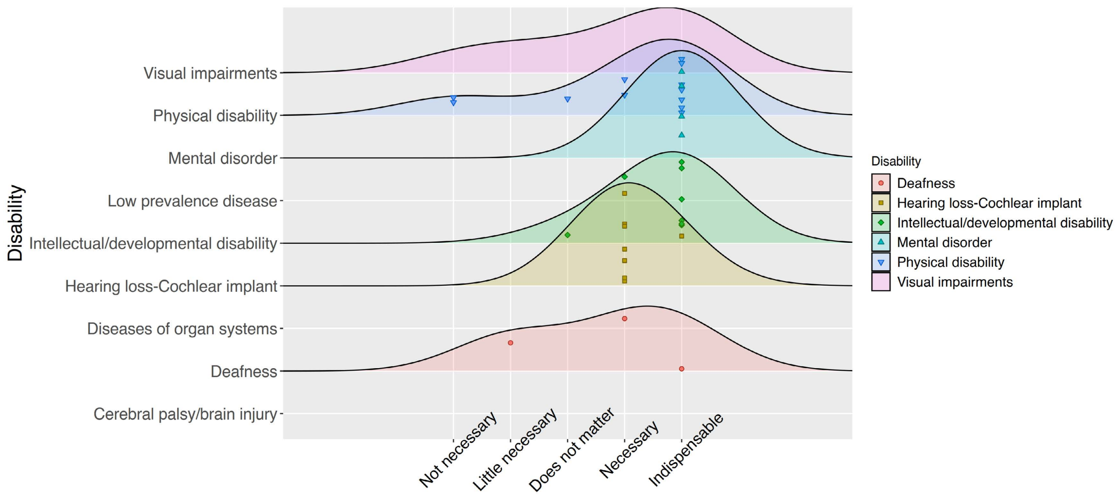


FIGURE 5. **Perceptions of surveyed persons with disabilities regarding the criterion "the educational resources must use mostly graphical contents."**

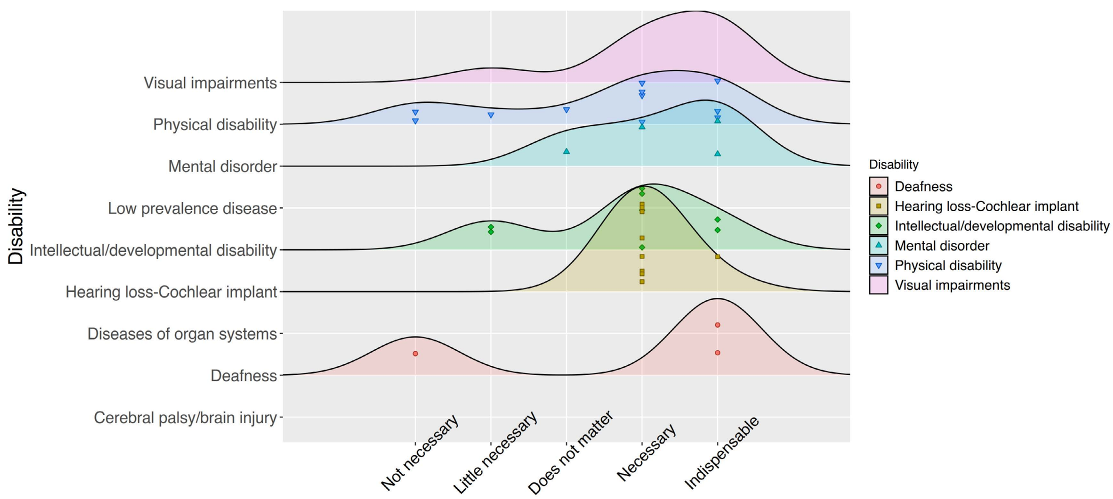


FIGURE 6. **Perceptions of surveyed persons with disabilities regarding the criterion "the educational resources must use mostly textual contents."**

degree and 13 had a university degree. Among the men, 2 had postgraduate studies, 20 had a university degree and 5 had a bachelor's degree. Regarding employment status, 19 men are not working and 8 are employed in various positions: teaching (3), administrative in a company (2), independent business (1) and operational (2). In the case of women, 11 are unemployed and 9 have the following jobs: teaching (2), independent business (4), management position (1), administrative position (1) and operational position (1).

Figure 5 shows the perception of the volunteers with respect to the possibility of using mostly graphic content in the educational resources uploaded in virtual learning environments. As can be seen, people with visual impairment, deafness or physical disability have different criteria, some consider it necessary, while others do not. However, for people with mental health problems and intellectual/development disabilities, they consider it ''indispensable'' and ''necessary'' to have graphic content.

On the other hand, Figure 6 shows the opinion of the volunteers regarding the criterion that learning objects deployed in virtual educational environments should use mostly textual content. In this aspect, it is observed that most deaf people believe that it is ''indispensable'' that the contents are mainly in textual format. In the case of people with hearing loss who use cochlear implants or hearing aids, they consider it ''necessary''. For people with intellectual/developmental disabilities the criterion varies from ''little indispensable'' to ''indispensable'', with no clear trend. Similarly, for survey participants with health problems, visual impairments and physical disabilities the criteria are scattered, with no clear trend.

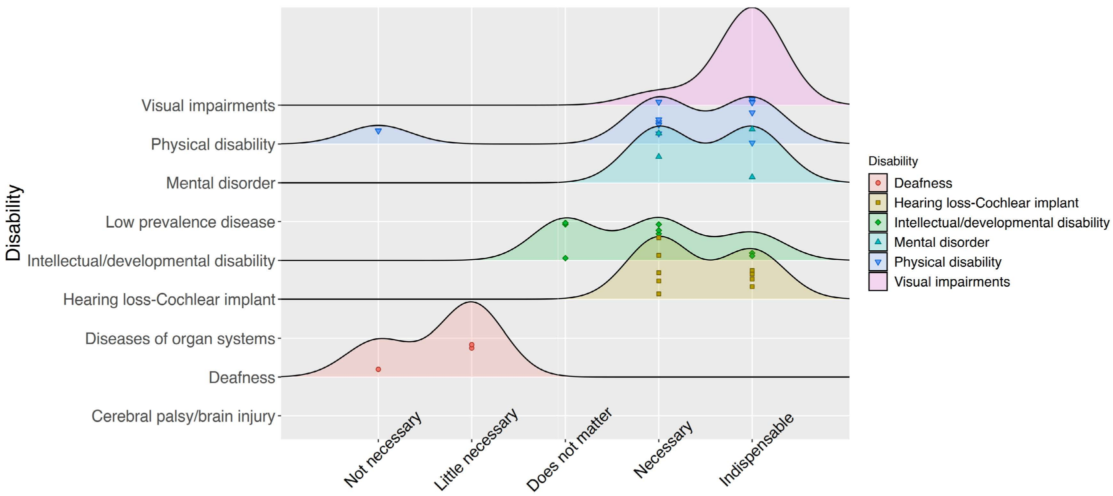


FIGURE 7. Perceptions of surveyed persons with disabilities regarding the criterion "the educational resources must use mostly auditory contents."

Finally, Figure 7 shows that deaf people consider that it is not necessary and very little necessary for learning objects deployed in virtual educational environments to use mostly auditory content. This is very coherent, since this type of resources cannot be used by them at all while using the learning object. In the case of people with hearing loss who use cochlear implants or hearing aids, they consider the criterion to be divided between "necessary" and "indispensable", since they can access these contents. As for people who are blind, the tendency is completely aligned with the "indispensable" option, since auditory content is the main resource they have to access the content of a learning object. For people with intellectual/development disabilities the criteria vary between "doesn't matter", "little indispensable" and "indispensable", with no clear trend.

Similarly, for survey participants with mental disorder and physical disability the criteria are scattered, with no clear trend.

IV. METHODOLOGY

A. SYSTEM ARCHITECTURE

The architecture proposal requires the loading of the learning object, which is usually packaged in SCORM, IMS, Common Cartridge educational format, being its common structure a compressed ZIP file. This is followed by the unpacking of the content and the extraction of its respective tags. Figure 8 shows the general architecture diagram.

1) MULTIMEDIA EDUCATIONAL RESOURCES

The multimedia educational resources layer contains the information extracted from the learning object in terms of textual content, video content, audio content and images. Using artificial intelligence techniques, the content is adapted in its specific modules for learners with different disabilities.

Audio content analysis module: automatic speech recognition for SubRip Subtitle file generation As seen in the statistical analysis conducted with 47 people with various types of disabilities, the audio aspect is a key feature for people with varying degrees of hearing loss. Therefore, it is essential that learning objects containing videos have subtitles in order to meet the needs of this group of people. Therefore, this section describes the module that allows extracting an audio track from a video contained in a learning object and from it obtaining a text file with the audio transcription. Likewise, based on this text file and the audio of the video, the module can also generate a SubRip Subtitle file (SRT).

Figure 9 shows the process carried out by the module for two possible scenarios: a) generation of an SRT file in case the video does not have subtitles and b) comparison of subtitle files created by external tools and SRT files generated by the module. The most important aspects of the stages carried out by the audio analysis module are detailed below.

In the first scenario, the module extracts the audio track from the video using a sampling frequency of 44.1 KHz and stores it in a Waveform Audio File Format (WAV) file (step 1). For this purpose, the open source, cross-platform tool FFmpeg1 is used. The WAV file is then analyzed through Automatic Speech Recognition (ASR) in order to extract the different sentences that make up the explanation given by a narrator or persons intervening in the video (step 3). This process is performed offline using the open source and multiplatform tool VOSK that works on top of the KALDI base tool. While the text file is extracted, the time lapses in which the sentences occur are detected and the corresponding timestamp is added, which makes it possible to create the SRT file itself (step 4). With these three steps the SRT file is incorporated to the video and with this it is possible to have subtitling in case it is required by the users of the system.

As for the second scenario, four more steps are executed than in the previous scenario. In this case, it is important

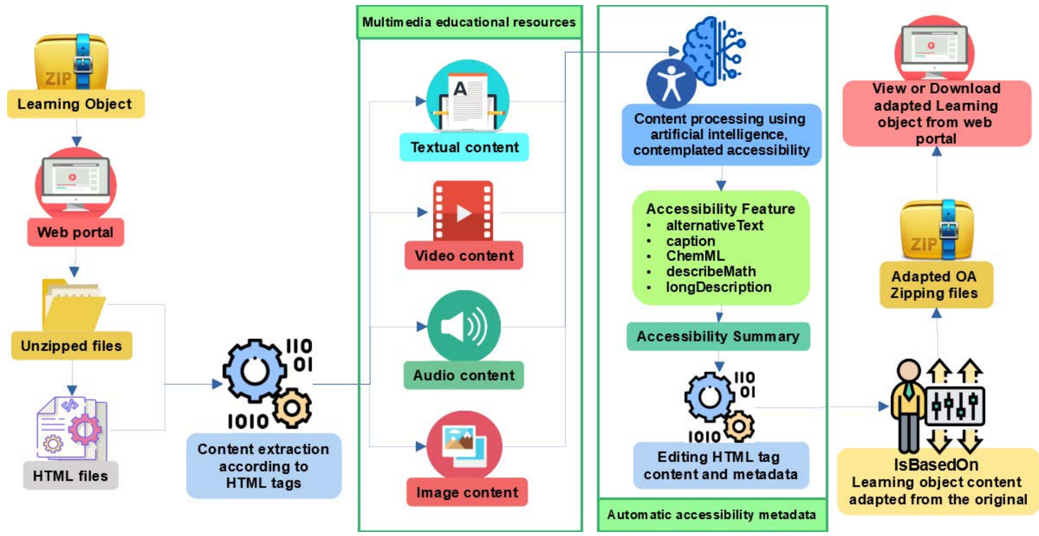


FIGURE 8. The general system architecture organized into modules and layers.

to point out that we start from a video file that already has its own subtitling. Therefore, the first thing to do is to extract the SRT file already included in the video (step 2). Once both SRT files are available (both the one generated with external tools and the one generated by the audio analysis module), the sentences are extracted (removing the timestamps) (step 5). The words of these sentences are converted into word embeddings (for each SRT file) (step 6). To carry out this task we use the neural network that comes pre-trained in the open source tool SpaCY, which is a popular Python library that contains the linguistic data and algorithms needed to process natural language texts [34]. These vector representations are then compared with cosine similarity metrics (Eq. 1) (step 7):

$$sim(SRT, CSTR) \equiv cos(vec(SRT), vec(CSRT)) = \frac{vec(SRT) \cdot vec(CSRT)}{|vec(SRT)| \cdot |vec(CSRT)|} \quad (1)$$

where:

- SRT represents the text (without timestamps) that has been generated by external tools.
- CSRT represents the text (without timestamps) produced by the audio analysis module.
- The *vec*() function is the one that allows to obtain the word embedding from a given sentence.

Figure 9 shows the different steps performed by the audio analysis module considering two scenarios: generation of SRT files for videos without subtitles, and comparison of the quality of subtitles generated versus those generated by external tools. With this, we can determine how effective is the result generated by the audio analysis module, as long as we start from a given gold standard, i.e., having an annotated corpus where we know exactly which words and sentences should be recognized by the tool. In our case, we used YouTube's automatic subtitling tool, as it is the one commonly used for the generation of educational content of various kinds. However, we should not lose sight of the fact that no tool is able to perform subtitling with 100% accuracy due to various aspects (noise, incorrect pronunciation of the speaker, idioms/localisms, etc.).

Image and textual content: In the extraction of the images present in the object, the <figcaption>, <img> tags and the information of the TAGs closest to the image are analyzed from the different HTML files: <h>, <p>, etc. The identified images go through different neural networks to generate information according to their content. The processing ends when editing the HTML file by adding information to the alternative text, which will support a correct labeling later. See Figure 10.

The proposal presented with its different multilevel deep learning networks, represent an adequate identification of description in 49% of images found with their different types or areas. The addition of several layers of deep learning considerably improves the classification of images, due to the number of images needed for training in specific classes and their characteristics.

For image classification, convolutional networks are used by Transfer learning described in [35]. For each category, it trained with more than 4k images. To generate a broad or complementary description of the classified images, natural language processing is used, improving the description and validating it semantically.

In order to improve the description of the images, natural language processing is used, which uses the information obtained from the <div>, <p>, <spam> or other tags with textual information closer to the image until reaching a <h1>-<h6> tag or another image. This description must comply with a semantic analysis that validates it. In case of being validated, this description is complemented to the result of the classification of the multilevel neural networks

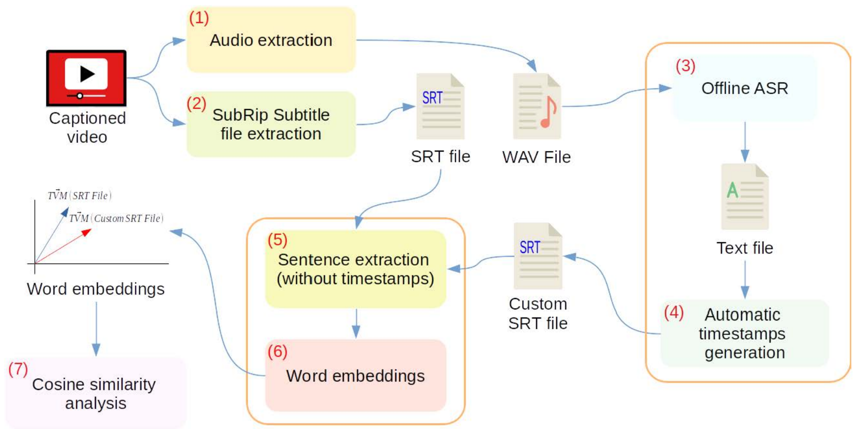


FIGURE 9. Audio analysis module.

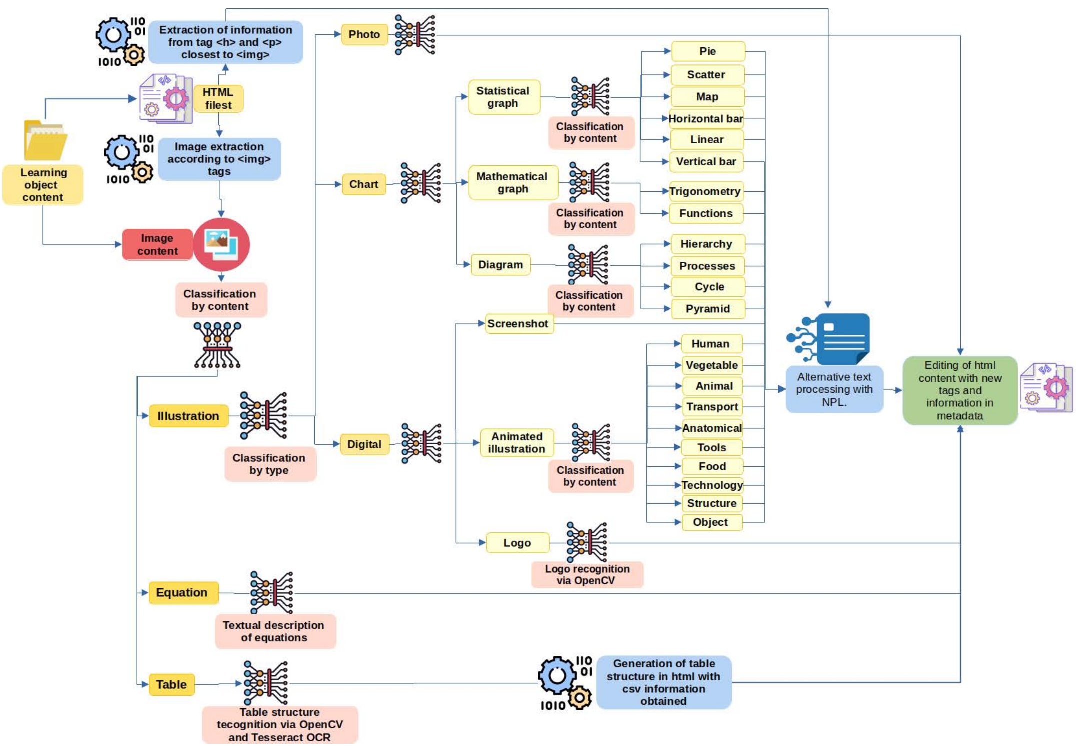


FIGURE 10. Multilevel deep learning proposed for classification and description of images.

processing (process in development). In case of not having a description of the semantic analysis verified by NLP, the closest title or subtitle prior to the image is used, complemented with the result of the classification of the multilevel Deep learning networks processing.

In the case of photographs, the neural network architecture is based on [36] and equations based on [37]. This architecture is based on the combination of CNN-RNN. CNNs preserves the spatial information and RNNs handles the sequential data. In these cases, use was made of LSTMs which are a modified version of recurrent neural networks. In the description of the photographs, the MS COCO dataset database was used, which has 180k training images with their respective annotations; as an addition to this database, 2k images with their description referring to educational topics were included. In the case of equations description, we made use of a database of 400k images with their annotations, this database was generated according to [37], as an aggregate we expanded characters and equations.

To generate a table in HTML sentences and tags, text recognition with Tesseract OCR engine is employed.

For logo recognition, the database of logos of brands from all over the world [38] are used because each logo has its own name and in most databases these are represented by the

```
<meta itemprop="accessMode" content="auditory">
<meta itemprop="accessMode" content="textual">
<meta itemprop="accessMode" content="visual">
<meta itemprop="accessibilityFeature" content="caption">
<meta itemprop="accessibilityFeature" content="alternativeText">
<meta itemprop="accessibilityFeature" content="synchronizedAudioText">
<meta itemprop="accessibilityFeature" content="describedMath">
<meta itemprop="accessibilityFeature" content="MathML">
<meta itemprop="accessibilityFeature" content="ChemML">
<meta itemprop="accessibilityFeature" content="audioDescription">
<meta itemprop="accessibilityFeature" content="transcript">
<meta itemprop="accessibilityFeature" content="ttsMarkup">
```

FIGURE 11. Automatic accessibility metadata.

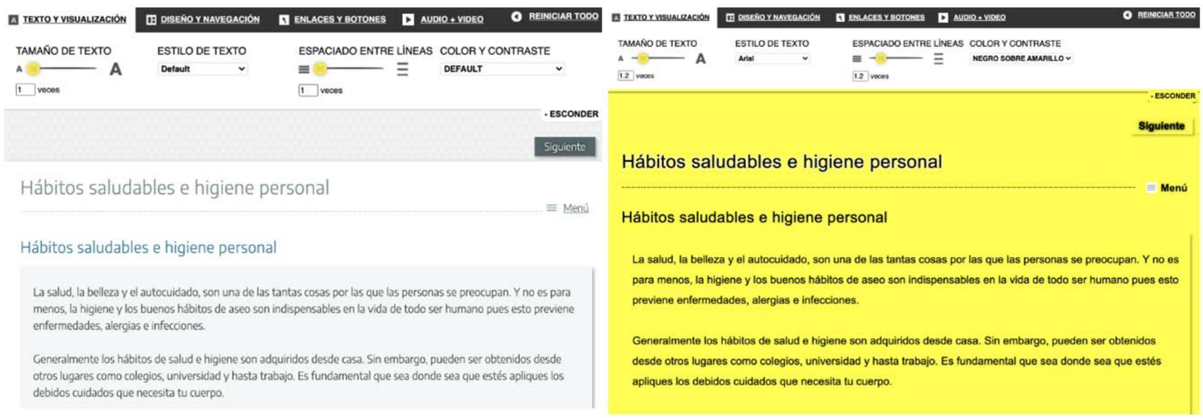


FIGURE 12. Accessible interface configuration. The screen capture of the left side shows the options to modify the user's preferences, whereas the image on the right side shows the result of applying high contrast on the learning object.

name of the company in general. The database has the largest collection of logos worldwide exceeding 300K.

2) MULTIMEDIA EDUCATIONAL RESOURCES

With the automatic identification and adaptation of the learning object, the appropriate labeling is generated considering the Schema.org guidelines and accessibility metadata as shown in Figure 11.

Additionally, the customization of the learning object interface is considered using the framework of the Fluid Infusion project [39] that combines JavaScript, CSS, HTML and user-centered design to support inclusive design on the web. This framework integrates at the top of the web page (Figure 12) an interface configuration palette with which the following can be adapted according to the user's needs or preferences:

- Text and display: Text size, text style, line spacing, color and contrast.
- Layout and navigation: Display the table of contents.
- Links and buttons: Emphasize links and magnify entries.
- Audio and videos: Display subtitles whenever possible, display transcripts whenever possible, default volume and default language.

With this it is possible to identify the automatic incorporation of the following accessibility metadata:

- accessibilityFeature: displayTransformability/ background-color
- accessibilityFeature: displayTransformability/ font-family
- accessibilityFeature: displayTransformability/font-size
- accessibilityFeature: displayTransformability/color
- accessibilityFeature: displayTransformability/ word-spacing
- accessibilityFeature: displayTransformability/ line-height
- accessibilityFeature: captions
- accessibilityFeature: synchronizedAudioText
- accessibilityFeature: highContrastDisplay
- accessibilityFeature:transcript
- accessibilityFeature: structuralNavigation
- accessibilityFeature:readingOrder
- accessibilityFeature: tableOfContents
- accessibilityFeature: index
- accessibilityFeature: audioDescription

V. EXPERIMENTATION AND RESULTS

In order to determine the accuracy of the audio analysis module, a process was carried out in which hundreds of videos were collected from the YouTube platform and 60 videos were selected from three different categories: chemistry, programming and mathematics. The videos were

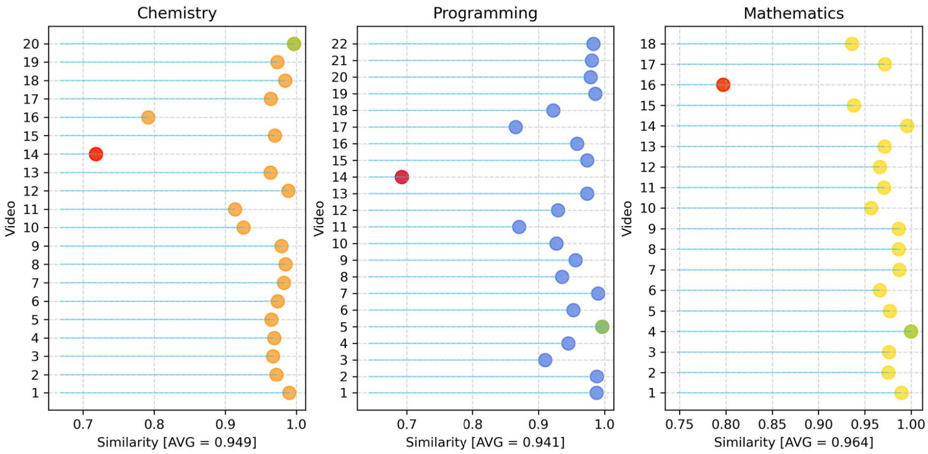


FIGURE 13. Similarity results obtained between the subtitles generated by external tools (YouTube®) and the audio analysis module.

reviewed manually in order to determine if the subtitling corresponded to the audio of the video.

Once this aspect was verified, we proceeded to use the audio analysis module based on the second scenario, i.e., comparing the similarity between the text already in the video and the text generated by the module.

As can be seen in Figure 13, the results are very positive, since on average the values for the three categories are 0.949, 0.941 and 0.964, respectively. The subtitles with lower similarity exceeded the value of 0.7. With this, we can determine that the quality of the texts generated by the audio module are comparable to those that can be achieved with external tools.

In relation to the description of images in learning objects, an analysis was made on the integration of the description or captioning of images in 15 learning objects with a total of 192 images. It was found that most of the images do not have any description in any of the forms normally used, identifying that the description is in the tag <alt>, <figcaption> or in any tag either <div>,<p>, <span> or other tag in which it presents a prefix (Figure, Fig, Image, Graphic). Table 2 shows the percentage of the use of the tags for the description of the images in its different ways, taking into account that the "Yes" represents a correct description, which provides relevant information for the understanding, in case it is not complete or relevant a "No" has been imposed.

Considering that a good practice for image description is to add it in <figcaption> and <alt> tags, the proposed tool to find a description in other tags or TAGs removes them and generates the description in <figcaption> and <alt> so that it can be read correctly by a screen reader.

Based on this analysis, in order to generate a description for 49% of the images, image processing using Artificial Intelligence is used, whose main objective is to generate information related to the image so that a user can know the content of the image, which may not be displayed in the browser or the user cannot perceive it because he/she has special education needs related to a visual disability.

TABLE 2. Results of the analysis carried out in 192 images from 15 learning objects. Most of images do not have any description such as <alt>, <figcaption> or other tags.

<figcaption>	other tags	<alt>	Images	Percentage (%)
No	No	No	94	48.96
No	No	Yes	14	7.29
No	Yes	No	36	18.75
No	Yes	Yes	26	13.54
Yes	No	No	0	0
Yes	No	Yes	22	11.46
Yes	Yes	No	0	0
Yes	Yes	Yes	0	0
		Total	**192**	**100**

This generated information will be present in the <figcaption> and <alt> tags, generating the alternativeText metadata or in specific cases the image will be converted into HTML content that can be interpreted in a browser and can be read by a screen reader.

The 94 images that do not have any description were classified. In Figure 14 we can see the results obtained from the classification using the ROC (Receiver Operating Characteristic) curve [40], which is a graphical representation of the ratio or proportion of true positives (TPR = True Positive Ratio) versus the ratio or proportion of false positives (FPR = False Positive Ratio), also according to the variation of the discrimination threshold (value at which we decide that a case is a positive), with an area under the curve of 0.65 in its lowest classification and an area of 0.98 in the highest classification.

In turn in Figure 15 we can see the result of the generation of the description of the subsequent images of the different classifications.

A. RECOGNITION OF INFORMATION PRESENT IN IMAGES OF TABLES

Only four images that represent tables were properly detected in all learning objects analyzed. In the same line, the HTML

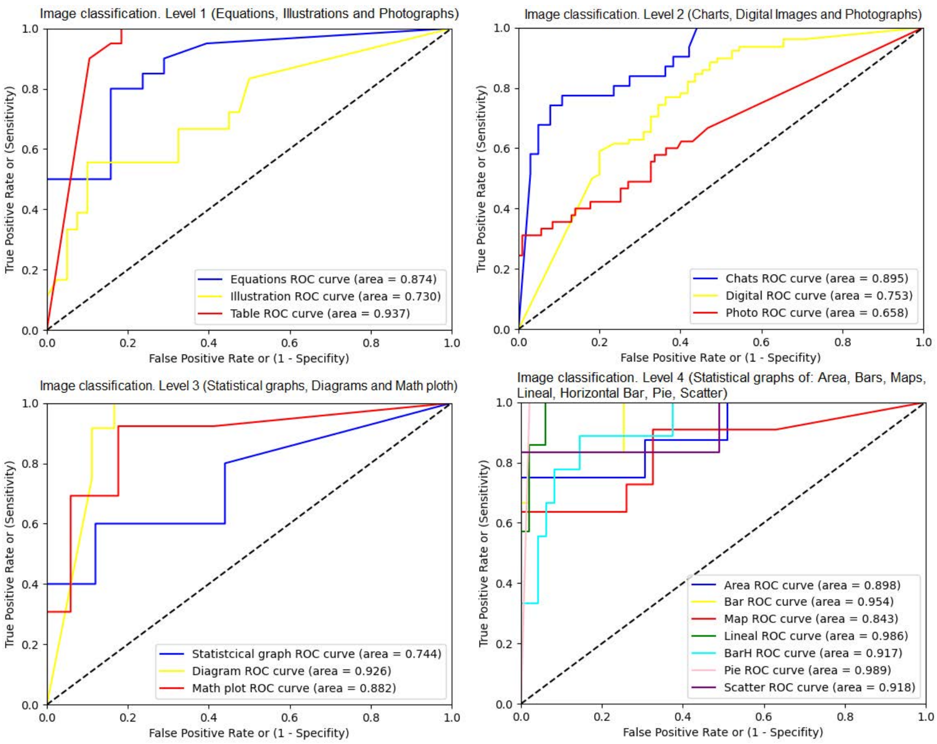


FIGURE 14. ROC curves for image classification according to the four levels defined.

code was successfully generated for these tables. These tables had a basic style without colors and defined cell lines.

To determine if the tool works properly, one of the tables was tested with different formats. As shown in Figure 15, a table that does not have a correctly defined grid, such as the external frames and the title and the color of the letters is soft, gives results with errors. In this case only the table was recognized and assembled in HTML with the cells that can be distinguished. In the following case, a table with a white grid was tested, which caused the cells not to be recognized. At the same time, the white text and the dark outline did not allow any text in the image to be recognized.

B. LOGO RECOGNITION

In the learning objects analyzed, 15 logos were found, most of which (12/15) were correctly recognized, with particular cases of regional or local logos that are not in the database. See Figure 15.

VI. LIMITATIONS

This research presented limitations during the process and in its search to answer the interaction questions from the experience of students with disabilities. The selection of the sample is limited. However, by obtaining a coefficient of 0.94 in Crombach's alpha, an optimal level of reliability is determined, thus identifying a reduction of bias. The automation of images responds to a diversity of categories with databases that are out of context, considering also that many images respond to mixed contents with texts in different languages and in many cases, the quality is not adequate for an optimal comparison; however, the bias was reduced by choosing a set of databases that cover the main disciplinary fields.

Another limitation is the selection of audios, but even so, it is considered that the evaluation process offers us a good overview of the scalable feeding of auditory information, identifying several relationships.

VII. DISCUSSIONS AND RECOMMENDATIONS

The objective of this work was to determine the barriers frequently detected in accessibility, from the own experience of students with disabilities in virtual environments. The results showed a lack of implementation of accessibility regulations in educational resources and learning objects, with special emphasis on images, video and their corresponding audio. While standards have been planned throughout history, the guidelines for their use are still subject to subjective criteria that depend on the design and implementation of a digital educational resource and student feedback. Adequate processes to meet reasonable accessibility of images and audios require automated tools that support the teacher to generate greater impact on applicability. Ensuring that a product complies with accessibility features strengthens the identification of metadata according to the preferences and needs of students with disabilities, as well as the satisfactory or not results of their teaching-learning process. Case reports

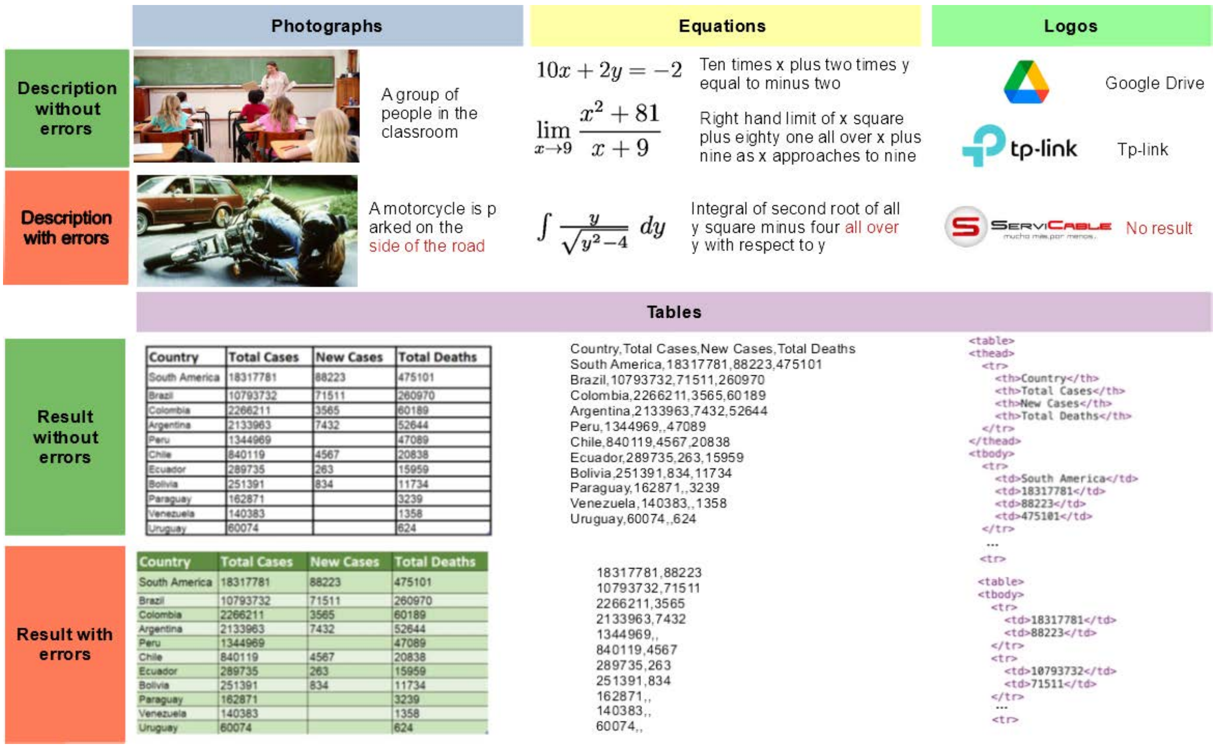


FIGURE 15. Results reached by the system with the automatic description process for images, photos, tables, and equations from learning objects.

are established with a limited number of subjects. The studies focus essentially on local experiences and rarely evaluate the positive effects of accessibility in interaction with learning objects in virtual environments. As a result of this research, we propose the development of an automatic adaptation tool to strengthen the accessibility and adaptability of OER considering standards and metadata.

A. RECOMMENDATION AND FUTURE DIRECTIONS

Through the findings and research perspectives on the various proposed solutions to improve the field of accessibility and adaptability in OER, it is relevant to explore the efforts generated to adapt resources and the effort required by each teacher, whether or not they have computer and accessibility knowledge. Although [9], [41], [42] apply numerous practices to incorporate accessibility, it is still complex to respond to automatic tools that consider accessibility guidelines and metadata established for the effect.

The advantages of a proper implementation of metadata publication is still not a general domain knowledge, as concluded by several authors [25], [43]. The effective applicability of accessibility metadata, could trigger a breakthrough in relation to the problem of finding accessible educational resources, effectively validated, and that respond to the variability in a student's learning, given their needs and preferences, considering that the efforts generated to create accessible OER enriches the universality of education.

Studies conducting experiments to develop recommender tools [31], [44], [45] that favor the information of accessible educational resources, as well as the possibility of feeding from diverse student profiles, seek to compare the effectiveness and degree of satisfaction of a student when being able to interact with appropriate resources. The needs and preferences of a student should be in sync with resources that meet those requirements, and education undoubtedly generates valuable educational material that could favor repositories and enrich the educational process that is strengthened in a virtual environment by not knowing borders and having an open availability.

Based on the research carried out, it is established that there are no automatic tools that favor the adaptability and accessibility of an OER considering the correct implementation of its metadata. It is considered that it is necessary to generate new metadata that respond to the guidelines determined by the Universal Design for Learning [46], so we propose as future work, to explore and investigate the strengths and weaknesses of OERs and the implementation of tools that facilitate the proper incorporation of their metadata from the students' experience and their variability in learning. Accessibility in virtual education is an issue that must be socialized, so it is emerging to contribute with OER that responds to the functional diversity of learning.

B. CONCLUSION

The objective of our study was to contribute to lay the foundations for an automatic adaptation of images and audios that meet accessibility standards and correct labeling through their metadata. Despite the limitations of this research, since it is based on scientific literature data, we consider that the bias was reduced by covering the identification of the main databases with more than 600k in their different images and audios.

Four main conclusions can be formulated. First, there is a paucity of applicability of accessibility metadata in OER. Moreover, available studies tend to focus more on design recommendations than on their effective adaptation for optimal interaction with students with disabilities. Second, the use of accessibility standards and metadata is subjective, in several cases responding to evaluative models that, although they consider accessibility as an evaluative metric, it is inconsistent to arrive at a common implementation process. Third, there is a lack of references that establish a significant sample of students with disabilities, their follow-up, monitoring and learning process, which requires more time to obtain reliable data.

In summary, this study reveals an automatic tool proposal that facilitates the implementation of adaptability and accessibility considering its metadata, information that can help other researchers and developers to incorporate the subject matter in the modeling of accessible resources considering the learner's needs and preferences. The integration of accessibility metadata in educational resources and learning objects has a great influence on the effective response of personalized search engines according to the interaction requirements of an educational resource.

Finally, this study reveals that, although contributions throughout history have generated standards and regulations that have motivated research, there is still a lack of automatic tools that favor ideal implementation. The information from quantitative, qualitative or mixed studies is insufficient to determine the impact on students with disabilities, so there is inconclusive data on the applicability of educational resources and search repositories.

ACKNOWLEDGMENT

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors and the Commission cannot be held responsible for any use which may be made of the information contained therein.

REFERENCES

[1] International Council for Open and Distance Education. (2015). *Annual Report 2015*. [Online]. Available: https://icde-beta.squarespace.com/s/annualreport2015website.pdf

[2] L. McAllister, K. Wylie, B. Davidson, and J. Marshall, "The world report on disability: An impetus to reconceptualize services for people with communication disability," *Int. J. Speech-Lang. Pathol.*, vol. 15, no. 1, pp. 118–126, Feb. 2013.

[3] United Nations. (2015). *Quality Education: Why It Matters*. [Online]. Available: https://www.un.org/sustainabledevelopment/education//

[4] United Nations. (2020). *United Nations Treaty Collection*. [Online]. Available: https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IV-15-a&chapter=4&clang=_en

[5] *Incheon Declaration and Framework for Action for the Implementation of Sustainable Development Goal 4*, UN, UNESCO, UNPFA, UNICEF, Paris, France, 2015.

[6] *Information Technology—Individualized Adaptability and Accessibility in E-Learning, Education and training—Part 1: Framework and Reference Model*, Standard ISO/IEC 24751-1:2008, Oct. 2020. [Online]. Available: https://www.iso.org/standard/41521.html

[7] A. Fidalgo, M. L. Sein-Echaluce, D. Lerís, and O. Castañeda, "Teaching Innova project: The incorporation of adaptable outcomes in order to grade training adaptability," *J. Univ. Comput. Sci.*, vol. 19, no. 11, pp. 1500–1521, 2013.

[8] A. Teixeira, C. J. Correia, F. Afonso, A. G. Cabot, E. G. López, S. O. Tortosa, N. Piedra, L. Canuti, J. Guzmán, and M. A. C. Solís, "Inclusive open educational practices: How the use and reuse of OER can support virtual higher education for all," *Eur. J. Open, Distance e-Learn.*, vol. 16, no. 2, pp. 56–65, 2013.

[9] A. Rodriguez-Ascaso, J. G. Boticario, C. Finat, and H. Petrie, "Setting accessibility preferences about learning objects within adaptive elearning systems: User experience and organizational aspects," *Expert Syst.*, vol. 34, no. 4, Aug. 2017, Art. no. e12187.

[10] A. C. Laverde, "Acerca de lo pedagógico en los objetos de aprendizaje-reflexiones conceptuales hacia la construcción de su estructura teórica," *Estudios Pedagógicos (Valdivia)*, vol. 35, no. 1, pp. 261–272, 2009.

[11] D. Wiley and J. Hilton, "Definiendo la pedagogía habilitada para REA," *Revista Mexicana de Bachillerato a Distancia*, vol. 11, no. 21, pp. 153–178, Feb. 2019.

[12] G. Rodríguez, J. Pérez, S. Cueva, and R. Torres, "A framework for improving web accessibility and usability of open course ware sites," *Comput. Educ.*, vol. 109, pp. 197–215, Jun. 2017.

[13] A. M. Teixeira, J. Mota, M. D. C. T. Pinto, and L. Morgado, "Can iMOOCs close the opportunity gaps? The contribution of social inclusive pedagogical design," *Revista Fuentes*, vol. 21, no. 2, pp. 239–252, 2019.

[14] Schema. (2011). *Creativework*. [Online]. Available: http://schema.org/CreativeWork

[15] IMS. (2012). *IMS Access for All V3.0 Public Draft Specification*. [Online]. Available: https://www.imsglobal.org/activity/accessibility#afav3p0I

[16] *Information Technology—Individualized Adaptability and Accessibility in e-Learning, Education and Training—Part 3: 'Access for All' Digital Resource Description*, Standard ISO/IEC 24751-3, 2008. [Online]. Available: https://www.iso.org/standard/43604.html

[17] *Information Technology—Individualized Adaptability and Accessibility in e-Learning, Education and Training—Part 2: 'Access for All' Personal Needs and Preferences for Digital Delivery*, Standard ISO/IEC 24751-2, 2008. [Online]. Available: https://www.iso.org/standard/43603.html

[18] M. Meyer, C. Rensing, and R. Steinmetz, "Multigranularity reuse of learning resources," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 7, no. 1, pp. 1–23, Jan. 2011.

[19] V. E. Robles-Bykbaev, M. López-Nores, J. J. Pazos-Arias, and D. Arévalo-Lucero, "SPELTA: An expert system to generate therapy plans for speech and language disorders," *Expert Syst. Appl.*, vol. 42, no. 21, pp. 7641–7651, Nov. 2015.

[20] C. Troussas, A. Krouska, and M. Virvou, "A multilayer inference engine for individualized tutoring model: Adapting learning material and its granularity," *Neural Comput. Appl.*, vol. 33, pp. 1–15, Feb. 2021.

[21] H. R. Amado-Salvatierra, J. R. H. González, and S. O. Tortosa, "Formalización de un marco metodológico para la implementación de un proyecto educativo virtual accesible," *Educación*, vol. 21, no. 2, pp. 349–367, 2018.

[22] C. Batanero, L. Fernández-Sanz, A. K. Piironen, J. Holvikivi, J. R. Hilera, S. Otón, and J. Alonso, "Accessible platforms for e-learning: A case study," *Comput. Appl. Eng. Educ.*, vol. 25, no. 6, pp. 1018–1037, 2017.

[23] S. G. Temesio Vizoso, "Metadatos de accesibilidad en recursos educativos: Análisis y propuesta," *Palabra Clave*, vol. 7, no. 1, p. 40, Oct. 2017.

[24] *Information Technology—Learning, Education and Training—Quality Management, Assurance and Metric—Part 3: Reference Methods and Metrics*, Standard ISO/IEC 19796-3, 2009. [Online]. Available: https://www.iso.org/standard/46159.html

[25] S. Sanchez-Gordon and S. Luján-Mora, "Research challenges in accessible MOOCs: A systematic literature review 2008–2016," *Univ. Access Inf. Soc.*, vol. 17, no. 4, pp. 775–789, 2018.

[26] L. A. Atiaja and R. G. Proenza, "The MOOCs: Origin, characterization, principal problems and challenges in higher education," *J. e-Learn. Knowl. Soc.*, vol. 12, no. 1, pp. 68–71, 2016.

[27] A. Yousef, M. Chatti, U. Schroeder, M. Wosnitza, and H. Jakobs, "MOOCs: A review of the state-of-the-art," in *Proc. 6th Int. Conf. Comput. Supported Educ. (CSEDU)*, Barcelona, Spain, 2014, pp. 9–16. [Online]. Available: https://files.ifi.uzh.ch/stiller/CLOSER

[28] M. Rohs and M. Ganz, "MOOCs and the claim of education for all: A disillusion by empirical data," *Int. Rev. Res. Open Distrib. Learn.*, vol. 16, no. 6, pp. 1–19, Dec. 2015.

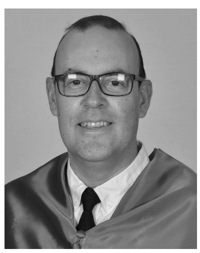
[29] A. I. Morales-Martín. (2018). *Challenges in Open Educational Resources: The Case of TOX-OER MOOC*. [Online]. Available: https://toxoer.files.wordpress.com/2018/02/toxoer-challenges-in-open-education-resources_2018_toxicology.pdf
[30] A. Devaux and M. Souto-Otero, "Validation and open educational resources (OER): Thematic report for the 2016 update of the European inventory on validation," Eur. Centre Develop. Vocational Training (Cedefop), Luxembourg, Tech. Rep. 4149, 2016.
[31] P. Ingavélez-Guerra, V. Robles-Bykbaev, S. Otón, P. Vera-Rea, J. Gálan-Men, M. Ulloa-Amaya, and J. R. Hilera, "A proposal based on knowledge modeling and ontologies to support the accessibility evaluation process of learning objects," in *Proc. Congreso Argentino de Ciencias de la Informática y Desarrollos de Investigación (CACIDI)*, Nov. 2018, pp. 1–5.
[32] C. Hodgkinson-Williams and P. Arinto, *Adoption and Impact of OER in the Global South*. Cape Town, South Africa: African Minds, 2017.
[33] S. Sanchez-Gordon and S. Luján-Mora, "An ecosystem for corporate training with accessible MOOCs and OERs," in *Proc. IEEE 3rd Int. Conf. MOOCs, Innov. Technol. Educ. (MITE)*, Oct. 2015, pp. 123–128.
[34] Y. Vasiliev, *Natural Language Processing With Python and SpaCy: A Practical Introduction*. San Francisco, CA, USA: No Starch Press, 2020.
[35] D. Pérez-Aguilar, R. Risco-Ramos, and L. Casaverde-Pacherrez, "Transfer learning en la clasificación binaria de imágenes térmicas transfer learning for binary classification of thermal images," *Mach. Learn.*, vol. 550, no. 26, p. 4, Jun./Dec. 2021.
[36] R. Luo, G. Shakhnarovich, S. Cohen, and B. Price, "Discriminability objective for training descriptive captions," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6964–6974.
[37] A. Mondal and C. V. Jawahar, "Textual description for mathematical equations," in *Proc. Int. Conf. Document Anal. Recognit. (ICDAR)*, Sep. 2019, pp. 1300–1307.
[38] Clio Awards. *Brands of the World an All Creative World Site*. Accessed: Nov. 17, 2021. [Online]. Available: https://www.brandsoftheworld.com/
[39] J. Treviranus, J. Mitchell, C. Clark, and V. Roberts, "An introduction to the floe project," in *Proc. Int. Conf. Univ. Access Hum.-Comput. Interact.* Switzerland: Springer, 2014, pp. 454–465.
[40] J. Cerda and L. Cifuentes, "Uso de curvas ROC en investigación clínica: Aspectos teórico-prácticos," *Revista Chilena de Infectología*, vol. 29, no. 2, pp. 138–141, Apr. 2012.
[41] A. Rodríguez-Ascaso and J. G. Boticario, "Accesibilidad y MOOC: Hacia una perspectiva integral," *Revista Iberoamericana de Educación a Distancia*, vol. 18, no. 2, pp. 61–85, May 2015.
[42] F. Iniesto and C. Rodrigo, "A preliminary study for developing accessible MOOC services," *J. Accessibility Des.*, vol. 6, no. 2, pp. 126–150, 2016.
[43] R. Navarrete and S. Lujan-Mora, "Microdata with schema vocabulary: Improvement search results visualization of open educational resources," in *Proc. 13th Iberian Conf. Inf. Syst. Technol. (CISTI)*, Jun. 2018, pp. 1–6.
[44] H. Imran, M. Belghis-Zadeh, T.-W. Chang, Kinshuk, and S. Graf, "PLORS: A personalized learning object recommender system," *Vietnam J. Comput. Sci.*, vol. 3, no. 1, pp. 3–13, 2016. [Online]. Available: https://scholar.google.cl/citations?hl=es&user=0yiQ7VoAAAAJ&view_op=list_works&sortby=pubdate
[45] O. C. Santos and J. G. Boticario, "User-centred design and educational data mining support during the recommendations elicitation process in social online learning environments," *Expert Syst., J. Knowl. Eng.*, vol. 32, no. 2, pp. 293–311, Apr. 2015, doi: 10.1111/exsy.12041.
[46] (Oct. 2018). *CAST*. [Online]. Available: http://udlguidelines.cast.org/

PAOLA INGAVÉLEZ-GUERRA (Member, IEEE) received the master's degree in multimedia design from the Universidad del Azuay, Ecuador, in 2014, and the master's degree in a specialty in intellectual diagnosis from the Universidad Tecnica Particular de Loja, in 2015. She is currently pursuing the Ph.D. degree in information and communication technologies with the University of Alcalá. For about ten years, she has been involved in ICT and disability research. For 18 years, she has been a Tenured Professor at Salesian Polytechnic University. She works as the Director of the biomedicine career. She is the author of book chapters and articles published in various magazines. Her main research interests include ICT and disability, accessibility, universal design of learning, and pedagogy. She received the title of a Systems Engineer from the Army Polytechnic School.

VLADIMIR E. ROBLES-BYKBAEV (Senior Member, IEEE) was born in Azogues, Cañar, Ecuador, in 1980. He received the degree in computer science from Universidad Politécnica Salesiana, Ecuador, in 2006, the M.S. degree in artificial intelligence, pattern recognition and digital imaging from the Polytechnic University of Valencia, Spain, in 2008, and the Ph.D. degree in information and communication technologies from the University of Vigo, Spain, in 2016. Since 2008, he has been an Assistant Professor with the Computer Science Department, Universidad Politécnica Salesiana. He is currently a Founding Member of the UNESCO Chair on Assistive Technologies for Educational Inclusion, Universidad Politécnica Salesiana. His research interests include the application of artificial intelligence techniques for improving the educational inclusion of children, youth, and older adults, as well as the rescue and preservation of cultural heritage of Andean people.

ANGEL PÉREZ-MUÑOZ was born in Quito, Pichincha, Ecuador, in 1995. He graduated in computer science from Salesian Polytechnic University, Ecuador, in 2019. Since 2019, he has been a Research Assistant with the Department of Computer Science, Salesian Polytechnic University. He is currently a member of the UNESCO Chair of Assistive Technologies for Educational Inclusion, Salesian Polytechnic University. His research interests include the application of artificial intelligence techniques to improve the educational inclusion of children, youth, and the elderly, as well as serious games and web accessibility.

JOSÉ HILERA-GONZÁLEZ received the B.S. degree in telecommunication engineering from the Polytechnic University of Madrid, Spain, in 1992, and the Ph.D. degree in science (mathematics) from the University of Alcalá, Spain, in 1997. He was a Lecturer with the Polytechnic University of Madrid. He is currently a Full Professor with the Polytechnic School, University of Alcalá. He is the Director of the master's program in software engineering for the web and the Coordinator of the ESVI-AL cooperation network on accessible virtual education, funded by the European Union's ALFA III Program. He has authored or coauthored over 150 scientific works (books, articles, papers, and research projects), some related to learning technology and accessibility. Since their creation, he has been organizing several editions of two international conferences on accessibility, including the International Conference on Quality and Accessibility in eLearning (CAFVIR) and the International Conference on Application of Information and Communications Technologies to improve Accessibility (ATICAcces).



SALVADOR OTÓN-TORTOSA received the degree in computer sciences engineering from the University of Murcia, in 1996, and the Ph.D. degree from the University of Alcalá, in 2006. He is currently an Associate Professor with the Computer Science Department, University of Alcalá. He coordinates the master on agile software development for the web. He is the Coordinator of the European Project EduTech about accessible virtual education, granted by the Erasmus+ Program of the European Union. He is the author or the coauthor of more than 80 scientific works (books, articles, papers, and research projects), the majority of them directly related to learning technology, accessibility, and interoperability. His research interests include learning objects and e-learning standardization and mainly in accessibility, metadata, distributed learning objects repositories, interoperability, and enterprise architectures.