

Received January 11, 2022, accepted February 5, 2022, date of publication February 10, 2022, date of current version February 28, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3150860

# Strategies and System Implementations for Secure Electronic Written Exams

OSVALDO GERVASI<sup>1</sup>, (Senior Member, IEEE), DAMIANO PERRI<sup>1,2</sup>,  
AND MARCO SIMONETTI<sup>1,2</sup>

<sup>1</sup>Department of Mathematics and Computer Science, University of Perugia, 06123 Perugia, Italy

<sup>2</sup>Department of Mathematics and Computer Science, University of Florence, 50121 Florence, Italy

Corresponding author: Osvaldo Gervasi (osvaldo.gervasi@unipg.it)

This work was supported in part by The European Chemistry Thematic Network (ECTN), <https://ectn.eu>, EchemTest® Project, through recurrent donations to the LibreEOL Project of the Department of Mathematics and Computer Science, University of Perugia; and in part by the Italian not for profit organization The International Conference on Computational Science and Its Applications (ICCSA), <https://iccsa.org>, Fiscal ID IT01814970768.

**ABSTRACT** The pandemic has accelerated the process of digitization in many fields, highlighting some critical issues that have slowed the development of digital technologies in many areas. One of these is undoubtedly that of skills assessment, which on a digital platform is a more streamlined, reliable and faster process and allows some individuals from disadvantaged groups to be able to carry out the exam more easily. We present some strategies for enabling online remote secure written exams and their implementation through the open-source system LibreEOL. The system can be used for university exams (in presence or from remote locations) and training courses of all types to evaluate the effectiveness of learning in itinere, i.e., before, during and after a course, to provide quantitative data on training effectiveness. We outline the architecture of the services underpinning the system and the logical organisation of the various user roles and content relating to questions and exam delivery. We describe the algorithms that have allowed us to develop an environment that appears to be an essential aid both for teachers and students in light of the questionnaires conducted on students. The system is successfully used in the EchemTest, promoted by the European Chemistry Thematic Network, and allows comparison of chemistry training outcomes in different European countries. Despite the importance, it has taken on in the last two years due to the pandemic, such a system makes it possible to carry out a valuable skills assessment.

**INDEX TERMS** E-assessment, cloud computing, distributed computing, artificial neural networks, human-computer interaction, user eXperience, COVID-19.

## I. INTRODUCTION

From the very beginning, the dissemination of knowledge and information has been one of the needs most keenly felt by humankind. Over time, the diffusion of ideas has adapted to technological development, taking on characteristics of structural maturity, speed of dissemination and social sharing of knowledge. Over the centuries, we have passed on knowledge orally to writing to arrive at modern times where digital information has become one of the predominant components. Nowadays, switching from traditional assessment methods to digital systems is an extraordinary leap forward in terms of modernity, respect for the student, the diversity with which

different people express their skills and the enhancement of the teacher's abilities.

The COVID-19 epidemic has hastened the digitization process in many industries, revealing several significant challenges that have hindered the development of digital technology in many domains. One of these is, without a doubt, skills testing, which on a digital platform is a more streamlined, reliable, and speedier procedure, allowing certain persons from disadvantaged groups to take the exam more readily. We discuss various ideas for facilitating online distant secure written examinations and how they may be implemented using the open-source system LibreEOL. Our experience in this direction started more than 20 years ago with our participation in a European Union initiative called *Leonardo 2* with a project called DASP [12], [13], [31]. From

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

that embryonic system, we moved in the following years to a full-fledged system capable of conducting computer-based written examinations for various university courses. In 2015, the system was adopted by the *European Chemistry Thematic Network* (ECTN) to carry out the Echemtest<sup>®</sup> sessions for the certification of competencies in the various fields of Chemistry [23].

Starting from January 2020 with the advent of the Pandemic of COVID-19, the system has been completely re-engineered to afford the crucial and delicate task of carrying out written exams in a secure way from remote. The reorganisation of the platform was carried out under the following guiding principles: user-friendliness and optimal user experience, full adherence to open standards, respect for user data protection by always managing the minimum amount of data possible, maximum protection from all possible attacks, scalability and elasticity of the service to ensure adequate response times. We are confident that we have achieved our objectives in light of the data collected through the questionnaires submitted to students and thanks to the volume of examinations taken.

The paper is structured as follows: in Section II an analysis of the research work carried out on the assessment systems is presented. The analysis shows several contributions affirming the importance of e-assessment and highlighting its salient aspects and undisputed advantages. These features revealed strategic importance in the period of the COVID-19 pandemic. Section III outlines the architecture of the services forming the heart of the system, especially concerning its execution in a Cloud environment. We point out the open-source solutions available to achieve the expected efficiency of the system, currently implemented in a commercial Cloud environment. Section IV describes the logical organisation of the various user roles, the content relating to the organisation of courses and questions, the various modality to deliver exams and the available statistics that may help the teacher to monitor the students' performances. Section V describes the algorithms implemented to validate the user's activity while conducting the exam. This part is the most important that has been developed recently, which has required an enormous effort to try to reduce to a minimum the forcing and evident controls on the user, adopting technologies that are non-invasive for the user but at the same time warrant that the teacher can control how the exam is taken by the user, guaranteeing its veracity. Section VI describes carefully the usability of the system, underlying the strategies adopted, inspired to the minimization of operation and the transparency of the operations necessary to achieve a given task successfully. Section VII presents in detail the results obtained, both in terms of quantity of exams carried out and of the satisfaction measured from the users analysing the responses to the questionnaires. Section VIII presents the results of the Data Protection Impact Analysis (DPIA) applied to the LibreEOL<sup>1</sup> platform, which is highly relevant given the very invasive

strategies adopted by many commercial e-assessment solutions. Section IX outlines the main objectives achieved during the development of this system and anticipates possible future developments.

## II. RELATED WORKS

Many scientists and researchers in the field have long been concerned with methods that accurately allow assessing the knowledge, skills, and competencies acquired by learners at the end of a term of study [29].

From early childhood education to university studies, assessment is based on standard protocols and results, measuring the results so that educational institutions responsible for student learning can make decisions based on objective and structured data. Of course, these results must be student-centred and reflect the knowledge, skills, and well-rounded skills preferred by today's students [6], [26]. The development of hardware and software technologies has made it possible to exchange and manage structured and complex information in an elastic way that society has undergone profoundly radical changes [20]. Thus was born the idea of bringing new technologies into the world of teaching. The first outcomes of this attempt undoubtedly show that learning seems to be significantly facilitated by electronic technologies [18], [35], [39].

Educational institutions quickly discovered that e-assessment allowed them to drastically reduce the time it took to create and manage a student's exam, benefiting all involved: the student, the teacher and the educational institution [2], [11]. Students prefer electronic assessments because they have more control over the situation and experience less stress throughout the exam; friendly interfaces and enjoyable environments are more appealing than the conventional ones. It is also quick and straightforward to use, with rapid feedback. The electronic evaluation may undoubtedly boost student performance by increasing motivation. The emotional component is equally crucial because it allows groups to communicate dynamically by neutralising distances [9], [32], [40]. Teachers also wink at e-assessment because it saves time in document correction, allows for the evaluation of a more significant number of students concerning the time available for this activity, aids in the monitoring of student performance and analysis of data obtained thanks to the many possible evaluations, and allows for the classification of concepts exposed to the students' understanding. Finally, because e-assessment provides structured data valuable for management and control available in a short time, the same educational institutions can speed up the decision-making process [5], [10], [15].

With the advent of the pandemic, the level of use of remote communication technologies has suddenly and unexpectedly increased, and e-assessment has played a fundamental role in all of this [42]; educational institutions were forced to resort to the help of online lessons and tests rapidly [4], [37]. This has allowed researchers in the sector to have an enormous amount of data to study, which will allow them

<sup>1</sup>The project URL is: <https://libreeol.org>.

to learn more about the impact these technologies can have on learning [1], [3], [8], [24], [41]. Furthermore, this has made it possible to uncover some critical points of the remote assessment process, such as authentication issues [28], [30], [38], anti-cheating methods [7], [22], and described their approach to validate the exam conducted from a remote location [17], [21].

### III. SYSTEM SERVICE ARCHITECTURE

The system architecture is distributed on a cloud platform through a complex subdivision of scalable and redundant microservices that guarantee very high reliability to service availability offered to users. The overall architecture is divided into different availability areas, and in particular, we have configured two availability zone, zone A and zone B, two regions far enough away to guarantee disaster recovery but still in the surroundings of the city of Milan. This means each microservice, like the file system, or the database, are replicated across multiple areas within the Milan area. This architectural model assures us that LibreEOL functions can continue to be provided to users even in case of faults. A synthetic schema of the system architecture is shown in figure 1.

The client requests are managed by an Application Load Balancer [14], [36], which is a significant application operating at level 7 on the ISO/OSI stack. Its task is to sort the number of active connections among the various Apache nodes, distributing the load over them in the most homogeneous way possible. The algorithm used for load sharing is Least Outstanding Requests (LOR) which assigns new HTTPS connections to nodes with fewer outstanding requests. Nodes are Docker containers running within a cluster managed by an orchestrator, which ensures that these are always properly functional and efficient. If a container no longer responds to requests or has saturated the RAM because of a memory leak, it is immediately restarted. The containers are all identical, running the Ubuntu Linux OS and an Apache webserver. HTTPS requests reach one of the Apache servers and are then handled by returning, for example the LibreEOL homepage. The container cluster uses horizontal autoscaling. Autoscaling increases or decreases the number of active containers in the Docker cluster based on CPU utilization percentage and makes user requests get handled correctly even during peak hours (i.e. when there are many exams simultaneously). If the CPU usage reaches a global average value of 60% on all the containers, the newly incoming requests will be routed to a new container created to empower those already present. As soon as the exams finish, the CPU goes back to lower usage levels and therefore, the autoscaling brings the number of active containers progressively back to a lower quantity, reducing platform costs and keeping the system efficiency at the highest levels all the time. In order to preserve the consistency of the data saved on the storage, the system uses a Network File System distributed throughout mount points individually hooked by each container. In this way, data written on container 1 located in availability zone A can also be correctly processed by container  $N$  in the availability

zone B. Correct data redirection can only be done in real-time by tracking the PHP session. A user could start his interaction with the system on node 1 of Availability Zone A and continue to use the contents through another node thanks to the load balancing performed by the load balancer. A Redis database is used to manage the user session, to assure consistency and reliability. That allows making computational processes' movements transparent to the user between one container and another. Further strength for the architecture created is the possibility of distributing the computation cost due to the SQL requests on the database. The database is a cluster of MySQL servers created through the use of a horizontally scalable architecture that includes a Writer (Master) node and  $N$  Reader (Slaves) nodes. The cluster is replicated on both Availability Zones to guarantee fault-tolerance and High Availability (HA).

LibreEOL is a Web App structured with a PHP-based backend, a JavaScript-based FrontEnd and MySQL is used as RDBMS. We have always tried to adopt standard solutions in developing the code, preferring HTML5 native approaches and respecting good practices in Human Computer Interaction and User eXperience. So, the languages used to create and develop LibreEOL have been JavaScript, PHP, SQL and Bash. A student's test execution is not needed to download additional components or install other software to the web browser, except technical cookies; this guarantees immediacy and ease in using the contents without introducing further overhead to the computer.

### IV. SYSTEM LOGICAL ARCHITECTURE

LibreEOL is a web application based on the Model-View-Controller (MVC) architectural pattern [25] that allows the management of electronic examinations for individual courses and enables the collection of statistical data that can be used to improve the quality of teaching.

#### A. USER ORGANIZATION

Users and content are divided into groups (i.e. Universities) and subgroups (i.e. Departments). Users are divided into the following roles: students, examiners, teachers and administrators. The examiner role is designed to support the teacher in conducting the examination sessions of a course. This role does not have any authority over the information content of questions relating to the course itself. The operations that a user can carry out within the system are determined by the role to which they belong. Each profile is associated with different interfaces with functionalities consistent with the operations allowed within the specific role. It is possible to connect it to an Identity Management service to assign roles automatically, to enhance the platform's usability.

#### B. CONTENT ORGANIZATION

Each teacher can create one or more courses made up of a set of questions classified according to the level of difficulty and topic.

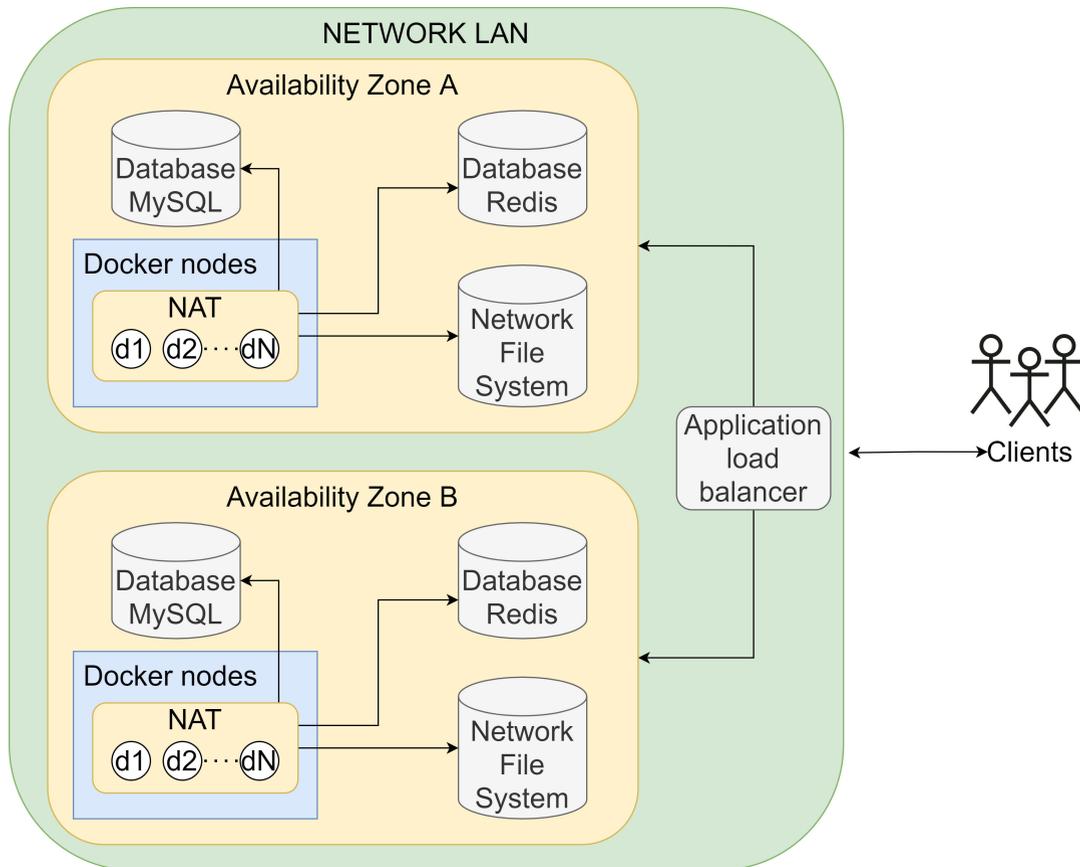


FIGURE 1. Schema of LibreEOL system architecture.

Several teachers and examiners may be associated with each course. The teacher is allowed to express his questions using a set of 11 different types: Multiple choice, Multiple choice with a score, Multiple responses, True/False, Yes/No, Essay, Hotspot, Numerical, Text match, Text match case sensitive and “QR code”. While Multiple Choice supports one answer as true and the others as false, in Multiple Choice with a score it is possible to assign a positive/negative floating point value to all answers. The Hotspot type allows selecting a portion of the image to indicate the answer. The Essay type allows answering by entering free text. A Numerical question allows providing a numerical result (an answer with a tolerance of 3% of the value given by the teacher is accepted). The “QR code” typology allows the upload of the images acquired with a smartphone of a free-written paper, as happens in in-person examinations.

### C. EXAM ORGANIZATION

Each teacher can create examination sessions according to the settings they have configured. In this way, it is possible to create complete examination sessions or mid-term evaluation tests covering a subset of the topics. Creating a new exam requires the teacher to indicate which exam settings to use. These include the possibility of activating or not activating the algorithms for conducting examinations remotely. Once

the exam session has been created, students can register and eventually take some demo tests provided by the teacher. On the day of the exam at the scheduled time, the teacher will provide the password to allow students to access the exam and take the test. If the exam session is held remotely, the teacher can monitor all the students through the “Live view” tool, which allows him/her to observe all the students taking the exam.

At the end of the examination, the teacher will access a panel summarising the tests taken and listing the candidates. This panel allows the teacher to view the individual tests and set the grade, either by accepting the value proposed by the system or by defining it, as is the case with “QR code” or open-ended questions (Essay).

Once the exam has been completed, the teacher archives the exam. The results could be sent by email to the students if the teacher enabled the option while creating the exam.

### D. STATISTICAL ANALYSIS

The system provides various statistics and reports, and it is in this ability, one of the central added values of conducting examinations on an electronic platform is hidden. The teacher can monitor the frequency of correct answers among the various students within a course and between courses of different years. Careful evaluation of these indicators allows

the teacher to calibrate the content delivered and the mode of delivery.

Administrators obtain a distribution of the tests by day and month, enabling them to proactively calibrate the resources made available to the system to carry out horizontal scaling (increasing active resources initially in anticipation of high usage peaks). It is also possible to plan vertical scaling of the infrastructure (switching to more powerful hardware in anticipation of high usage peaks).

## V. ALGORITHMS IMPLEMENTED TO VALIDATE STUDENTS'S ACTIVITY

The central aspect that deserves serious consideration when one decides to take exams remotely is to ensure that the test measures the actual skills, quality and quantity of a student's knowledge. Furthermore, it is also essential to ensure that the test has legal validity, i.e. it is necessary to prevent the student from receiving suggestions or information from the outside. The technology developed by us is divided into various components that analyze the ambient audio, the video stream captured by the webcam, the mouse movements and the keyboard. In the design phase, we took particular care of efficiency, and we intended to focus on a code capable of running on inexpensive devices with low computational power. Students often own old laptops and smartphones. Therefore, our technique is completely based on geometry and basic mathematical operations that make our algorithm fast and executable on most of the hardware used by students. Neither a dedicated graphics accelerator is required. All the algorithms presented have properly been run on dual-core CPUs produced since 2010.

### A. ALGORITHM FOR VIDEO STREAM ANALYSIS

Our objective is for the student to take the test without being distracted or receiving illicit assistance. The system requires a camera and microphone to verify how the test is being carried out remotely; thus, the test may be conducted from remote locations. The images captured by the camera represent sensitive data and are rich in information. We built our algorithm to keep the number of photographs and sounds saved on the server to a bare minimum and keep the server's information for as little time as possible. Each student had to provide permission for this sensitive data to be processed before taking the exam.

#### 1) IDENTIFICATION OF THE FACE MOVEMENTS

A very significant aspect for identifying a student misbehaviour concerns the identification of video distractions, intended as too wide and sudden movements of the head. In figure 2 the 68 landmarks produced by the neural network that we use for this function are shown. Neural networks perform the task of automatically classifying classes of objects in the most varied fields of application [33], [34]; in our specific case, the network has been developed in such a way that it can recognise the landmarks of the user's face.



FIGURE 2. The identification of the landmarks computed by the neural network from the collected webcam flow of images.

Starting from analysing the subject's face and its movements in space, we can assume with high reliability whether his/her conduct has been inappropriate. This is obtained using a neural network that provides landmarks extraction and detects 68 points on the subject's face representing them by X, Y coordinates in a Cartesian plane.

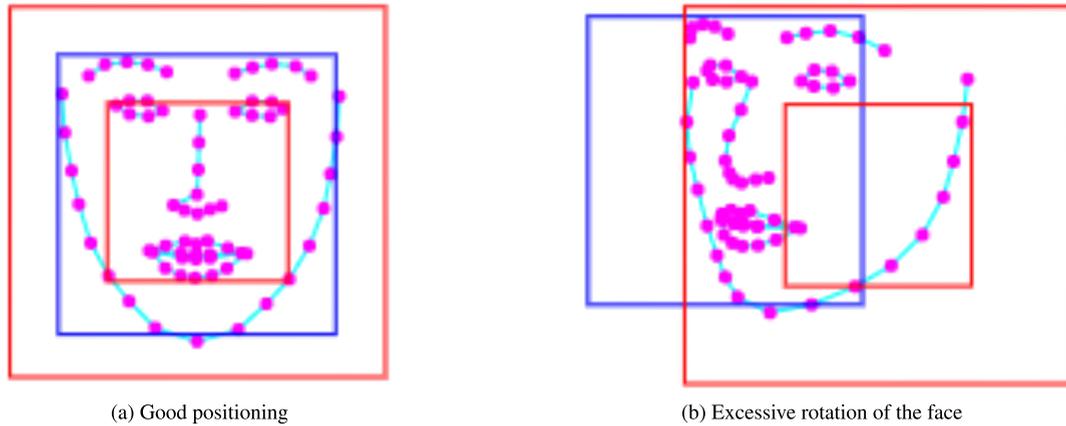
In figure 3(a), we can graphically analyze the algorithm computation. The photograph shows explicitly a student who is staring at the monitor. It is possible to notice that the face is contained within two red-bordered rectangles, an inner and an outer one. The blue rectangle represents the polygon tangent to the outermost points of the student's face and represents our centre of gravity. These rectangles represent the tolerance of the system. The algorithm computes as follows: in the beginning, the neural network computes the landmarks, and the algorithm evaluates the relative tolerance threshold so that the student can move, breath, yawn. These two rectangles are periodically recalibrated over the student's face in order to ensure that they are always consistent with his/her position. It is, in fact, taken into account that during an exam that can even last 2 hours, the student can move slightly to get more comfortable. If the learner turns away from the screen and does not look at it, the blue square will touch or even exceed the boundaries set by the red rectangles' corners.

In figure 3 (b) an example about this behaviour has shown.

In that case, a photograph of the event is taken, and a distraction is counted.

#### 2) MONITORING OF ALL STUDENTS TAKING THE EXAM

Low-resolution images are extracted from the webcam images at 2-second intervals and displayed on the exam's associated *Live View* page so that the teacher can monitor the progress of the exam as a whole. Based on the images shown, the teacher can send Chat messages to individual students or



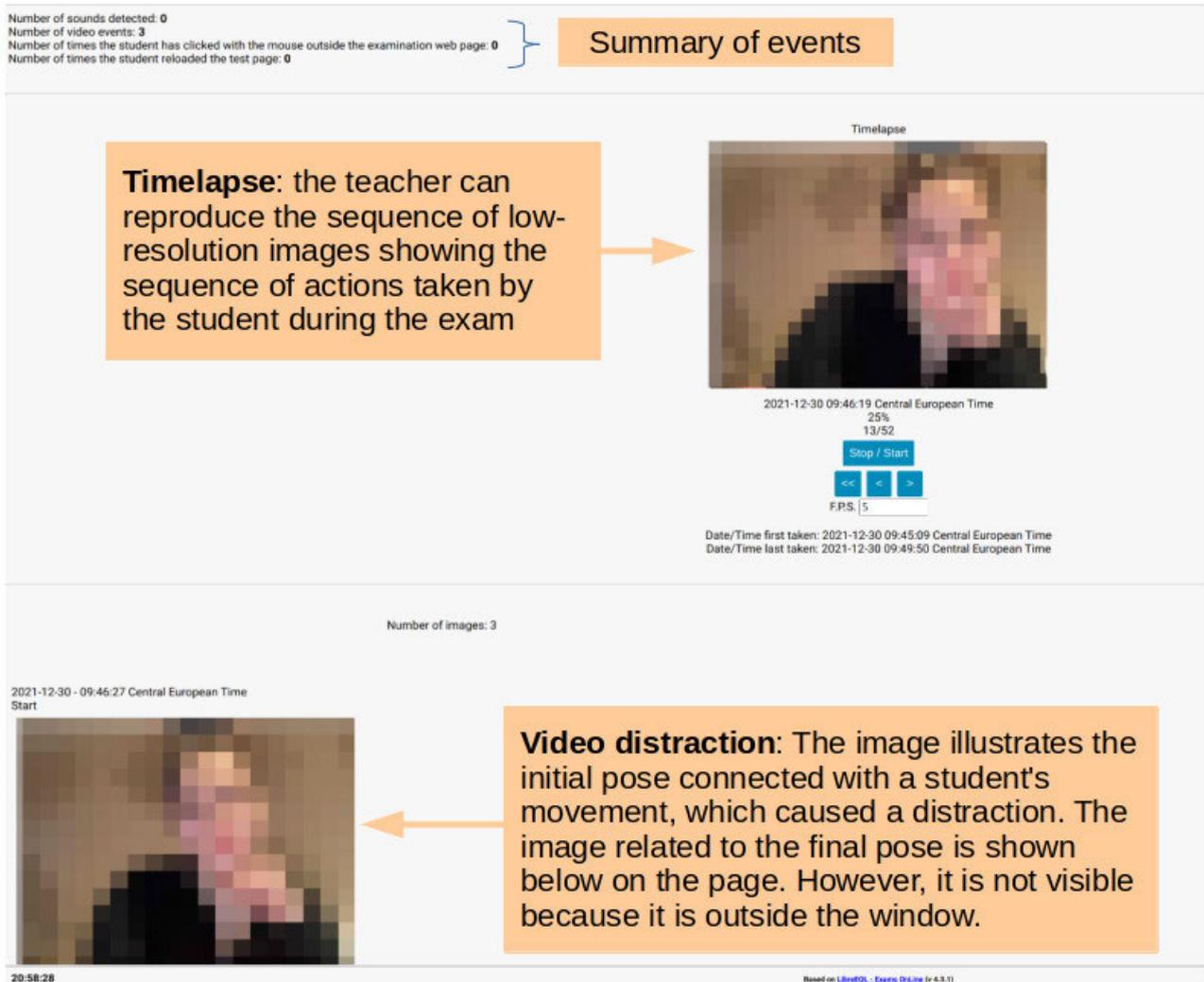
**FIGURE 3.** Areas computed by our algorithm when the user is properly positioned in front of the webcam (a) and when she/he turns the head, causing a distraction to be detected by the system (b).



**FIGURE 4.** Annotated panel showing the Live View of a running exam (the image has been altered to protect the personal data of the persons in the picture).

all the participants. This tool has an enormous impact and power on the teacher and allows them to establish important contact with students during the exam. The images shown in the “Live View” panel are also displayed on the page dedicated to each student, showing the various video distractions, any audio recorded, the times the client was disconnected

from the server and for how long, and the number of times the exam page was reloaded. This page is essential for assessing the correctness of the test taken by the student. Figure 4 shows the Live View environment described above while the exam is in full swing. Figure 5 shows a student’s summary page with indications of the events that characterized the test and allow



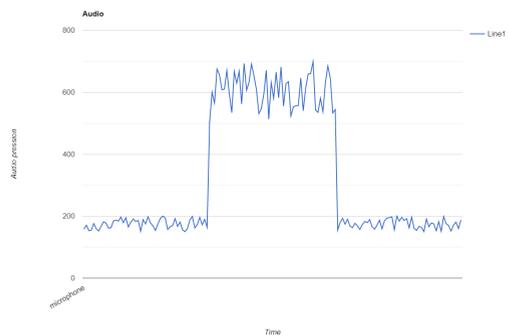
**FIGURE 5.** Annotated panel showing the summary information of a remote exam (the image has been altered to protect the personal data of the person in the picture).

the teacher to make a final decision on the correctness of the test.

**B. ALGORITHM FOR ANALYSING ENVIRONMENTAL AUDIO**

Through JavaScript code we obtain the audio stream from the microphone. The audio is analysed with a sampling rate of 48K times per second and a bit rate of 64KBps. The audio we capture is stereophonic. We first calculate the average value of the ambient sound pressure. If we detect a peak, i.e. a sound with an intensity more remarkable than the average of the last period, then, for the duration of the event, we record an audio file that in this first phase will be stored in the RAM memory of the computer. The audio file is recorded in OGG format. This is an open format, open-source and released under a BSD license.

The OGG format can store audio encoded using the lossy compression algorithm 'Vorbis', also released under a BSD licence. The recording phase, monitoring of ambient sound



**FIGURE 6.** A sample audio monitoring graph.

pressure and coding of the audio flow, takes place entirely on the client, i.e. it is carried out by the student's computer, with a low expenditure of computational resources. An example that graphically shows the variation of the ambient sound pressure during a speech can be seen in figure 6. The system then checks the length of the recorded audio file. If it is

less than 1 second long, it is not uploaded to the server. If, on the other hand, its duration is greater than or equal to one second, then it is securely stored in the system, and the audio event associated with the student is recorded in the database. Monitoring takes place throughout the examination.

### C. MOUSE AND KEYBOARD MONITORING ALGORITHM

To prevent students from obtaining content and information on the Internet, we built algorithms that ensure that they only sit the test with the knowledge they have gained throughout their studies. The examination screen is enlarged and displayed in full-screen mode when the test is started. Any windows that may have been open are then covered by the application screen, which will occupy the entire visible area of the student's screen. At the same time, mouse monitoring is started; in particular, we want to prevent the mouse from leaving the examination screen and clicking on external areas, such as on a second monitor. To do this, we first install an event listener on the web page that will monitor the property "focus". This property tells us whether or not the web page is in the foreground, or whether the user is clicking or acting on elements in the background or on other monitors. The loss of focus, in fact, guarantees that the student has clicked outside the examination window.

If a student clicks outside the examination window, a picture of the event is taken, and the timestamp of the event is stored in a server-side counter. Before starting an exam, the teacher can define the maximum number of times a student can click outside the exam window. It is essential to provide a certain margin from our tests, and not be too strict on this aspect. We think it is correct to set the 'focus lost' parameter to 3. It is, in fact, possible that during the exam, the student has notification from the antivirus program requesting an interaction, or he has audio problems and needs to click on the headset control panel, or it is possible that he accidentally clicks out of the exam. In any case, once the occurrences the student loses focus on the page exceed the parameter set by the teacher, the exam is closed and the student is ejected from the session. In addition to this, a series of checks are carried out on the keyboard and on the mouse's interaction with the page. These controls prevent the copying/pasting of text to and from the examination page.

```

1 $(document).mouseleave(function () {
2     console.log("Focus lost !");
3 });
4 $(document).mouseenter(function () {
5     console.log('Focus in');
6 });
7 $(window).blur(function () {
8     if(!document.activeElement instanceof
9         HTMLIFrameElement ){
10         console.log("Focus lost !");
11     }
12 });

```

LISTING 1. Mouse control code.

Furthermore, the page printing functionality, the screen recording functionality and the print screen functionality are inhibited.

## VI. SYSTEM'S USABILITY

The system's usability is one of the essential aspects of designing a software application. A modern web application must be easy to use, must not require lengthy explanations of how the graphical interface works, be respectful of minorities and consider users' disabilities. Furthermore, applications of this type should not require any special skills on the part of the users. We have also considered the different types of users, and the disabilities that people using the system may have.

For example, some students have difficulty using the mouse or carrying out drag and drop operations, so we have designed an interface that does not require the mouse, drag and drop movements or operations requiring complex motor coordination. A low percentage of students also have Specific Learning Disorders (SLD) to varying degrees. An e-assessment system allows the teacher to offer these students a test tailored to their needs, including more time for the test and questions formulated to facilitate comprehension or mathematical calculations.

When developing the user interface, we took the importance of colours seriously, so we chose a colour palette that would not cause confusion or problems for people with colour blindness. According to the most recent estimates, 4.2 per cent of the world's population is colour blind, and in Italy itself there are 2.5 million people affected by this condition [16], [19], [27]. In the light of these considerations, all the messages that the graphical interface proposes to users have been calibrated on an appropriate shade of colour compatible with Protanopia (insensitivity to red), Protanomalia (insufficient sensitivity to red), Deuteranopia (insensitivity to green), Deuteranomalia/Teranomalia (insufficient sensitivity to green), Tritanopia (insensitivity to blue, violet and yellow) and Tritanomalia (insufficient sensitivity to these colours). In addition, we have taken particular care to ensure the compatibility of the system with the main web browsers (Google Chrome, Microsoft Edge, Mozilla Firefox and Opera) and also compatibility with old and resource-poor computers; this has involved taking care to minimise the computational impact of the application on the client's hardware resources.

In many disciplines such as mathematics, chemistry, physics or the humanities, it is crucial to write by hand on sheets of paper, carry out calculations better, prove theorems, and make graphs or drawings. To this end, a particular type of question has been created that allows students to securely upload photographs of the paper they have produced, through a QR Code using a smartphone. The smartphone is necessary to take high-resolution photos of the paper produced by the student. Once the camera has framed the QR Code, this is decoded and the web page dedicated to uploading the images appears. Each QR Code uniquely associates the paper produced by the student with his examination.

We submitted a questionnaire to the users, allowing them to voluntarily express a series of considerations on the qualitative assessment and problems encountered during the use of the system. We obtained a satisfaction index of 81.4% of the sample, indicating an important general appreciation level. Thanks to an agile programming approach and constant interaction between developers and end-users, the system's usability is continuously improved.

## VII. DISCUSSION OF RESULTS

From 21 February 2021 to 21 September 2021, we kept a direct communication channel open between the platform and the students by delivering an online questionnaire. That allowed us to collect valuable information directly from the primary users of the platform. The students voluntarily participating in the anonymous questionnaire were 1423, expressing their opinions and views on the system we have developed. The questionnaire consisted of eleven questions divided into eight sections. The preliminary questions were technical and aimed at understanding the tools used by the students to access the examination platform. The first question asked which browser the student was using, while the second asked about the operating system. The results show that 79.4% of the students use Google Chrome, 8% use Firefox, and 5.8% use Safari. The remaining 6% of the students use less popular browsers such as Brave, Chromium, Microsoft Edge, Internet Explorer, Opera and Samsung's Browser Internet. The second question got the following responses: 62.5% of the students use Windows 10, 11.9% of the students use macOS, in the third position we find Windows 7, used by 8.2% of the students. The results obtained from these first two questions are highly significant as they show us that the development environment on which to program and develop the software code must first be tested on the Google Chrome browser. As far as the functionality of the application we have created is concerned, the operating system is of less importance since we do not use any particular APIs that would make Windows rather than other systems decisive. In the preliminary software's development, we have tried to devise a code that runs on as many devices as possible.

The next question asked the students whether they experienced particular events during their examination, such as unexpected notifications, irrelevant error reports or unexpected behaviour. 84% of the students answered that they had not experienced any particular events, while 16% answered that they thought the system had behaved unexpectedly. These students, in particular, were asked, employing an open-ended question, to describe the event that had occurred during the examination session. We collected a total of 228 answers. Some of them explained that the cause of the problems they experienced was slow internet line at home, which delayed the loading of images and texts. Others described that notifications from the operating system, not caused by LibreEOL, had appeared in the foreground and distracted them, such as notifications of antivirus software, new emails or system updates. The most interesting answers are related to the

algorithm we programmed for face tracking. In fact, some students reported that they had received warnings following small movements of the face that are part of the natural behaviour in concentration situations. These latter responses were precious to us, as we immediately took them into account and made adjustments, corrections and calibrations to the system to address these issues. In particular, we calibrated the face at the start of the test and at regular intervals to ensure that the face tracking also took into account the natural posture adjustments that people make when they sit in front of the computer for one or more hours. In addition, we have made sure that the distraction notifications we give to students are even better calibrated. These notifications are important because they allow the student to be warned if the system detects incorrect conduct. Nevertheless, it is essential to consider that the number of notifications and warnings that an automated system can give to a person while he or she is taking an exam must be wisely limited. There may be situations that are not caused by a person's failure to comply with the most basic behavioural rules but are caused by other problems, such as a webcam with poor resolution, or poor room lighting, or an incorrect framing of the face that could generate a high number of alerts. In this case, we stop sending notifications to students when a certain threshold is exceeded. According to our tests, the optimal maximum number of alerts per person is 6. More warnings could only be counterproductive and lead to the person becoming impatient. It should also be borne in mind that the teacher is responsible for monitoring the students' video streams throughout the exams. So, the teacher can also view the counters for the number of facial distractions, and the number of audios detected and analyse them in real-time via a secondary browser window. Thanks to these measures that we have implemented in the platform, we have not received any further reports of inappropriate warnings given to students.

The next question asked the students whether they experienced any additional discomfort during the examination. This question was answered by 100% of the students. Among them, 75.9% answered that they had not experienced any discomfort, while 24.1% (343 students) stated that they had experienced further problems. These students were then asked to express what kind of discomfort they had experienced while using the system through an open-ended text field.

The students answered that they had experienced anxiety were 76, while 12 students stated they had experienced stress during the examination. Students who experienced anxiety represented 5% of the total number of students interviewed, while those who experienced stress represented 0.8% of the total. Their valuable evaluations and the relevant advice provided to us by this small percentage of students contributed to an highly significant improvement in the usability and characteristics of the system. We realized that to reduce the students' stress and anxiety levels, we needed to implement a series of algorithms to help them feel less alone. The algorithms we

introduced provide visual and textual feedback to the student, reassuring them that they are progressing well.

The first algorithm relates to the auto-saving function of the exam, which occurs automatically at periodic intervals. This function allows the student's answers to be stored in the system's database before the exam is finally handed in. Auto-saving allows the exam to be retrieved if the student's computer crashes, the power goes out, a network disconnection occurs, or other problems prevent the exam from continuing. Each time the exam is auto-saved, a small green text appears to the student informing him that the test has been correctly saved in the database.

The second algorithm implemented carries out a periodic check on the photographs that are loaded into the system relating to the QR codes displayed on the student's exam page. When the system detects that the student has uploaded a photograph via QR code, a photograph relating to an exercise that he had to carry out on paper, the system will inform him that the photograph has been saved correctly on the exam page and not only on the smartphone screen.

The third algorithm we implemented allows direct communication via chat messages between the student and the teacher. One of the reasons that caused the most anxiety was the lack of communication between the teacher and student during the exam. This feature allows students to feel part of a community and not just an individual mechanically assessed by a software programme.

The next question asked for an overall assessment of the system's quality: 81.4%, i.e. 1158 students, said that the system has a high level of quality and praised its features; 10.3% of the students considered the software invasive from the point of view of privacy; 8.3% of the students felt that the system was not appreciated and that it should not be used.

Thanks to the data collected, we received a confirmation of the goodness of the solutions adopted to manage better and protect personal data. In fact, through application's design, only a minimal amount of sensitive data (photos, sounds, text of chats) are collected. They are saved on the file system only for the time needed by the teacher to carry out the examination procedures.<sup>2</sup>

We have met with student representatives several times, collected suggestions and objections, and publicised the guides we have produced and all the documentation available on the website. We have explained to them the techniques adopted to protect personal data, receive their active collaboration and calm the students' spirits.

All facial detection and tracking algorithms only work on the student's computer, and no biometric data is processed or sent from the clients to the server.

<sup>2</sup>An exception to this procedure is represented by the information collected for students who have been excluded from the exam, either by the teacher or the system. This information is saved encrypted on a file system outside the web domain for 130 days so that, in the event of legal proceedings, the organization will be able to exercise its defence by producing this evidence.

The last question was an open-ended one. We asked the students to express any criticism or constructive proposals for improving the software. A total of 248 people responded, most of the criticisms and suggestions related to topics we discussed earlier, such as more effective face calibration and the request to have visual feedback on events that occur during the exam (such as self-saving) and the request to have the possibility to speak even only verbally with the teacher. These were all valuable suggestions that motivated us to refine and release changes that had already been planned and developed in the test environment, such as the chat environment, real-time calibration to adapt the control system to the natural movements of the person, refining the online documentation, enhancing support via social networks and email.

## VIII. DATA PROTECTION ISSUES

The European Union adopted in 2016 the General Data Protection Regulation (GDPR).<sup>3</sup> A processing operation is likely to pose a high risk to the rights and freedoms of data subjects because of the systematic monitoring of their behaviour, the large number of data subjects whose sensitive data are perhaps processed, or even because of a combination of these and other factors. In such cases, GDPR obliges data controllers to conduct an impact assessment before commencing the processing, consulting the supervisory authority if the technical and organizational measures they have identified to mitigate the impact of the processing are not deemed sufficient, i.e., when the residual risk to the rights and freedoms of data subjects remains high.

We have drafted the Data Protection Impact Analysis (DPIA) for LibreEOL using the program made available by the French Data Protection Authority, CNIL,<sup>4</sup> a multi-lingual open-source distribution made available for all Linux distros and Windows 10 under the Windows Subsystem for Linux (WSL) as *appimage*.<sup>5</sup>

Following the document's structure proposed by the CNIL, the operations involving the management of personal data (images of the student, of the room in which the examination is conducted and environmental audio) have been detailed, specifying the technical solutions adopted and the associated risk. This phase of analysis allowed us to adopt appropriate changes to the code of the proposed application, aimed at minimizing the amount of data collected, the length of the data retention period before deletion and the methods of secure and encrypted storage of data that must be retained for more extended periods.

In some situations, unfortunate choices had to be made, in the sense that it was necessary to give up handling data that would have allowed better support to users but could give rise

<sup>3</sup>The General Data Protection Regulation is a European Union regulation on the processing of personal data and privacy, adopted on April 27, 2016, published in the Official Journal of the European Union on May 4, 2016, and entered into force on May 24 of the same year and operational as of May 25, 2018. It is known as Regulation 2016/679.

<sup>4</sup>Please visit the URL: <https://www.cnil.fr/fr/analyse-dimpact-la-cnil-publique-la-version-3-0-de-son-logiciel-pia>

<sup>5</sup>Please visit the URL: <https://appimage.org/>

to ambiguous interpretations on the appropriateness of such processing.

The outcome of the impact assessment was that the risk was negligible.

## IX. CONCLUSION AND FUTURE WORK

We have created a platform for secure remote online exams that respects personal data and the users' primary needs. Security by design and privacy by design were the foundations of our work. During the COVID-19 pandemic, the platform proved helpful, with the University of Perugia adopting it as the official tool for conducting exams remotely. The high level of participation in the questionnaire we administered, together with the high level of acceptance and the impressive number of examinations taken in one year (over 50,000), are, in our opinion, the best proof of the validity of our work.

Thanks to the difficult period that we have experienced at world level, we believe that, at least in our university, an extraordinary cultural revolution has taken place, driven by students and led to a profound transformation among the teaching staff. They have experimented mainly with an efficient and effective way of evaluation, made possible thanks to a re-thinking of how evaluation is carried out. We are confident that this acquired patrimony will remain over time and facilitate the teachers' evaluation work and allow the students to expound their competencies and knowledge more completely.

Our future work on the platform will be inspired by a desire to increase support for teachers and students by identifying new forms of questions and answers that make the assessment process as natural, intuitive, and immediate as possible. We also wish to increase support for those who experience difficulties in their daily lives, for example, by implementing screen-reading assistants or voice assistants for entering answers to questions via voice feedback provided by the student.

## ACKNOWLEDGMENT

The authors would like to thank the European Chemistry Thematic Network (ECTN) and the EchemTest® Project for the fruitful collaboration. The authors would also like to thank the Data Protection Officer (DPO) of the University of Perugia, Dr. Flavia Cristiano, for the careful review of the data protection impact analysis related to LibreEOL and for the stimulating and appropriate guidance provided on GDPR issues.

## AUTHOR CONTRIBUTIONS

Conceptualization—Osvaldo Gervasi, Damiano Perri, and Marco Simonetti; data curation—Osvaldo Gervasi, Damiano Perri, and Marco Simonetti; investigation—Osvaldo Gervasi, Damiano Perri, and Marco Simonetti; methodology—Osvaldo Gervasi, Damiano Perri, and Marco Simonetti; software—Damiano Perri; supervision—Osvaldo Gervasi; validation—Osvaldo Gervasi, Damiano Perri, and Marco Simonetti; writing (original draft)—Osvaldo Gervasi,

Damiano Perri, and Marco Simonetti; and writing (review and editing)—Osvaldo Gervasi, Damiano Perri, and Marco Simonetti.

## REFERENCES

- [1] S. R. Alotaibi, "A novel framework of success using of E-Assessment during corona pandemic," *Int. J. Emerg. Technol. Learn. (iJET)*, vol. 16, no. 12, p. 215, Jun. 2021.
- [2] N. Alruwais, G. Wills, and M. Wald, "Advantages and challenges of using e-assessment," *Int. J. Inf. Educ. Technol.*, vol. 8, no. 1, pp. 34–37, 2018.
- [3] A. A. Ayyoub and O. Jabali, "University students' evaluation of E-assessment in light of the coronavirus pandemic," *Cypriot J. Educ. Sci.*, vol. 16, no. 4, pp. 1434–1449, Aug. 2021.
- [4] F. M. Azmi and H. B. Khoshaim, "The COVID-19 pandemic and the challenges of E-assessment of calculus courses in higher education: A case study in Saudi Arabia," *Int. J. Learn., Teaching Educ. Res.*, vol. 20, no. 3, pp. 265–281, Mar. 2021.
- [5] R. E. Bennett, "Inexorable and inevitable: The continuing story of technology and assessment," *Comput.-Based Test. Internet, Issues Adv.*, vol. 1, no. 1, pp. 201–217, 2006.
- [6] N. A. Buzzetto-More and A. J. Alade, "Best practices in e-assessment," *JITE-Res.*, vol. 5, no. 1, pp. 251–269, 2006.
- [7] D. Bylieva, V. Lobatyuk, S. Tolpygin, and A. Rubtsova, "Academic dishonesty prevention in e-learning university system," in *Proc. World Conf. Inf. Syst. Technol.*, Cham, Switzerland: Springer, 2020, pp. 225–234.
- [8] V. Deepika, K. Soundariya, K. Karthikeyan, and G. Kalaiselvan, "'Learning from home': Role of e-learning methodologies and tools during novel coronavirus pandemic outbreak," *Postgraduate Med. J.*, vol. 97, no. 1151, pp. 590–597, 2021.
- [9] J. Donovan, C. Mader, and J. Shinsky, "Online vs. traditional course evaluation formats: Student perceptions," *J. Interact. Online Learn.*, vol. 6, no. 3, pp. 158–180, 2007.
- [10] A. J. Edson, "Learner-controlled scaffolding linked to open-ended problems in a digital learning environment," *ZDM*, vol. 49, no. 5, pp. 735–753, Oct. 2017.
- [11] M. A. H. Eljinini, S. Alsamarai, S. Hameed, and A. Amawi, "The impact of e-assessments system on the success of the implementation process," *Int. J. Modern Educ. Comput. Sci.*, vol. 4, no. 11, p. 76, 2012.
- [12] O. Gervasi, R. Catanzani, A. Riganelli, and A. Laganá, "Integrating learning and assessment using the semantic web," in *Computational Science and its Applications—ICCSA 2005*, O. Gervasi, M. L. Gavrilova, V. Kumar, A. Laganá, H. P. Lee, Y. Mun, D. Taniar, and C. J. K. Tan, Eds. Berlin, Germany: Springer, 2005, pp. 921–927.
- [13] O. Gervasi and A. Laganá, "Eol: A web-based distance assessment system," in *Computational Science and its Applications—ICCSA 2004*, A. Laganá, M. L. Gavrilova, V. Kumar, Y. Mun, C. J. K. Tan, and O. Gervasi, Eds. Berlin, Germany: Springer, 2004, pp. 854–862.
- [14] O. Gervasi, D. Perri, M. Simonetti, and S. Tasso, "High performance computing and computational intelligence applications with multi-chaos perspective," in *Multi-Chaos, Fractal and Multi-Fractional Artificial Intelligence of Different Complex Systems*, Y. Karaca, D. Baleanu, Y.-D. Zhang, O. Gervasi, and M. Moonis, Eds. Amsterdam, The Netherlands: Elsevier, Apr. 2022, p. 40.
- [15] L. Gilbert, D. Whitelock, and V. Gale, "Synthesis report on assessment and feedback with technology enhancement," Univ. Southampton Institutional Repository, Southampton, U.K., Tech. Rep., Jul. 2011. [Online]. Available: <http://eprints.soton.ac.uk/id/eprint/273221>
- [16] N. Gordon, "Colour blindness," *Public Health*, vol. 112, no. 2, pp. 81–84, 1998.
- [17] I. G. Hormigo, M. E. Rodríguez, and X. Baró, "Design and implementation of dashboards to support teachers decision-making process in e-assessment systems," in *Engineering Data-Driven Adaptive Trust-based e-Assessment Systems*. Cham, Switzerland: Springer, 2020, pp. 109–132.
- [18] H. Ibrahim and A. S. H. Alqahtani, "The impact of adopting web 2.0-based E-book on student learning skills," *EURASIA J. Math., Sci. Technol. Educ.*, vol. 14, no. 6, pp. 2509–2522, Apr. 2018.
- [19] L. Jefferson and R. Harvey, "An interface to support color blind computer users," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, Apr. 2007, pp. 1535–1538.
- [20] R. A. Khan and D. A. Khan, "Cloud migration: Standards and regulatory issues with their possible solutions," *Int. J. Adv. Neww. Appl.*, vol. 10, no. 6, pp. 4113–4119, 2019.

- [21] Y. Khelifi and A. Hassan El-Sabagh, "A novel authentication scheme for e-assessments based on student behavior over e-learning platform," *Int. J. Emerg. Technol. Learn.*, vol. 12, no. 4, pp. 62–89, Apr. 2017.
- [22] C. Kienert, N. D. Vos, M. Knockaert, and J. Garcia-Alfaro, "The influence of conception paradigms on data protection in E-learning platforms: A case study," *IEEE Access*, vol. 7, pp. 64110–64119, 2019.
- [23] A. Laganà, O. Gervasi, S. Tasso, D. Perri, and F. Franciosa, "The ECTN virtual education community prosumer model for promoting and assessing chemical knowledge," in *Computational Science and its Applications—ICCSA 2018*, O. Gervasi, B. Murgante, S. Misra, E. Stankova, C. M. Torre, A. M. A. C. Rocha, D. Taniar, B. O. Apduhan, E. Tarantino, and Y. Ryu, Eds. Cham, Switzerland: Springer, 2018, pp. 533–548.
- [24] C. Lazzari, S. McAleer, A. Nusair, and M. Rabottini, "Psychiatric training during covid-19 pandemic benefits from integrated practice in interprofessional teams and ecological momentary e-assessment," *Rivista di psichiatria*, vol. 56, no. 2, pp. 74–84, 2021.
- [25] A. Leff and J. T. Rayfield, "Web-application development using the model/view/controller design pattern," in *Proc. 5th IEEE Int. Enterprise Distrib. Object Comput. Conf.*, 2001, pp. 118–127.
- [26] T. Love and T. Cooper, "Designing online information systems for portfolio-based assessment: Design criteria and heuristics," *J. Inf. Technol. Educ., Res.*, vol. 3, no. 1, pp. 65–81, 2004.
- [27] M. F. Marmor, "Vision, eye disease, and art: 2015 keeler lecture," *Eye*, vol. 30, no. 2, pp. 287–303, Feb. 2016.
- [28] H. Mellar, R. Peytcheva-Forsyth, S. Kocdar, A. Karadeniz, and B. Yovkova, "Addressing cheating in e-assessment using Student authentication and authorship checking systems: Teachers' perspectives," *Int. J. Educ. Integrity*, vol. 14, no. 1, pp. 1–21, Dec. 2018.
- [29] D. Nicol, "E-assessment by design: Using multiple-choice tests to good effect," *J. Further Higher Educ.*, vol. 31, no. 1, pp. 53–64, Feb. 2007.
- [30] A. Okada, D. Whitelock, W. Holmes, and C. Edwards, "E-authentication for online assessment: A mixed-method study," *Brit. J. Educ. Technol.*, vol. 50, no. 2, pp. 861–875, Mar. 2019.
- [31] F. Giorgetti, O. Gervasi, and A. Laganà, "Distance assessment system for accreditation of competencies and skills acquired through in-company placements (dasp), 1999," in *Proc. INET Internet Global Summit, 9th INET Int. Conf.*, Jun. 1999, pp. 22–25.
- [32] K. Penny and J. Brodie, "Student and staff perceptions of the use of multiple choice testing in higher education assessments," in *Proc. EDEN Conf.*, Fairfax, VA, USA: George Mason Univ. Libraries, 2014, pp. 184–192.
- [33] D. Perri, M. Simonetti, A. Lombardi, N. Faginas-Lago, and O. Gervasi, "Binary classification of proteins by a machine learning approach," in *Computational Science and its Applications—ICCSA 2020*, O. Gervasi, B. Murgante, S. Misra, C. Garau, I. Blečić, D. Taniar, B. O. Apduhan, A. M. A. C. Rocha, E. Tarantino, C. M. Torre, and Y. Karaca, Eds. Cham, Switzerland: Springer, 2020, pp. 549–558.
- [34] D. Perri, M. Simonetti, A. Lombardi, N. Faginas-Lago, and O. Gervasi, "A new method for binary classification of proteins with machine learning," in *Computational Science and its Applications—ICCSA 2021*, O. Gervasi, B. Murgante, S. Misra, C. Garau, I. Blečić, D. Taniar, B. O. Apduhan, A. M. A. C. Rocha, E. Tarantino, and C. M. Torre, Eds. Cham, Switzerland: Springer, 2021, pp. 388–397.
- [35] D. Perri, M. Simonetti, S. Tasso, and O. Gervasi, "Learning mathematics in an immersive way," in *Software Usability*. London, U.K.: IntechOpen, Mar. 2021, pp. 1–15.
- [36] D. Perri, M. Simonetti, S. Tasso, F. Ragni, and O. Gervasi, "Implementing a scalable and elastic computing environment based on cloud containers," in *Computational Science and its Applications—ICCSA 2021*, O. Gervasi, B. Murgante, S. Misra, C. Garau, I. Blečić, D. Taniar, B. O. Apduhan, A. M. A. C. Rocha, E. Tarantino, and C. M. Torre, Eds. Cham, Switzerland: Springer, 2021, pp. 676–689.
- [37] R. Peytcheva-Forsyth and L. Aleksieva, "Forced introduction of e-assessment during COVID-19 pandemic: How did the students feel about that? (Sofia University case)," in *Proc. AIP Conf.*, vol. 2333, 2021, Art. no. 050013.
- [38] N. Sclater, G. Conole, and B. Warburton, "E-assessment," in *Contemporary Perspectives in E-Learning Research*. Evanston, IL, USA: Routledge, 2006, pp. 165–177.
- [39] M. Simonetti, D. Perri, N. Amato, and O. Gervasi, "Teaching math with the help of virtual reality," in *Proc. Int. Conf. Comput. Sci. Appl.*, Cham, Switzerland: Springer, 2020, pp. 799–809.
- [40] E. Sorensen, "Implementation and Student perceptions of e-assessment in a chemical engineering module," *Eur. J. Eng. Educ.*, vol. 38, no. 2, pp. 172–185, May 2013.
- [41] C. St-Onge, K. Ouellet, S. Lakhal, T. Dubé, and M. Marceau, "COVID-19 as the tipping point for integrating e-assessment in higher education practices," *Brit. J. Educ. Technol.*, pp. 1–18, 2021, doi: 10.1111/bjet.13169.
- [42] T.-H. Wang, "Developing a web-based assessment system for evaluating examinee's understanding of the procedure of scientific experiments," *EURASIA J. Math., Sci. Technol. Educ.*, vol. 14, no. 5, pp. 1791–1801, Feb. 2018.



**OSVALDO GERVASI** (Senior Member, IEEE) received the Ph.D. degree. He is an Associate Professor at the Department of Mathematics and Computer Science, University of Perugia. He teaches networks architecture and virtual reality systems for the bachelor's students in computer science, and the distributed systems and high-performance computing and human-computer interaction for the master's students in computer science "intelligent and mobile computing" at the University

of Perugia. He is the Project Leader of the open-source e-assessment platform LibreEOL (<https://www.libreeol.org/info>). He has published more than 100 papers on scientific journals and peer-reviewed international conferences proceedings. He has edited 80 books of international conference proceedings. His research interests include parallel and distributed computing, cloud and high performance computing, GPGPU computing, artificial intelligence, virtual and augmented reality, and e-Learning. He is a Senior Member of ACM and a member of the Internet Society. He was invited seven times to deliver a plenary lecture at international conferences. He was the President of the Open-Source Competence Centre (CCOS) of the Umbria Region, Italy, established under the Regional Law 11/2006, from 2007 to 2013. He is the President of the not-for-profit organization ICCSA, which manage yearly the International Conference on Computational Science and Its Applications (<http://www.iccsa.org>). He served on the Board of Directors of The Document Foundation (TDF), from 2016 to 2020, which releases the popular open-source Office Suite and LibreOffice.



**DAMIANO PERRI** received the joint B.Sc. and M.Sc. (*summa cum laude*) degrees from the Departments of Mathematics and Computer Science of the University of Florence and the University of Perugia, where he is currently pursuing the joint Ph.D. degree in computer science under the joint Ph.D. Program in mathematics, statistics, and computer science. He is an expert developer of synthetic environments and their applications in everyday life. Since 2016, he has been the Technical Manager of the LibreEOL Project (<https://www.libreeol.org/info>). His research interests include GPGPU computing, parallel and distributed environments, virtual and augmented reality, artificial intelligence, and neural networks. In particular, his recent research interest includes investigating predictive models in order to optimize the most important operators in emerging convolutional neural networks (CNNs).



**MARCO SIMONETTI** received the degree in electronic engineering from the University of Bologna, with a thesis on chaotic carrier transmission systems. He is currently pursuing the joint Ph.D. degree in computer science with the Departments of Mathematics and Computer Science of the University of Florence and the University of Perugia, through the joint Ph.D. Program in mathematics, statistics, and computer science. He worked in the field of clinical engineering for ten years and participated with several teams in the implementation of various automated decision-making systems for health administrations. He works as a High School Mathematics Teacher. His research interests include deep learning and neural networks, quantum computing, and mathematics education.

...