Erasmus+

ΙΔΡΥΜΑ ΚΡΑΤΙΚΩΝ ΥΠΟΤΡΟΦΙΩΝ IKY

CAT SL

Computer-assisted teaching of Sign Language using Computer Vision and Machine Learning (CAT-SL)

Ref. Project: 2020-1-EL01-KA203-079232

Intellectual Output 2: Automated Sign Language teaching System and Service for European SLs

MAIN AUTHORS: K. Karampidis, G. Papadourakis D. Kosmopoulos, K. Antzakas, K. Konstantinopoulos, V. Megalooikonomou, O. Golfi, K. Diakogiorgi, B. Elsendoorn, S. Chatzis, P. Paramithiotis, N. Escudeiro, P. Escudeiro

This project has received funding from Call 2020 Round 1 KA2 - Cooperation for innovation and the exchange of good practices, KA203 – Strategic Partnerships for higher education grant agreement 2020-1-EL01-KA203-079232.

Document history

Date	Name	Affiliation	Position/Project Role	Action/ Short Description
31/6/2023	D. Kosmopoulos	UPATRAS	Technical Manager	Review of document template
17/10/2022	K. Karampidis	HMU	Content supervisor for HMU	Collection of contributions and initial version

Table of contents

Executive Summary

The CAT-SL project aimed to create an automated system for teaching sign languages, focusing on Greek, Dutch, and Portuguese. The system combines computer vision, machine learning, and feedback mechanisms to facilitate the learning process. The project's primary goal was to overcome the shortage of sign language tutors in higher education institutes and special education schools, thus contributing towards a more inclusive society.

The CAT-SL system was developed based on the specific requirements captured in the O1: Analysis of Sign Language Teaching Methods and Challenges for an Automated Teaching System. It includes modules for system administration, course administration, content management, evaluation, body tracking, and sign learning & avatar for visual feedback. The system is expected to facilitate the teaching process of sign language courses in Higher Education institutes, in special education schools, and for individuals interested in learning or mastering their sign language skills.

The report provides extensive details on the project's methodology, which involved a modular development and integration approach. The different components of the system were developed separately and then integrated into a unified educational platform using Moodle's capabilities for course delivery and student engagement.

The report also outlines the different functional and non-functional implementations of the system, details the use cases for different user groups, and describes the various evaluation criteria used to assess the system's effectiveness.

Overall, the project has resulted in a robust automated system that can facilitate sign language learning and contribute towards narrowing the communication and cultural gap with the Deaf community. Future steps include expanding the system to cover more sign languages and developing specialized courseware for teaching foreign sign languages to deaf users.

List of abbreviations, acronyms and definitions

Abbreviation / Acronym	Definition
AI	Artificial Intelligence
API	Application Programming Interface
CERF	European Framework of Reference for Languages
CSL	Cypriot Sign Language
DSL	Dutch Sign Language
ECTS	European Credit Transfer and Accumulation System
ELP	European Language Portfolio
EU	European Union
GSL	Greek Sign Language
JSON	JavaScript Object Notation
LGP	Portuguese Sign Language
MT	Machine Translation
NLP	Natural Language Processing
PHP	Hypertext Preprocessor
SL	Sign Language
UI	User Interface
WebRTC	Web Real-Time Communication

List of figures

List of Tables

1. Introduction

Many educators need to receive appropriate training in Sign Languages, the native languages of the Deaf, so that they can communicate with deaf students and to teach them how to develop linguistically. Sign Languages in the curricula of Primary Education University Departments or Departments dealing with Special Education are very useful and contribute towards this goal. However, it is hard for Higher Education Institutes to find and employ tutors as experts in Sign Language, so many educators are not adequately trained as students to face the challenges of communicating with deaf children in the classroom. Similar considerations apply for teachers of SLs in special education schools.

Here we present the CAT-SL Automated Sign Language Teaching System for European SLs. The system implements a server-based learning service. A stand-alone system is available as well for experimentation with the signing, independently of the course pipeline. It is the first such service that we know of for learners of SLs, that combines computer vision, machine learning and feedback mechanisms for EU SLs and especially for Greek, Dutch and Portuguese. The development was based on the open-source platform of Moodle [Moodle].

Our development was based on the analysis of the requirements as they were captured in the O1: Analysis of Sign Language Teaching Methods and Challenges for an Automated Teaching System. One-by-one the user requirements were mapped to design and implementation. The design implementations were mapped to software modules. Afterwards, the software modules were integrated into the system to offer an educational service for learning SLs. The user is able to access the full service by connecting a color web camera.

The system is expected to facilitate the teaching process of SL courses in Higher Education institutes, in special education schools, but also to help individuals interested in learning or mastering their SL skills. Today it is almost impossible to find a tutor in a SL different from the official local SL, however this can be partially handled by CAT-SL, since it is able to cover multiple languages. In the long term, this is expected to narrow the communication and cultural gap with the Deaf community and contribute towards a more inclusive society. The approach can be extended to more SLs in the EU and globally.

The service was implemented for a small number of languages (Dutch, Greek, Portuguese). However, the same methodology is applicable to several more languages in Europe and beyond. This can be achieved by focusing on a common approach, without making our approach dependable on a specific language. Furthermore, through the CAT-SL paradigm it is possible to author more courses in various educational levels and for various proficiency levels in each SL.

The document is structured as follows: Following the introduction, we explain the methodology we used in developing the CAT-SL Automated Sign Language Teaching System. This includes an overview of the modular design approach, the roles of each project partner, and the key technologies used. The next section provides a detailed breakdown of the functional and non-functional implementations of the system. We also offer a comprehensive description of the system's features and capabilities, including system administration, content management, and evaluation. We then discuss the specifics of our sign recognition technology, discussing how we use machine learning and computer vision to analyze and score sign language movements.

Next, we provide a detailed overview of the use cases of the CAT-SL system, illustrating how

the platform can be used by different user roles. We then discuss our stand-alone practice application and go into detail on how we utilize the VirtualSign avatar for visual feedback.

Finally, we discuss the evaluation of our system. We provide insights into the functionality and user-friendliness, robustness, assessment of learning effects, type of feedback provided to students, and how CAT-SL can be used by Deaf users. We conclude with a summary of our findings and suggestions for future steps.

2. Methodology

The methodology follows a modular development and integration approach, which was coordinated among the project partners. The project employs a server-client architecture, where the Moodle platform serves as the client, and dedicated servers process video inputs for hand sign recognition. The project tasks span four main areas: Moodle module development, video processing, gesture scoring, and content creation.

The contributions of the partners can be summarized as follows: UPATRAS focused on hand and body tracking methods to identify and classify hand and body gestures. CUT developed machine learning algorithms for understanding sign languages. IPP contributed an avatar for sign language using Unity3D. The HMU worked on the overall service application and user interface, integrated into Moodle using PHP and JavaScript. Kentalis was responsible for user validation through methods such as surveys and usability testing.

In the development of the CAT-SL Automated Sign Language Teaching System, the first task was to develop individual modules, with each partner contributing to specific areas of expertise as detailed above. Following this, all the developed modules were integrated into a unified educational platform, using Moodle's capabilities for course delivery and student engagement.

For the backend, the project uses Moodle as the primary Learning Management System, taking advantage its open-source capabilities and robust community support. The RecordRTC plugin, developed by HMU, integrates directly into Moodle's Quiz and QuestionType subsystems. This arrangement enables a seamless interface with the specialized scoring algorithms hosted on UPAT servers. The choice of Moodle is based on its modular architecture, scalability, and extensive user base, which make it a reliable platform for our specialized application.

Once the initial version of the platform was ready, a series of tests were conducted to validate the recognition and modeling results. These tests ensured that the machine learning algorithms accurately identified and scored the signs, offering a means of student assessment. Parallel to that, the validity of the feedback mechanism was also tested to ensure that users received accurate, near real-time evaluations of their performance.

Next, comprehensive usability tests were conducted to assess the overall service. These tests tested the system's intuitiveness and ease of use, employing metrics such as user satisfaction and task completion rates. The final stage of development involved evaluations by real users to obtain valuable feedback and insights into any areas requiring improvement.

The methodology employed an iterative approach. After each phase, whether it was module development, integration, or testing, the system underwent further refinement. This iterative process allowed for the continual improvement of the platform, making adjustments and

enhancements based on both quantitative metrics and qualitative feedback.

The recognition part concerns the abilities of machine learning systems to represent and identify the motions performed by hands or other body parts, in order to facilitate feedback to the students.

In our solution for sign recognition, we integrate the MediaPipe framework, specifically leveraging its Hand and Pose modules. These modules constitute an optimized pipeline that facilitate the detection and tracking of hand and pose motions, both vital for our sign recognition system.

MediaPipe is a general open-source framework by Google, which serves as a platform for building and implementing machine learning pipelines. The Hand and Pose modules are part of this framework and particularly notable for their application in real-time perception tasks. The Hand module is designed to identify and track 21 hand landmarks across frames in a video sequence, a crucial task for hand gesture recognition. The Pose module is designed to estimate body poses, identifying 33 body landmarks including the position of the head, shoulders, elbows, hips, etc.

In the context of computer vision and machine learning, landmarks refer to specific points of interest within an image or a sequence of images, such as a video stream. These points are typically associated with particular anatomical or structural features, such as the corners of the eyes, the tip of the nose, or joints like elbows and knees in the case of pose landmarks. By accurately detecting and tracking these landmarks, algorithms can interpret the pose, movement, and gestures of humans which in turn facilitates motion analysis and gesture recognition. Landmarks are represented by coordinates of the 3D space (x, y, z).

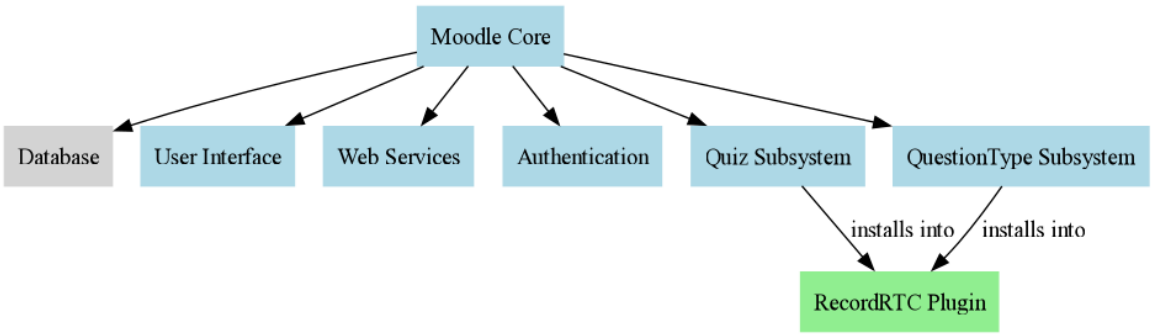


Figure 1 The Moodle architecture and the RecordRTC Plugin

The CAT-SL project is integrated into Moodle through the RecordRTC plugin, developed by the HMU. This plugin ties into Moodle's quiz and questionType subsystems, enabling users to answer questions via video. Once the video answer is uploaded, a separate API, co-developed by HMU and UPAT, handles synchronization and grading. This API compares the student's answer with a template and updates the Moodle subsystems accordingly. The choice of Moodle as the backbone for the CAT-SL project stems from its scalability, extensibility, and strong community support. Its open-source nature ensures flexibility, while its well-established plugin architecture makes it amenable for custom developments like the RecordRTC plugin.

The underlying technology of these modules uses advanced machine learning algorithms and models. The Hand module uses a palm detector followed by a hand landmark model to

recognize the 21 hand landmarks. The Pose module uses a pose landmark model to identify the positions and orientations of the 33 body landmarks.

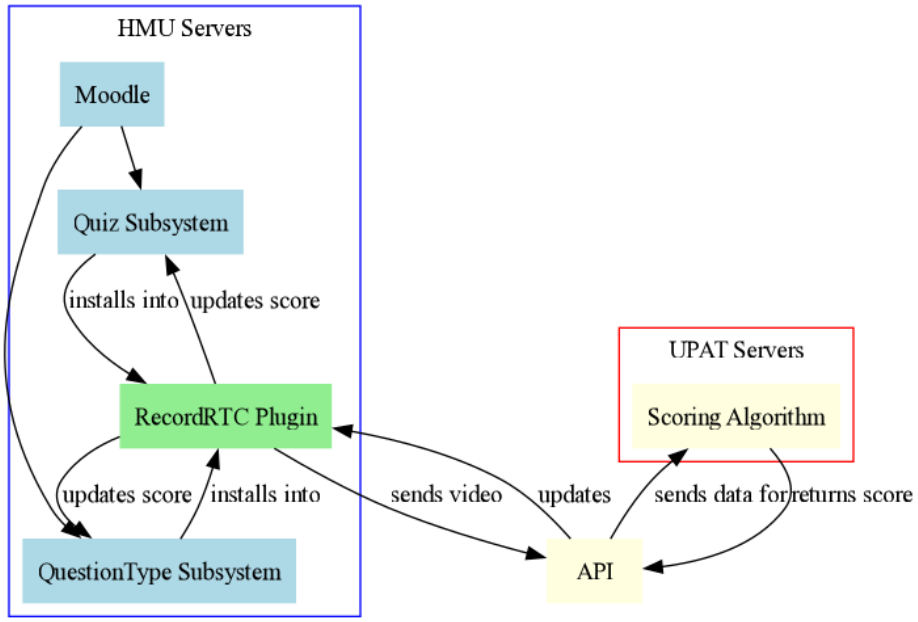


Figure 2 Architecture of the system communication between the plugin and the recognition servers

Figure 3 Pipeline of the scoring procedure

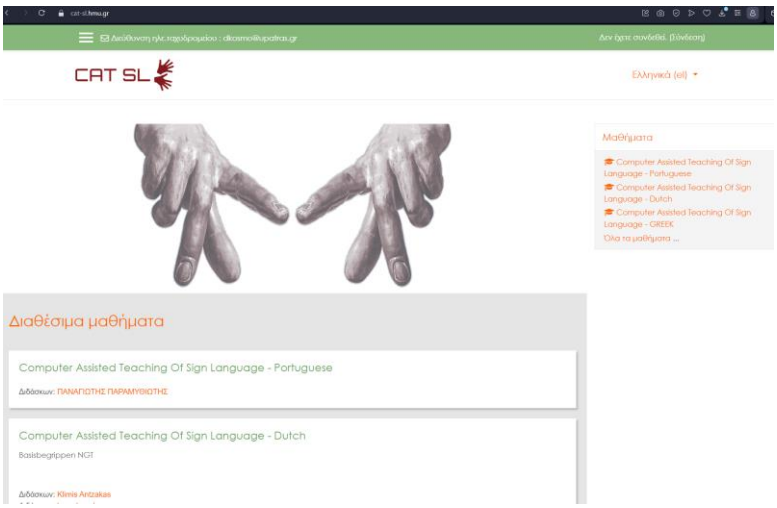


Figure 4 The e-learning system web page, accessible at https://cat-sl.hmu.gr/

3. Functional Implementations

The CAT-SL system contains modules for: System administration, Course administration, Content management, Evaluation, Body Tracking, Sign Learning & Avatar for visual feedback.

The related modules were developed by the consortium after reviewing the related literature (i.e., linguistic implementations by each of the involved SLs) and the technical capabilities offered by state-of-the-art technology. These modules were validated by the users. The implementations are described in the following by presenting the requirements as defined in the and the respective menus in the platform.

3.1. System Administration

The system administration (SA) includes a full range of functions for the management and configuration of system parameters and attributes, data, users, courses and interactivity settings. The following basic functions were included, which cover: authentication, management of rights and roles, user management, import and export of users and resources, management of language packs and log, etc.

	Description	Example Use Case
SA1	Supports multiple roles like Admin, Teacher, Student, Guest	Role-based access control
SA2	Enables creation of user groups for collaboration	Study groups, project teams
SA3	Access resources from external websites	Embedding YouTube videos
SA4	Single interface for all admin tools	Centralized control
SA5	Allows disk space quotas for users and courses	Resource management
SA6	Enables user-specific settings based on roles	Bandwidth limitations
SA7	Monitors visits and other usage statistics	Analytics
SA8	Provides event log storage and analysis	Security audits
SA9	Manages interactivity modules (e.g., body tracking)	Module installation

Table 1: The system administrator functions.

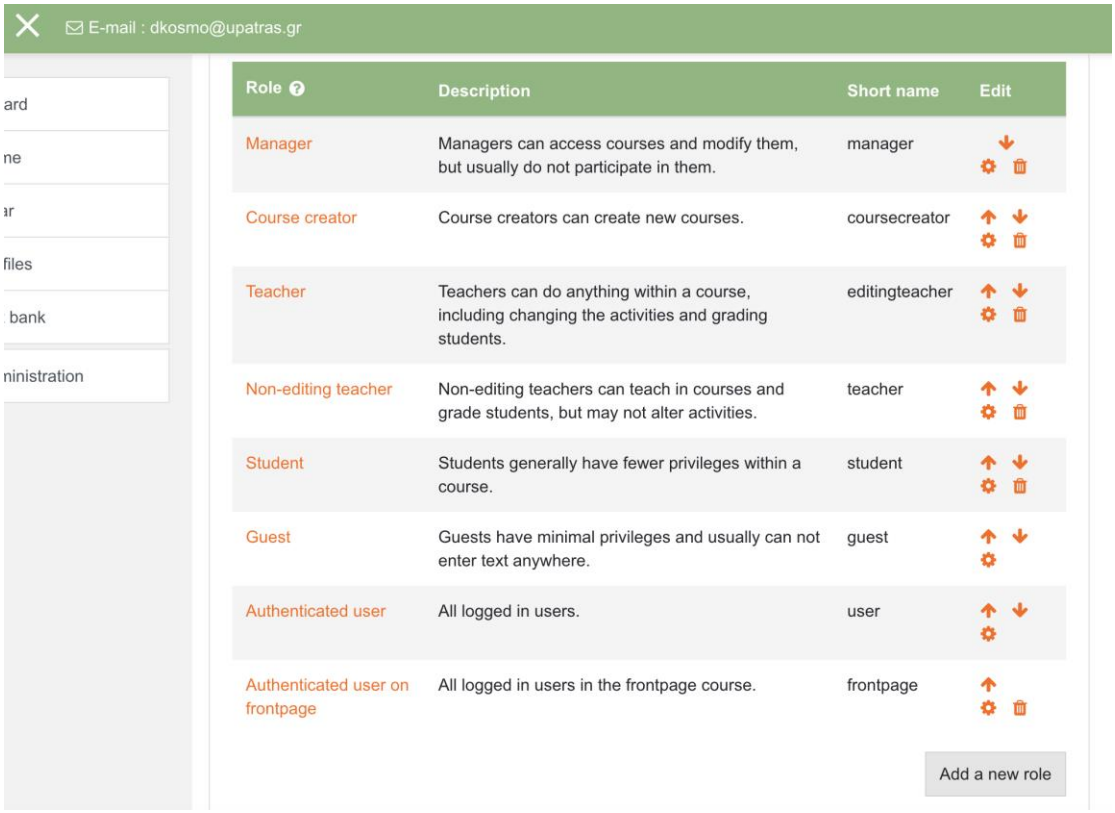


Figure 5 SA1: The system supports several standard roles (e.g., Administrator, Teacher, Student, Guest) and has the potential to create additional roles.

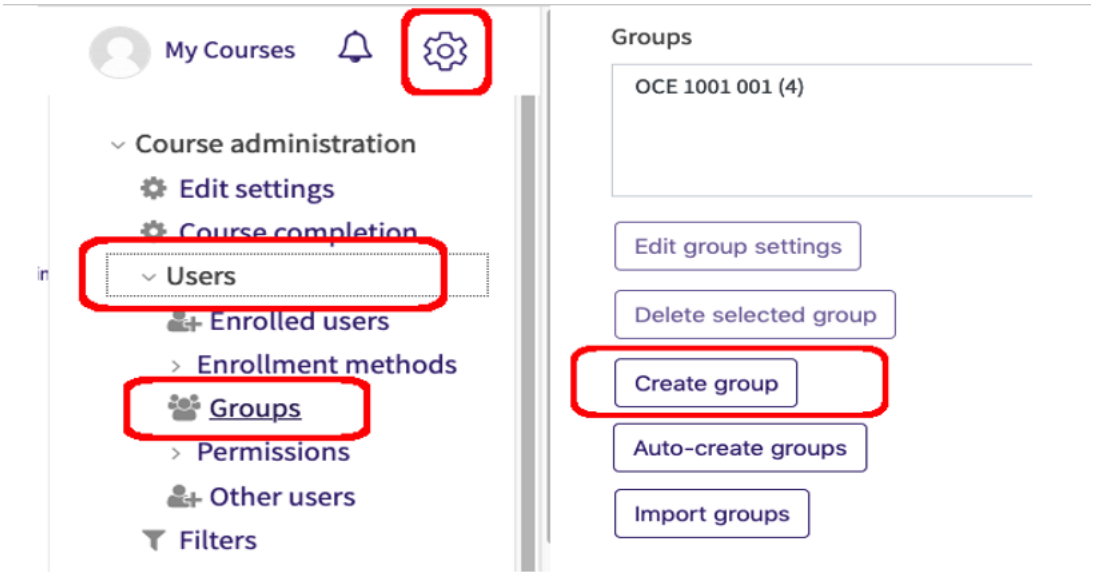
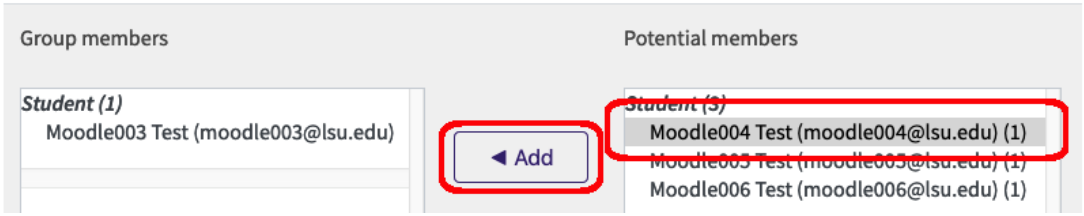


Figure 6 SA2: The system is also able to create user groups to collaborate, communicate and share content. These can also accommodate different groups of users attending different courses.



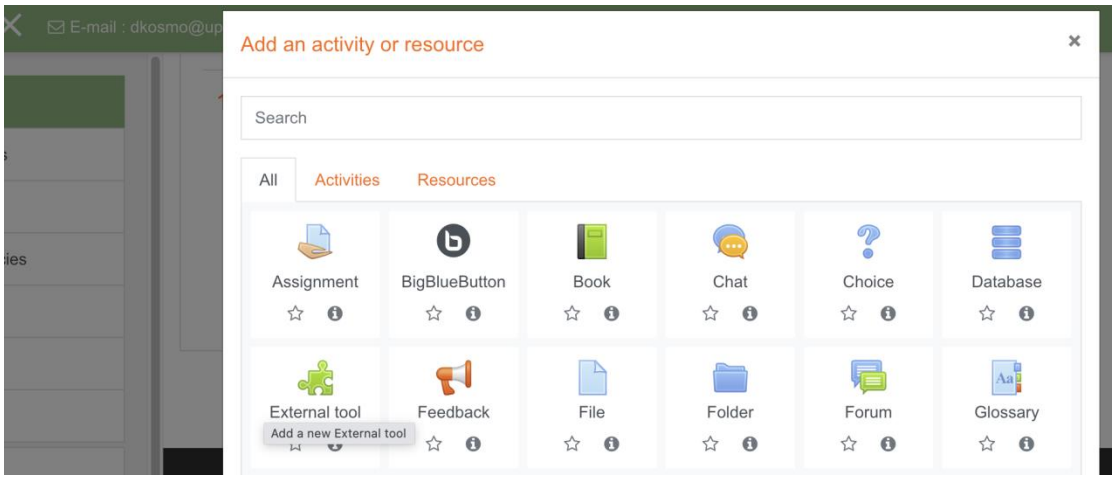


Figure 7 SA3: The platform enables users to access resources from external websites.

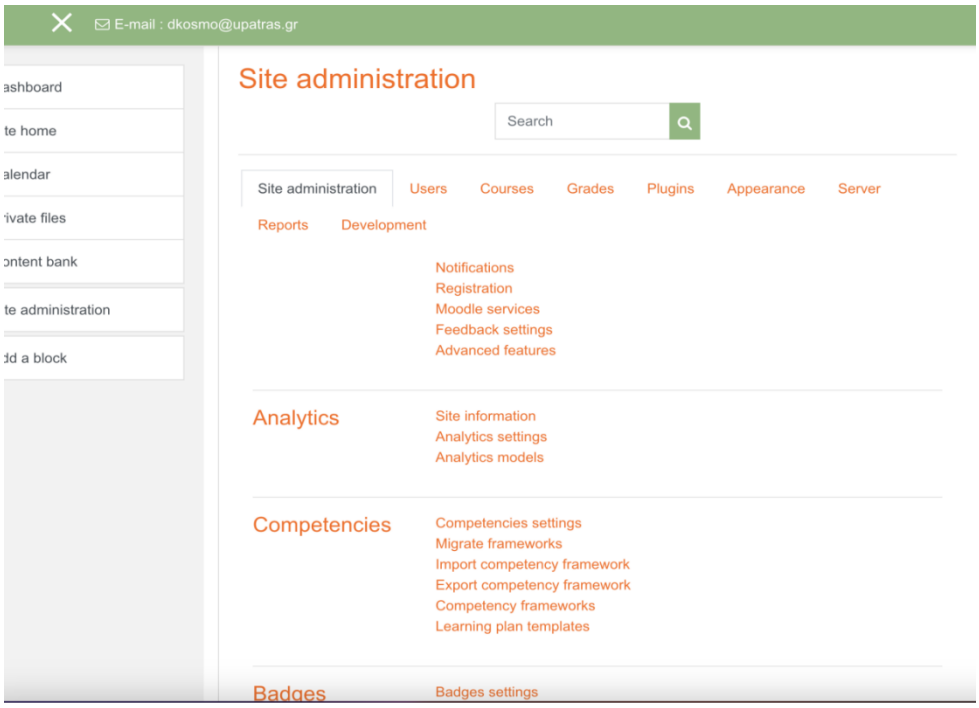


Figure 8 SA4: It is possible to access all administrative tools and functionalities from a single interface.

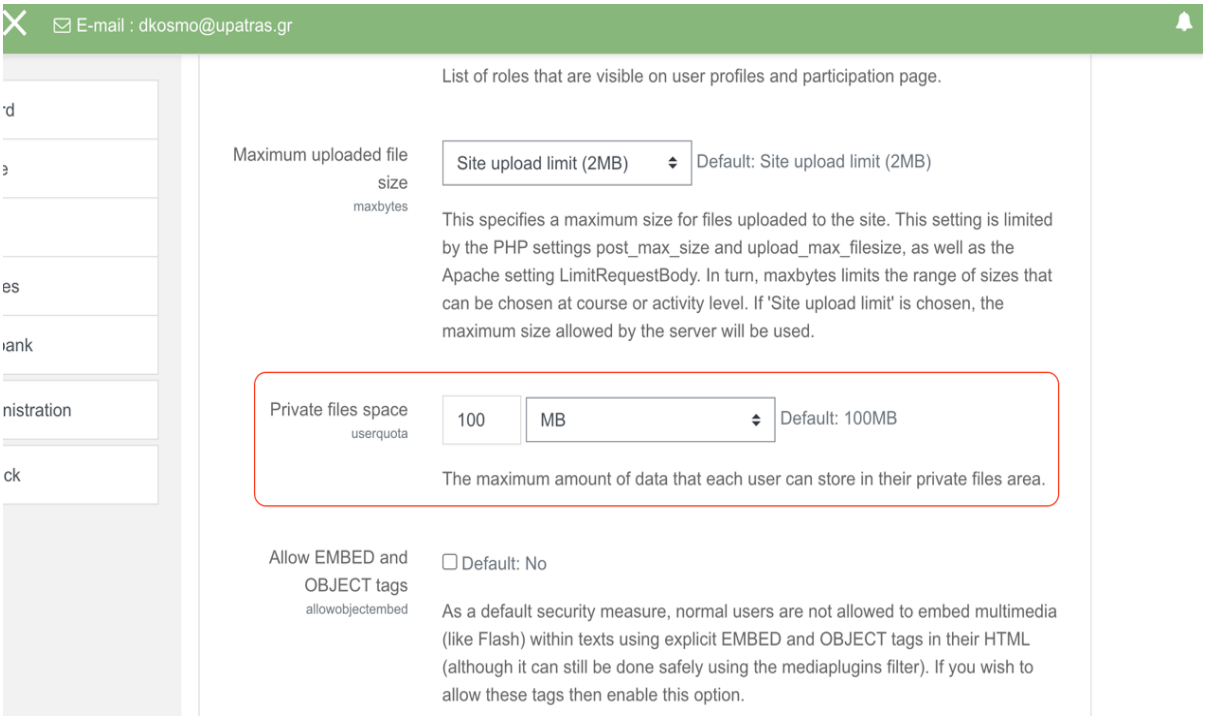


Figure 9 SA5: System Administrators are able to set quotas on the disk space for individual users, courses and organizations.

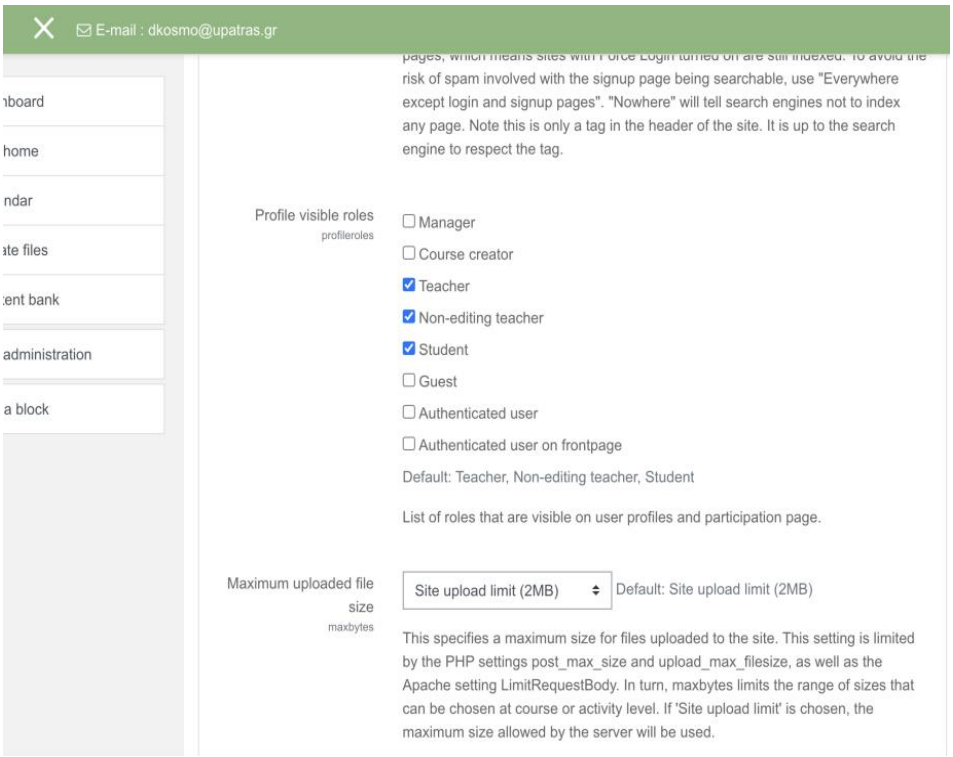


Figure 10 SA6: The administrator is able to set specific settings for the rights of users based on user roles, including settings for bandwidth on e-resource access.

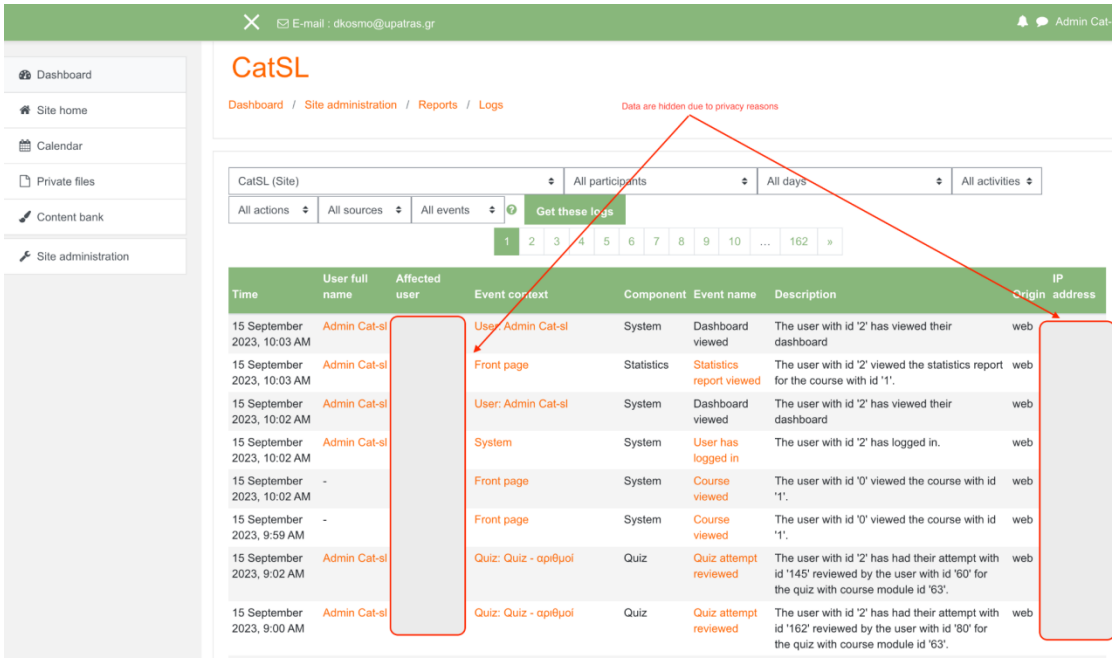


Figure 11 SA7: The administrator is able to monitor visits and other statistics of the platform (i.e. number of users, visit period, etc.).

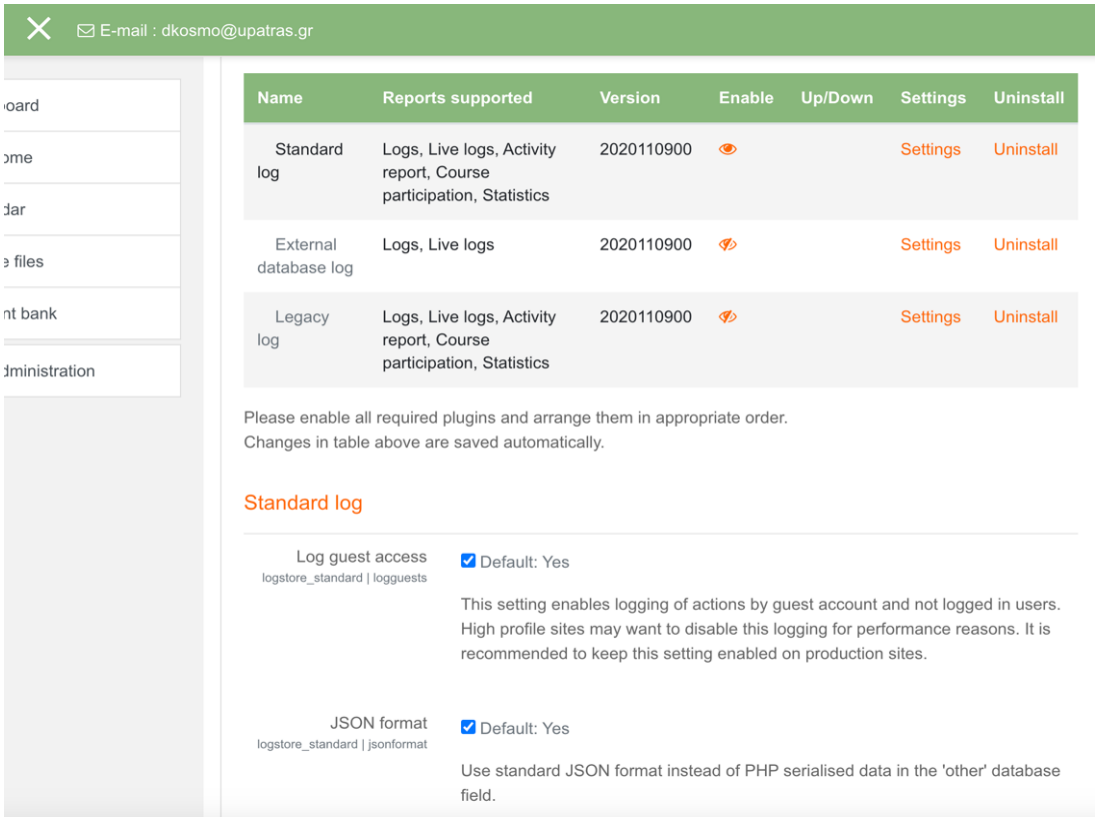


Figure 12 SA8: Event log storage and log analysis functionality for the needs of system administrators.

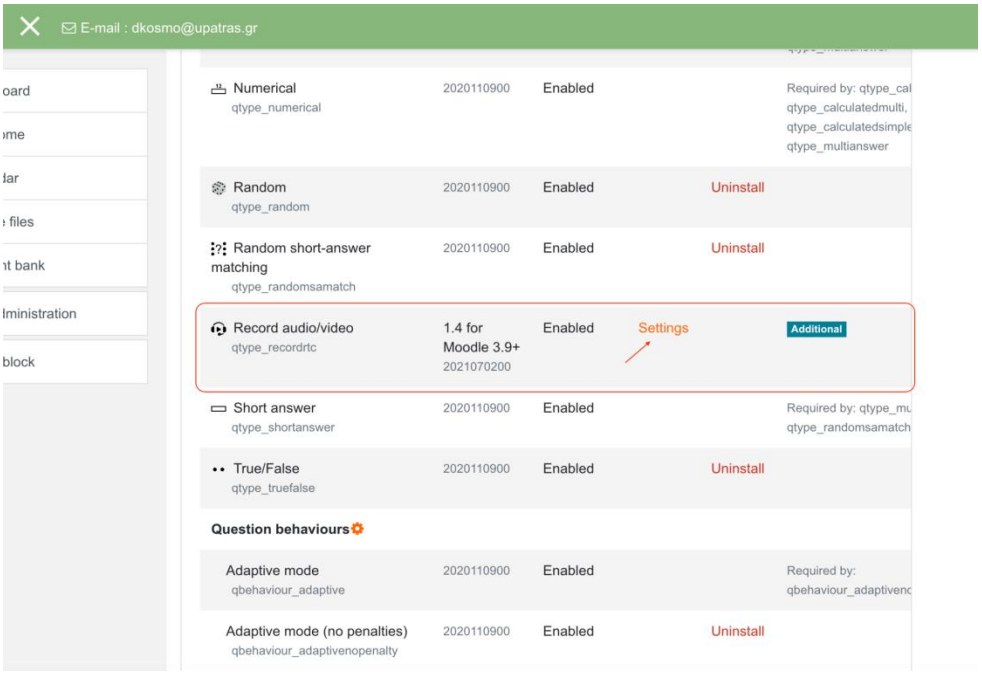


Figure 13 SA9: The administrator is able to install and maintain the interactivity modules (body tracking, sign modeling, avatar).

Figure 14 SA10: The following Language packs are available: Greek, Portuguese, Dutch and English.

3.2. Course Administrator

Course administration (CA) provides tools for synchronous and asynchronous e-learning, creating, editing, saving and deleting e-learning courses, encouraging student participation in the learning process and ensuring better interactivity within the teaching process. It also provides an opportunity to test, assess and oversee the student's performance. It also gives the opportunity for self-assessment.

	Description	Example Use Case
CA1	Assign roles and rights to users	User permission settings
CA2	Generate standard and custom reports	Tracking user activity

	Description	Example Use Case
CA3	Tools for communication and interaction	Calendar, email, chat
CA4	Upload syllabuses with specific access rights	Course outline access
CA5	Create and categorize courses	Course organization
CA6	Embed media content	Multimedia lessons
CA7	Set sequence for accessing educational materials	Structured learning path
CA8	Customize content access based on performance	Adaptive learning
CA9	Time-based availability of course settings	Scheduling
CA10	Choose what to include in course backup	Data backup
CA11	Automatic notifications for activities	Update alerts
CA12	Control group memberships and tools	Collaborative learning
CA13	Publish course-related info	Transparency
CA14	Enable online submissions	Assignment submission
CA15	Conduct user surveys	Feedback collection
CA16	Enable self-assessment options	Interactive self-tests
CA17	Enable anonymous evaluations	Bias-free assessment
CA18	Import/export tests and questionnaires	Reusability
CA19	Create question repositories	Efficient question management
CA20	Set weights for automatic evaluation	Grading customization
CA21	Personal feedback methods	Automated or teacher-provided feedback
CA22	Built-in email system	Direct communication
CA23	Import/export courses with standard support	Cross-platform compatibility
CA24	Preview student view without logging out	Interface testing
CA25	Create personal portfolio of courses	Teaching record
CA26	Create a reference body motion	Motion-based learning

Table 2: The course administration functions.

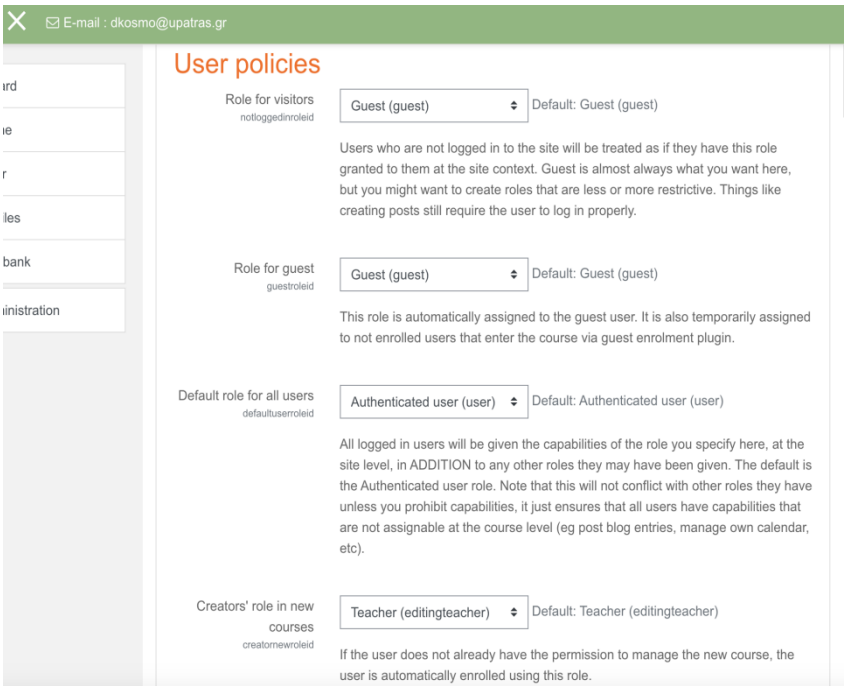


Figure 15 CA1: Course administrators are able to assign different roles and rights to the different users, as well as assign access rights to various e-resources within the system.

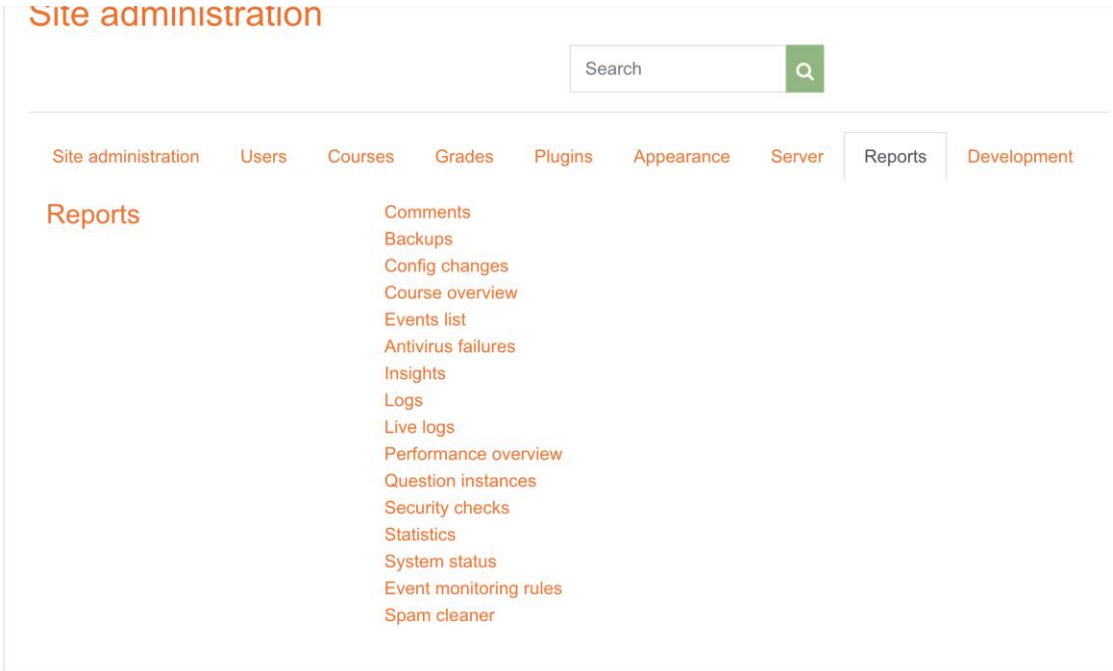


Figure 16 CA2: Course administrators are able to produce standard user activity and system access reports, and create customized reports without the need for additional programming.

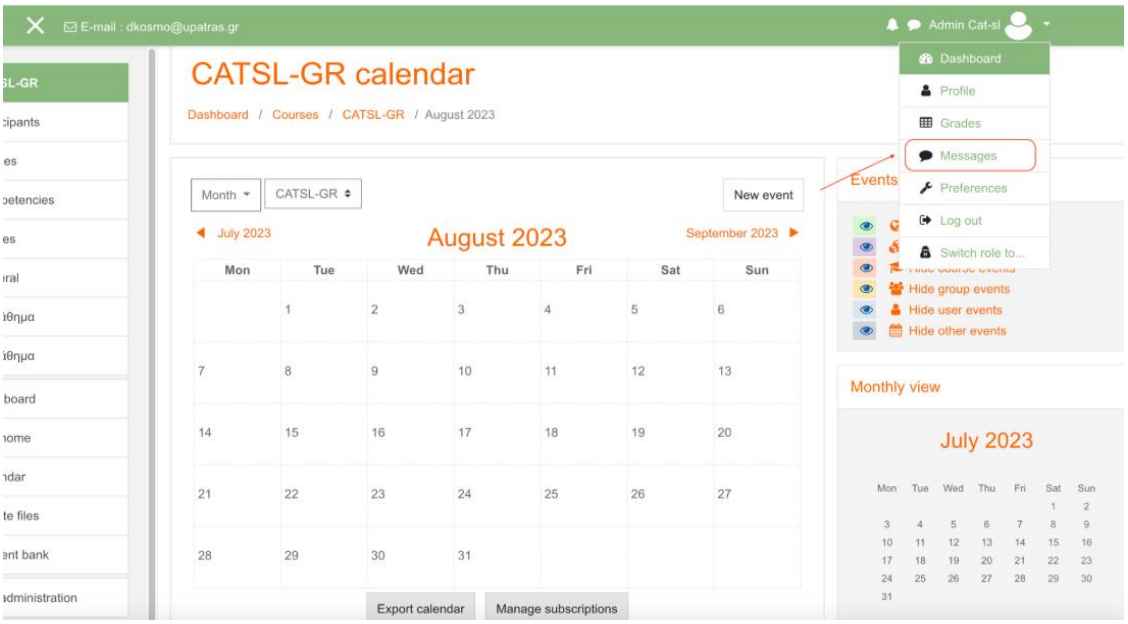


Figure 17 CA3: The system disposes of available tools for communication and interaction such as a calendar, messages and announcements, email service, tasks, and chat.

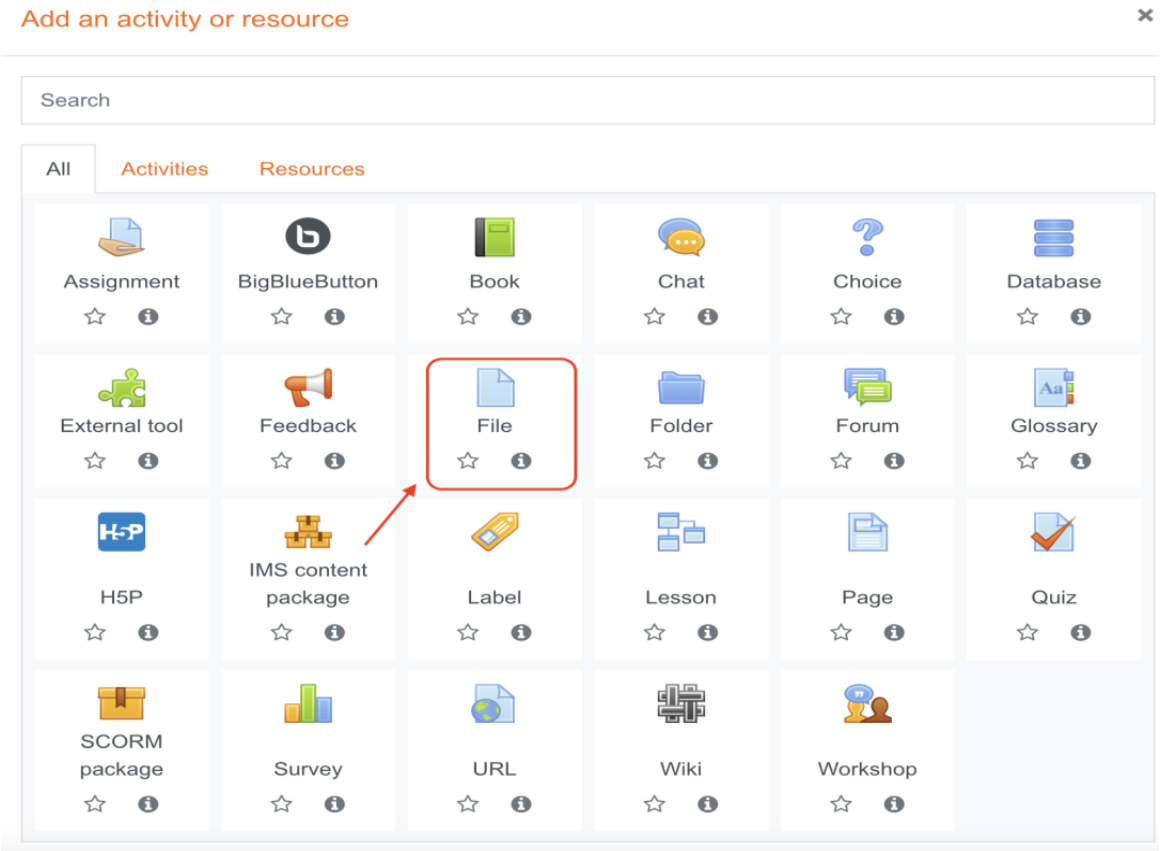


Figure 18 CA4: The system allows the upload of syllabuses that are accessible through students' specific access rights.

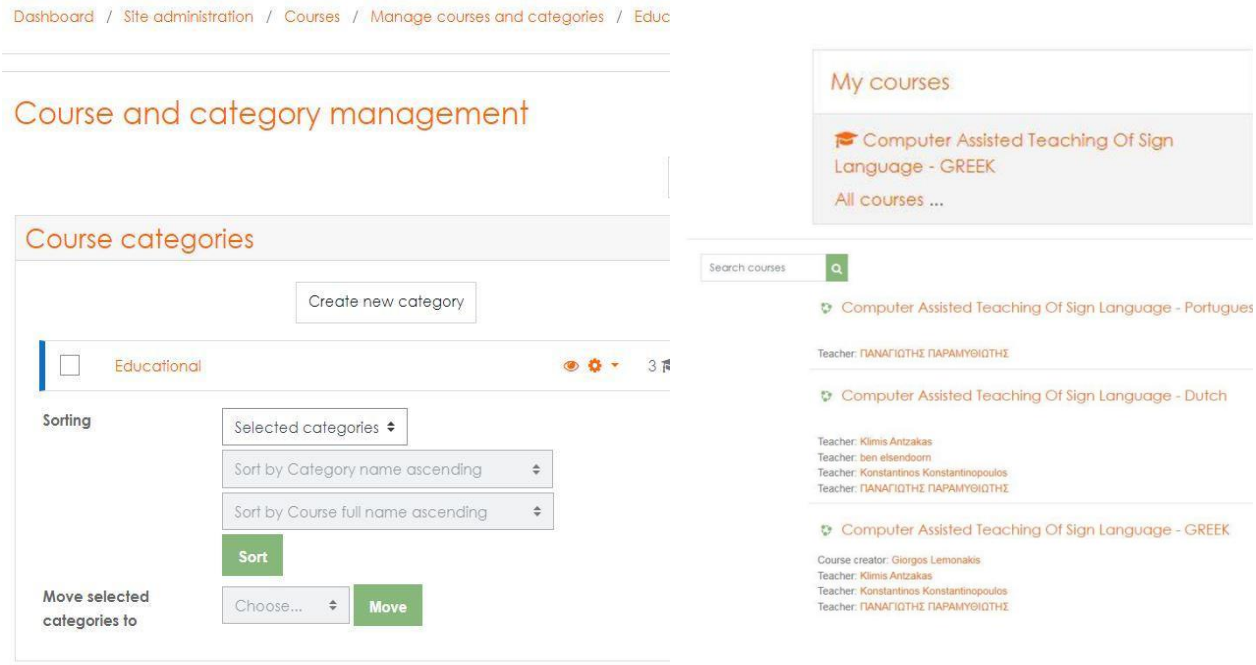


Figure 19 CA5: The system allows to create course categories and to search and view training courses according to specific rights and roles of individual users in the system.

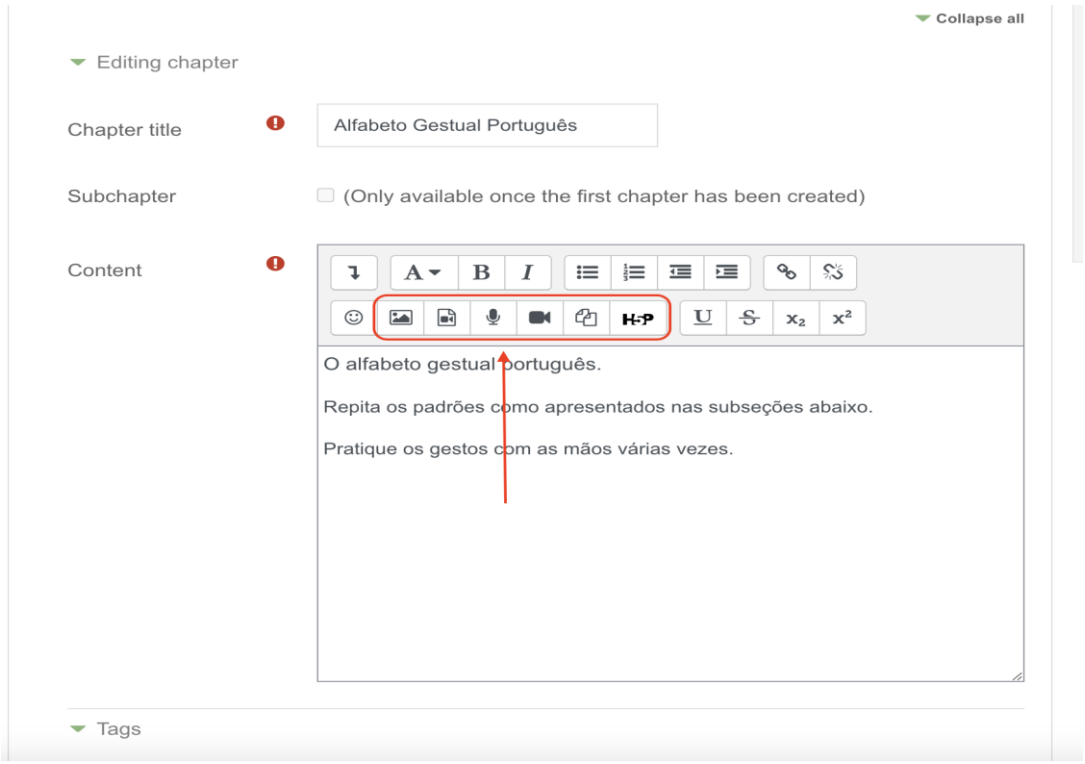


Figure 20 CA6: The system allows to integrate and embed images, presentations, and video content.

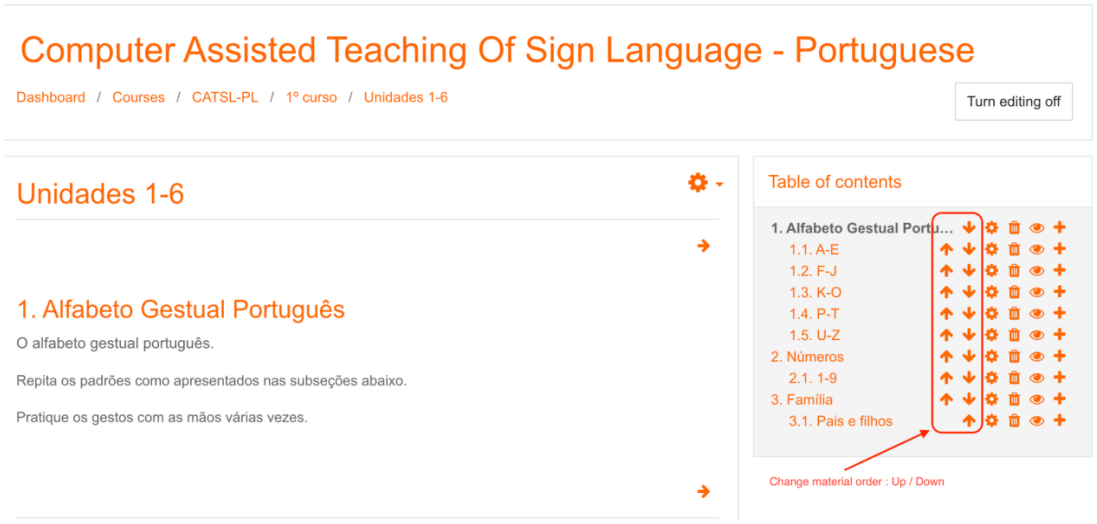


Figure 21 CA7: Course administrators are able to set an order for accessing educational materials with a view to their utilization in a particular sequence.

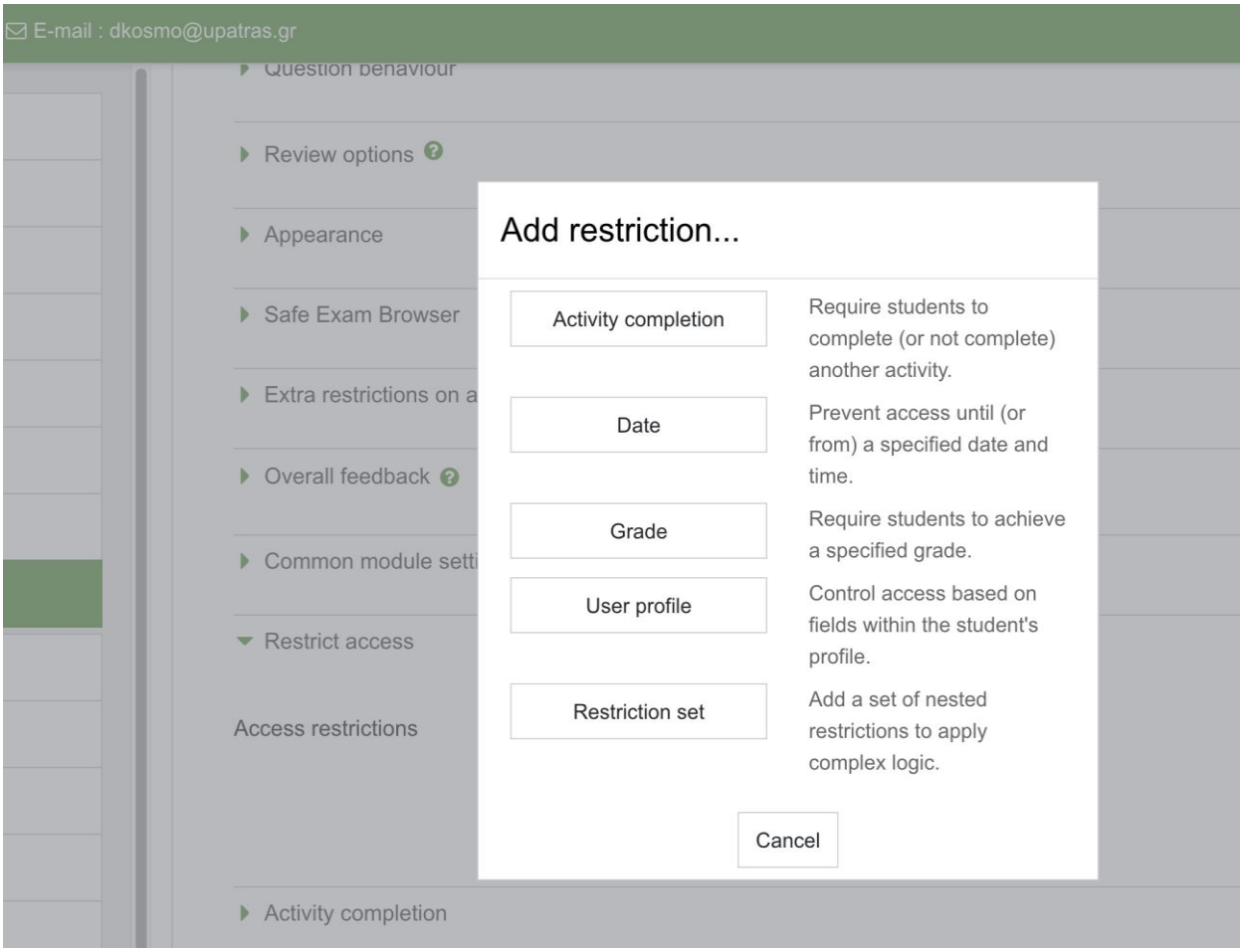


Figure 22 CA8: Course administrators are able to define access to different content, depending on individual performance and student progress.

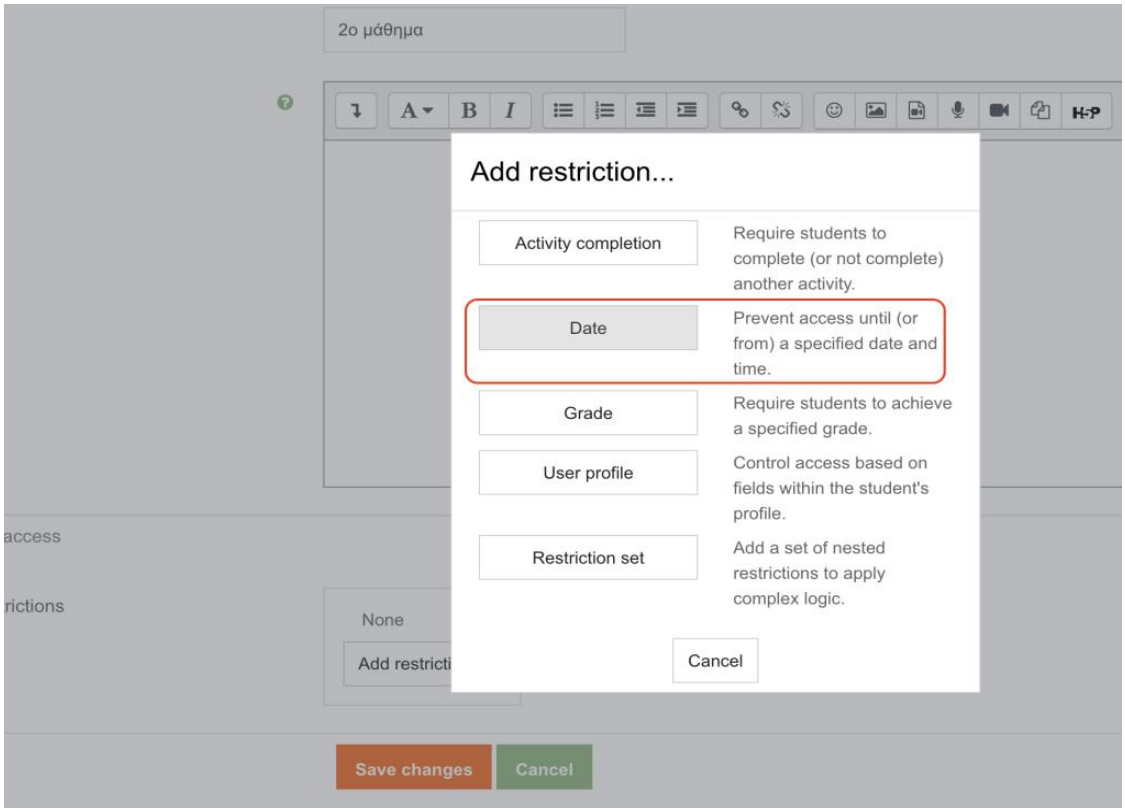


Figure 23 CA9: Course administrators are able to change the course settings and make certain tools and parts of the course available on a specific date and time.

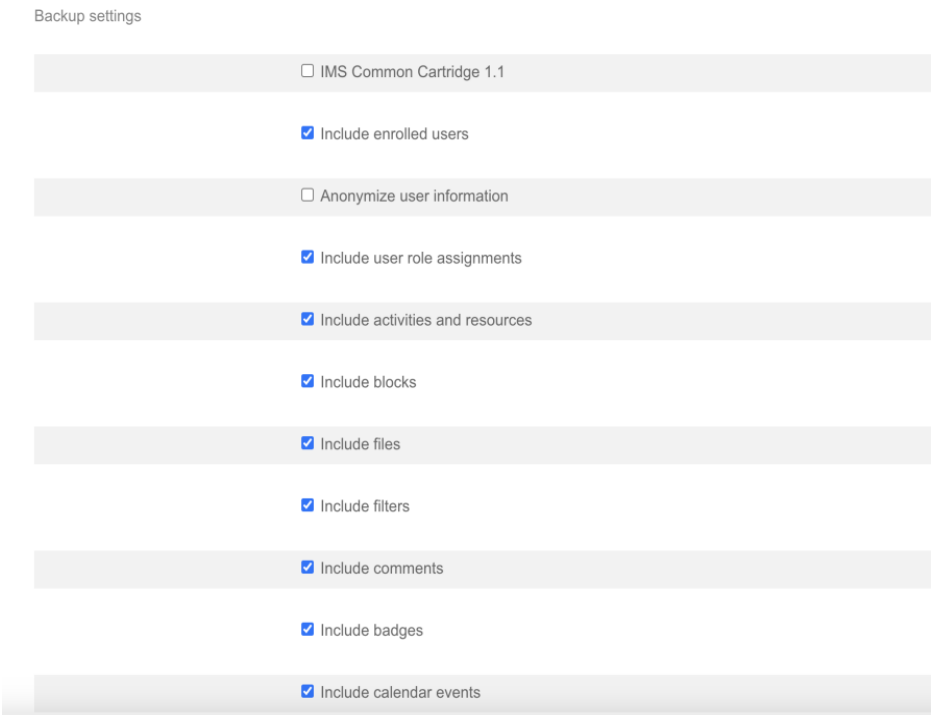


Figure 24 CA10: Course administrators are able to choose what to include on a course backup.

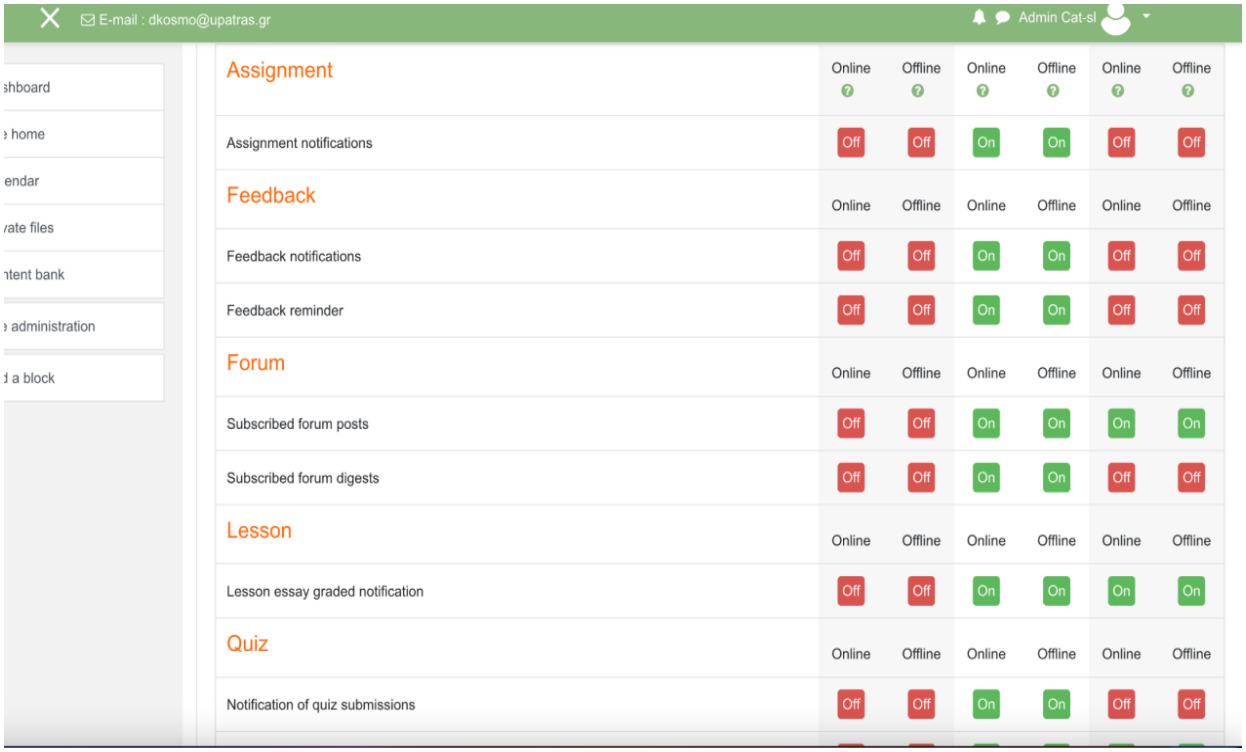


Figure 25 CA11: Course administrators are able to automatically notify users about new activities, publications, assignments, examinations, tests, or changes in the course.

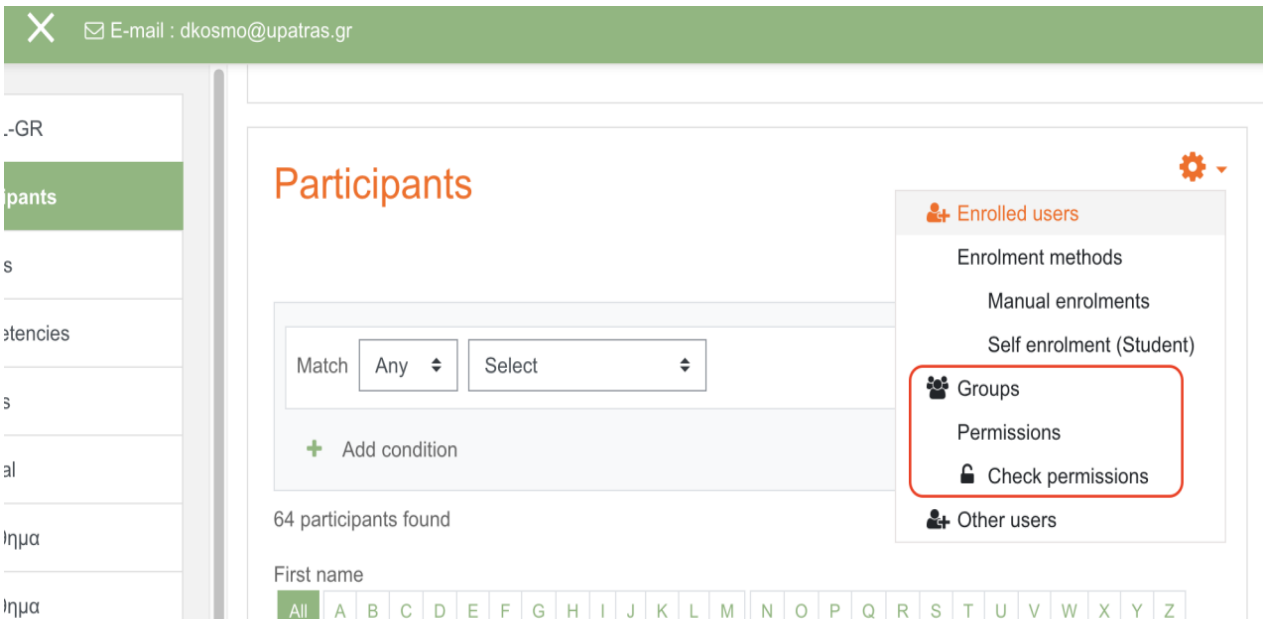


Figure 26 CA12: Course administrators are able to create groups, to control a group's membership by assigning specific rights to users, and to determine what tools are available to certain groups.

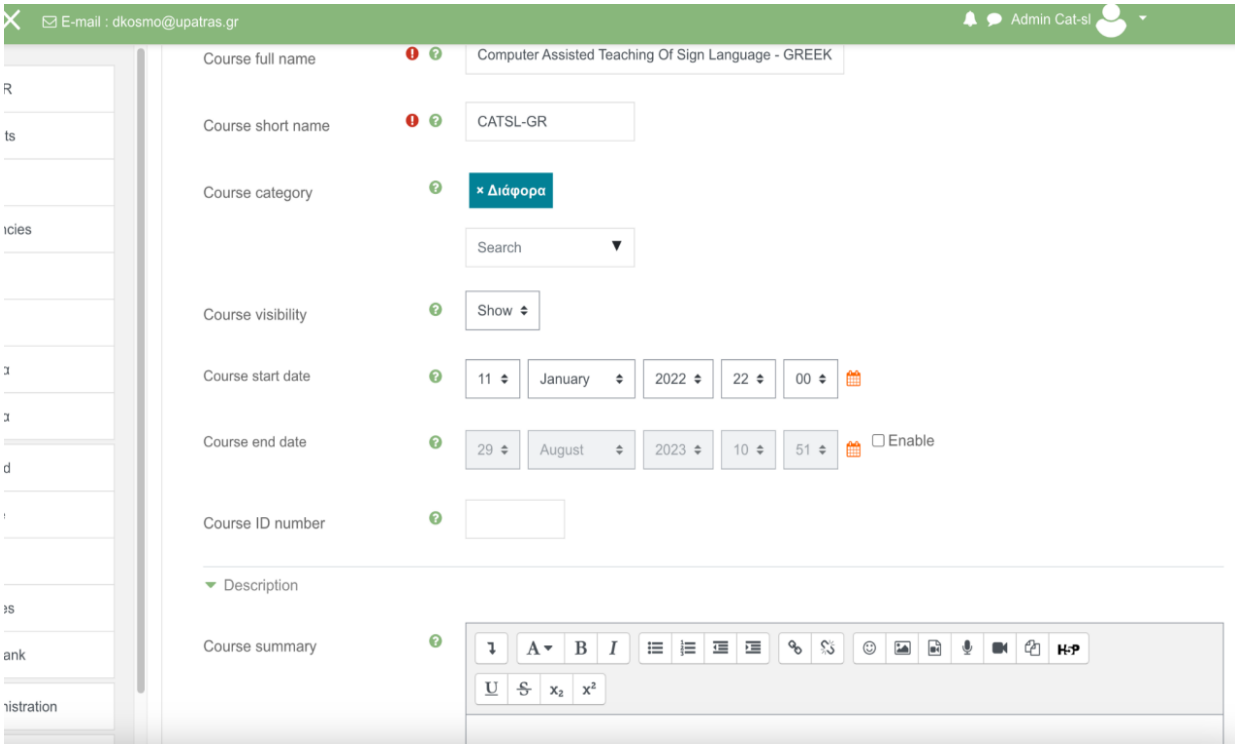


Figure 27 CA13: Course administrators are able to publish information related to the teacher and the course.

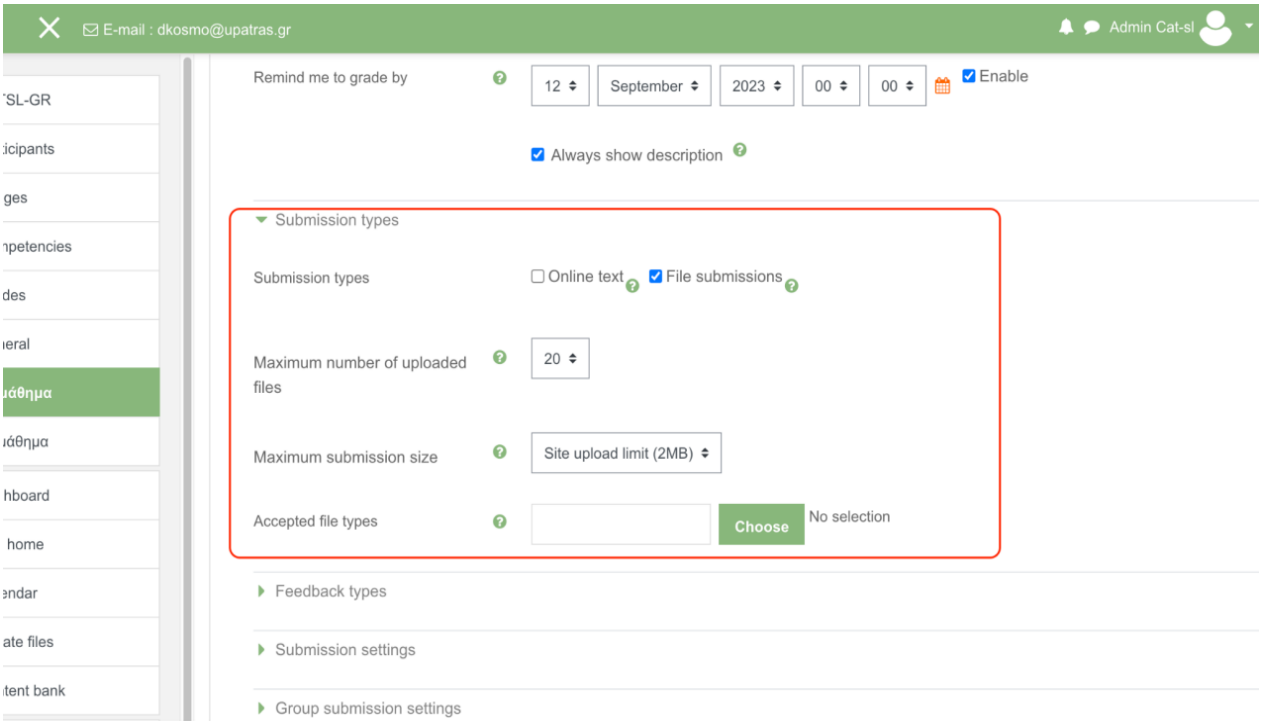


Figure 28 CA14: Course administrators can enable online submission of assignments and tests.

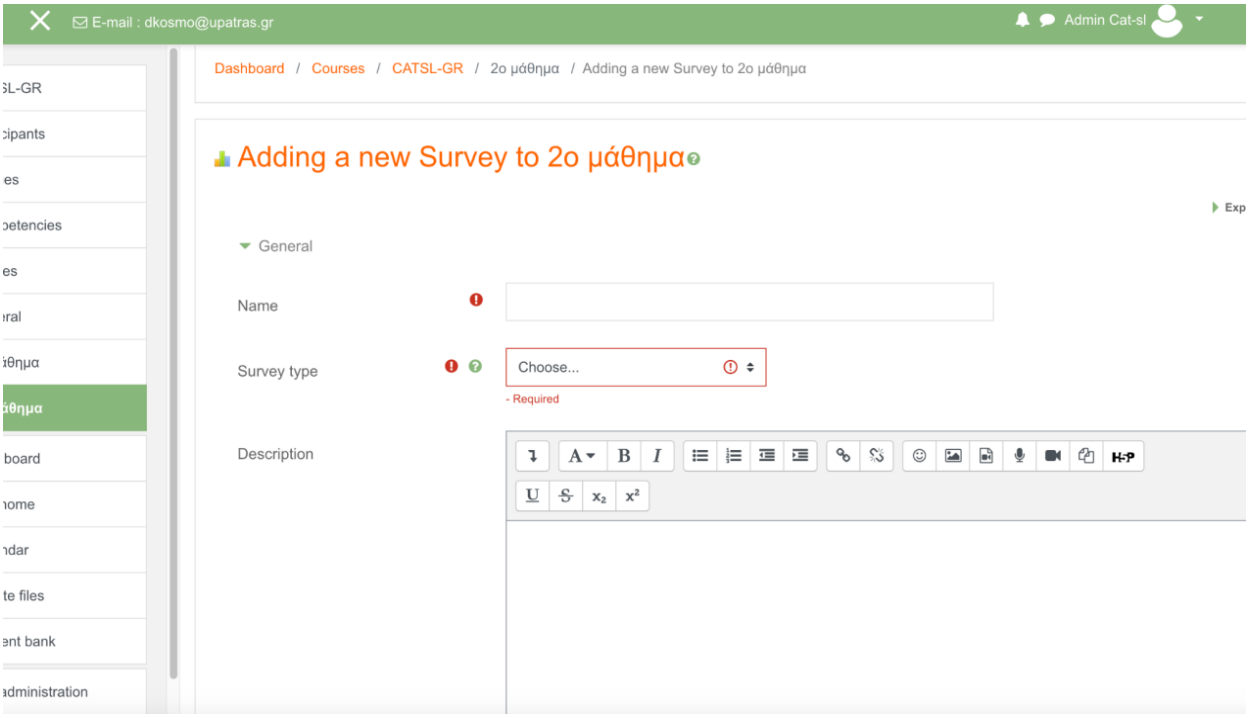


Figure 29 CA15: Course administrators can conduct surveys among users in the system.

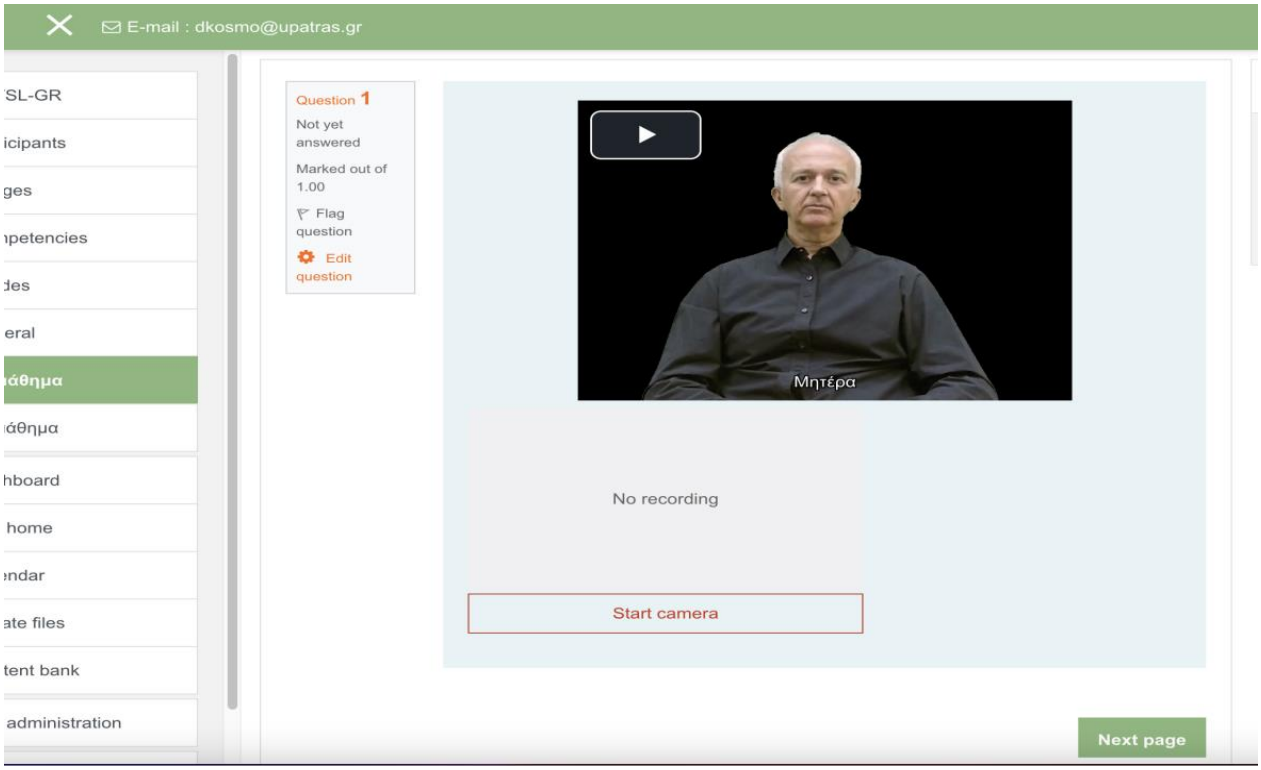


Figure 30 CA16: Course administrators can enable the self-assessment especially for interactive content in SL.

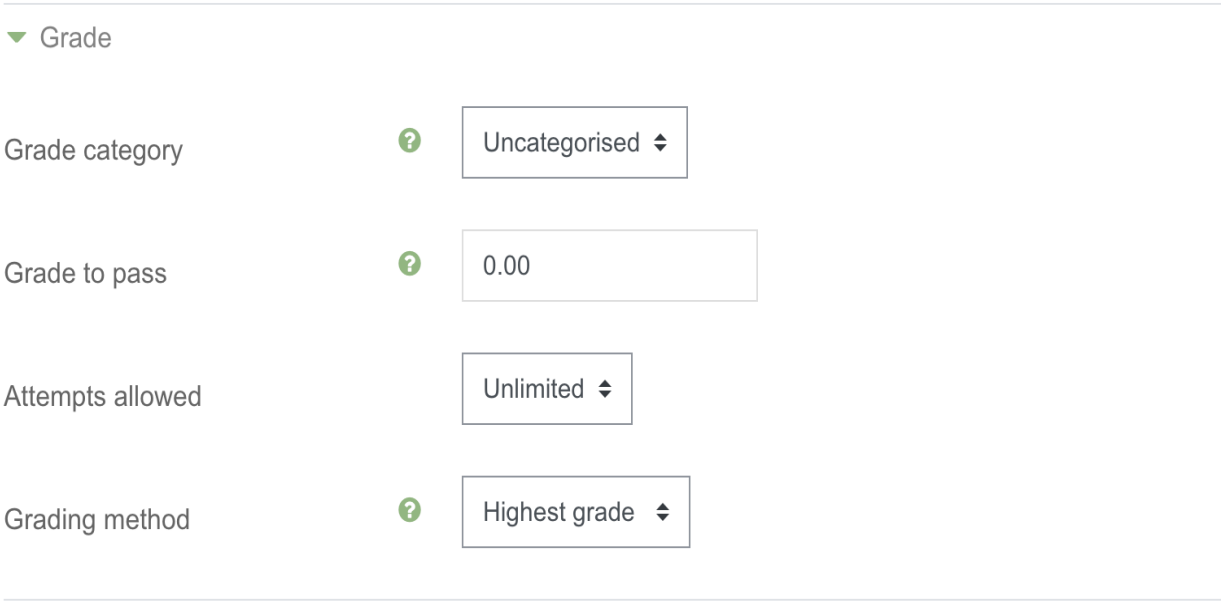


Figure 31 CA17: Course administrators can enable anonymous evaluation of students by the teacher.

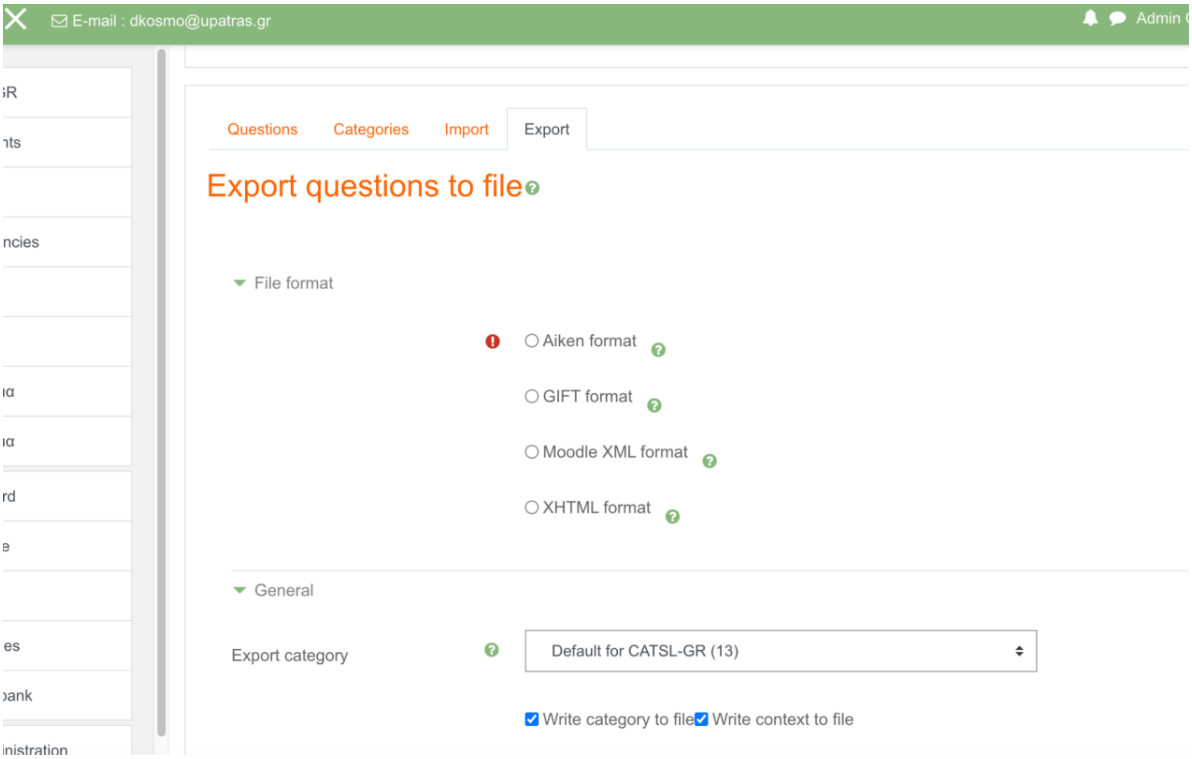


Figure 32 CA18: Course administrators can import and export tests and questionnaires.

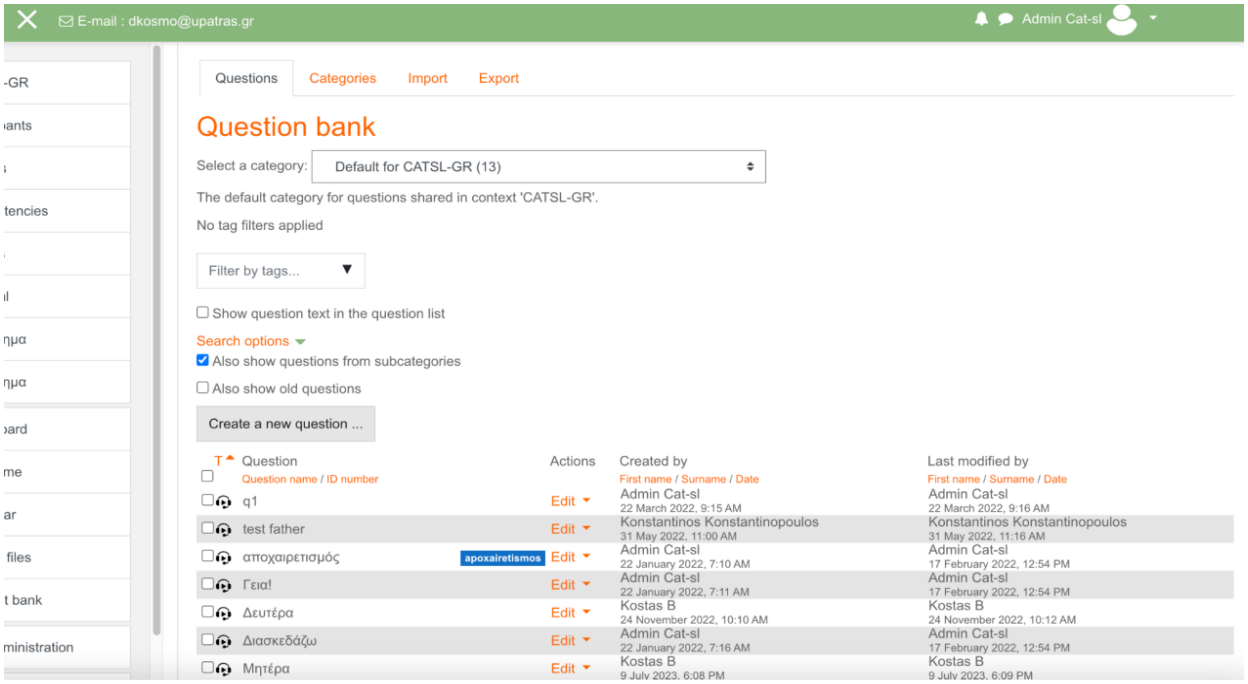


Figure 33 CA19: Course administrators can create a repository of questions (question bank), allowing for repeated use of the questions in different courses and tests without the need for recreation.

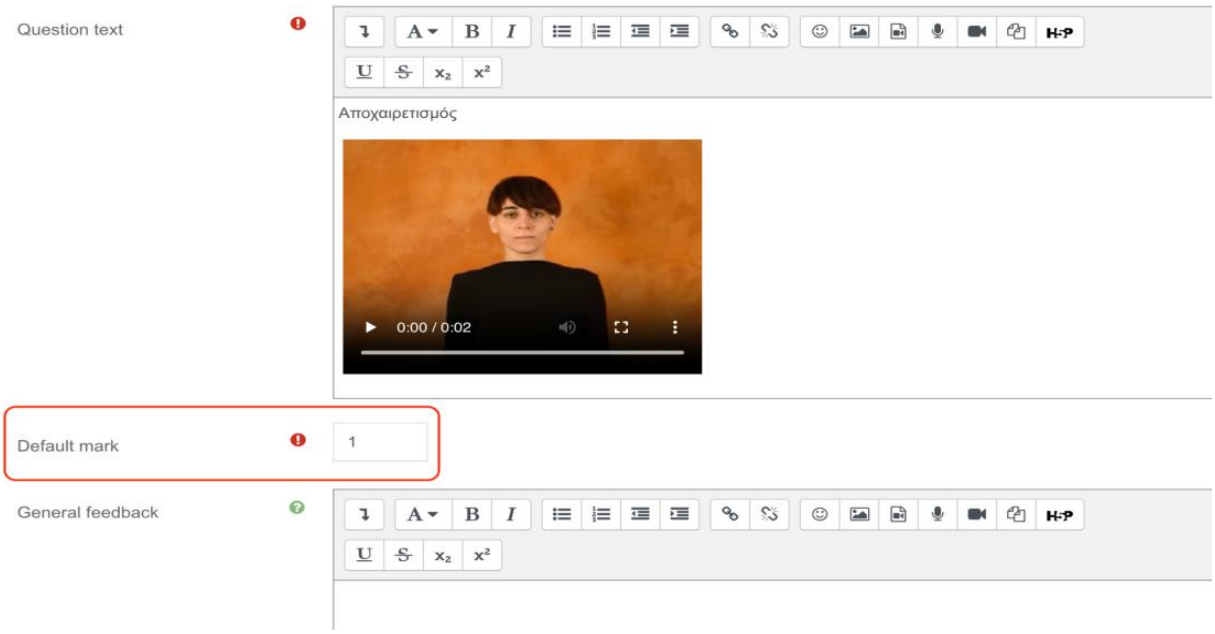


Figure 34 CA20: Course administrators can set different weights for the automatic evaluation of questions and award partial credit.

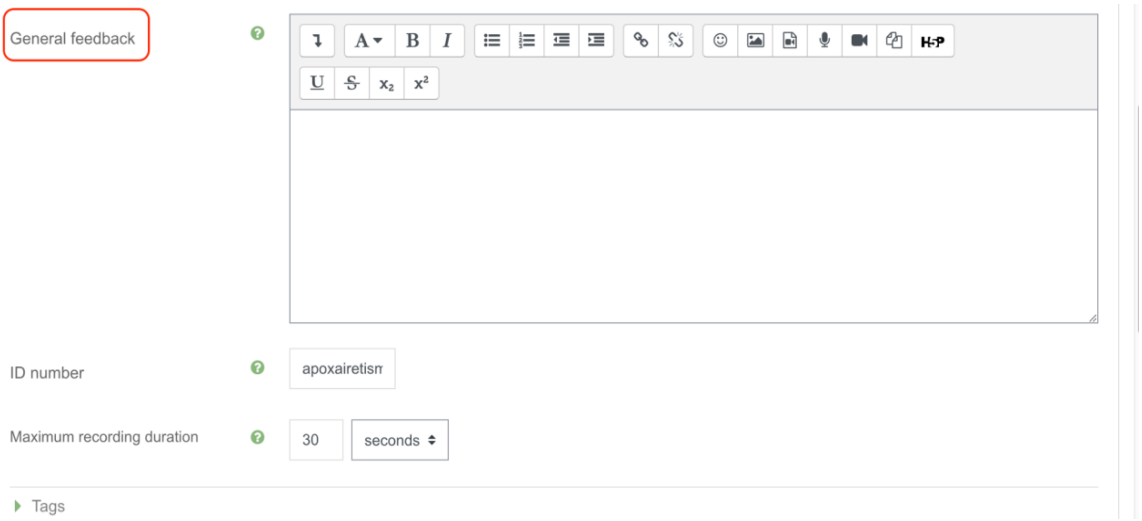


Figure 35 CA21: The system gives the possibility for personal feedback by the teacher or the system itself by using automated methods.

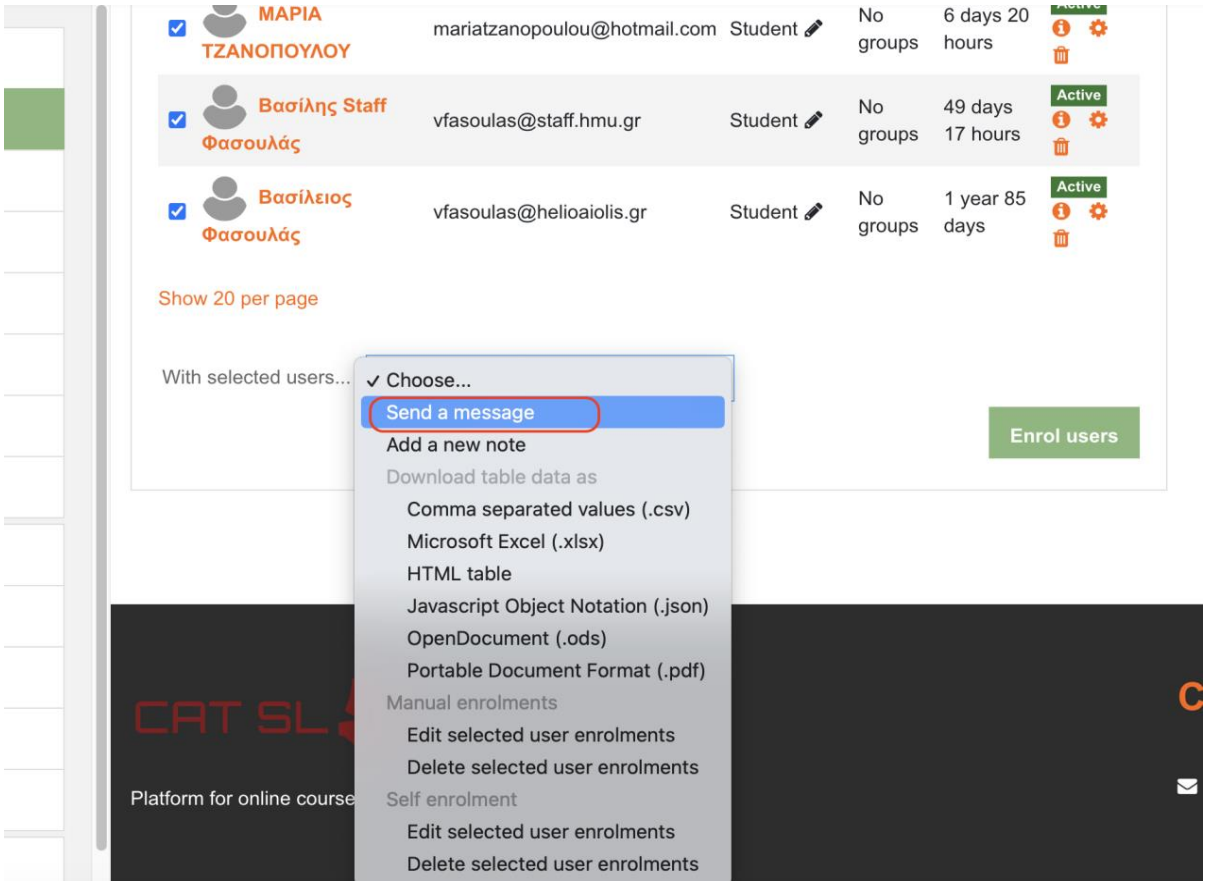


Figure 36 CA22: The platform possesses a built-in system for sending and receiving emails among the course administrator and the students.

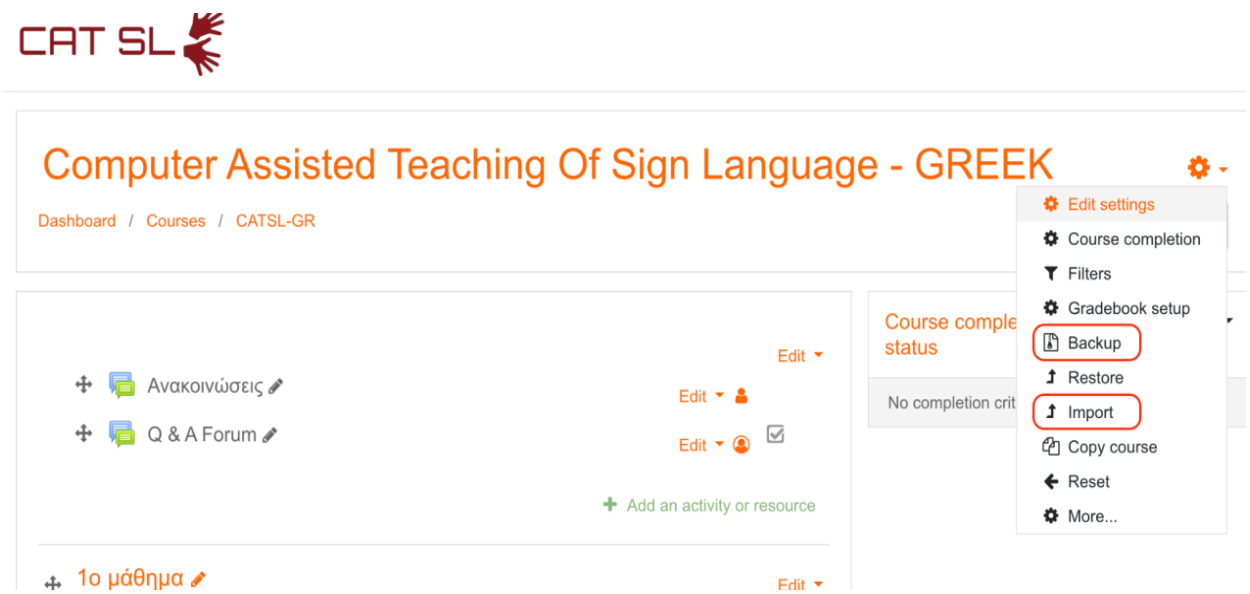


Figure 37 CA23: The system is able to import and export courses and supports accepted standards.

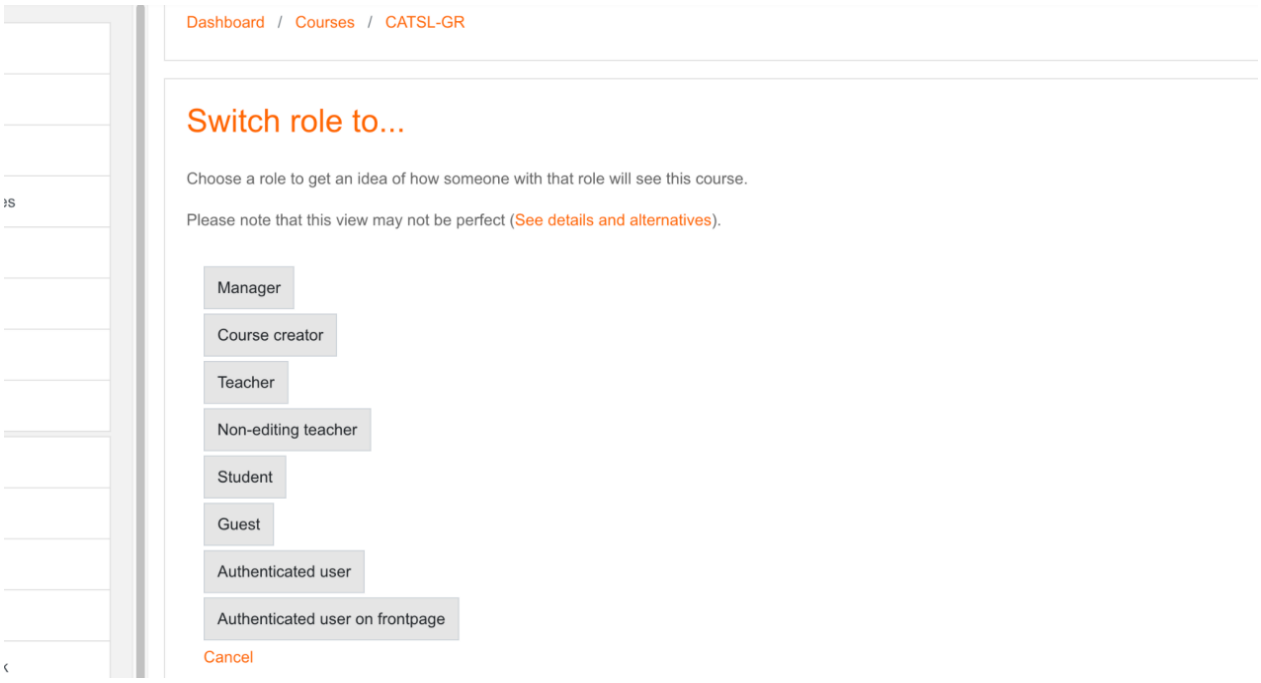


Figure 38 CA24: The course administrator is able to check the student view of the course without logging out of the system and logging in as a student or switching from one interface to another.

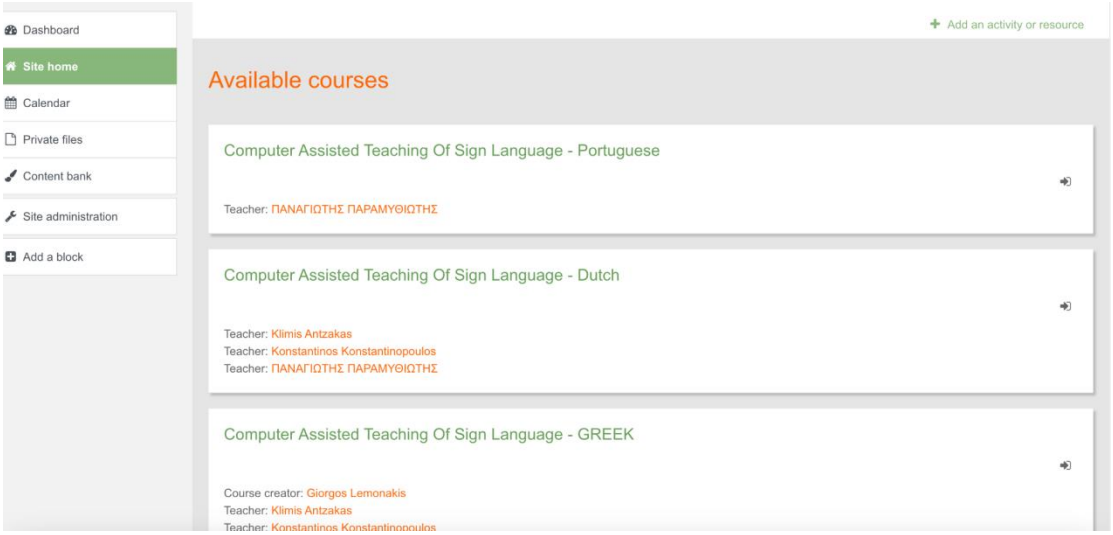


Figure 39 CA25: The teacher is able to create a personal portfolio of courses.

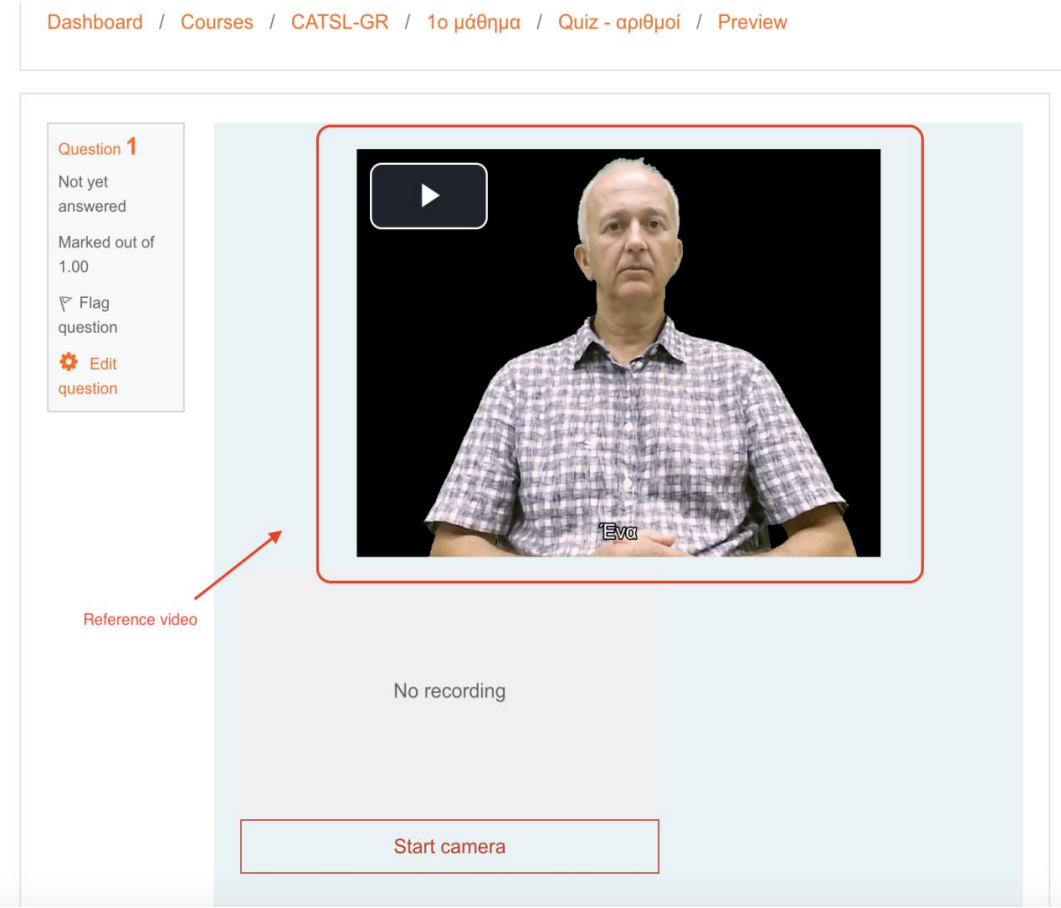


Figure 40 CA26: The teacher is able to create a reference video displaying the teacher signing.

3.3. Content Management

Content management (CM) implementations address the learning content of all users on the platform in order to facilitate work with the learning material.

	Description	Example Use Case
CM1	Store and manage any multimedia content	Course material organization
CM2	Set different access rights to content	Privacy settings
CM3	Track history of each file or directory	Auditing
CM4	Easily view and manage course-related files	Teacher's administrative tasks

Table 3 The content management functions.

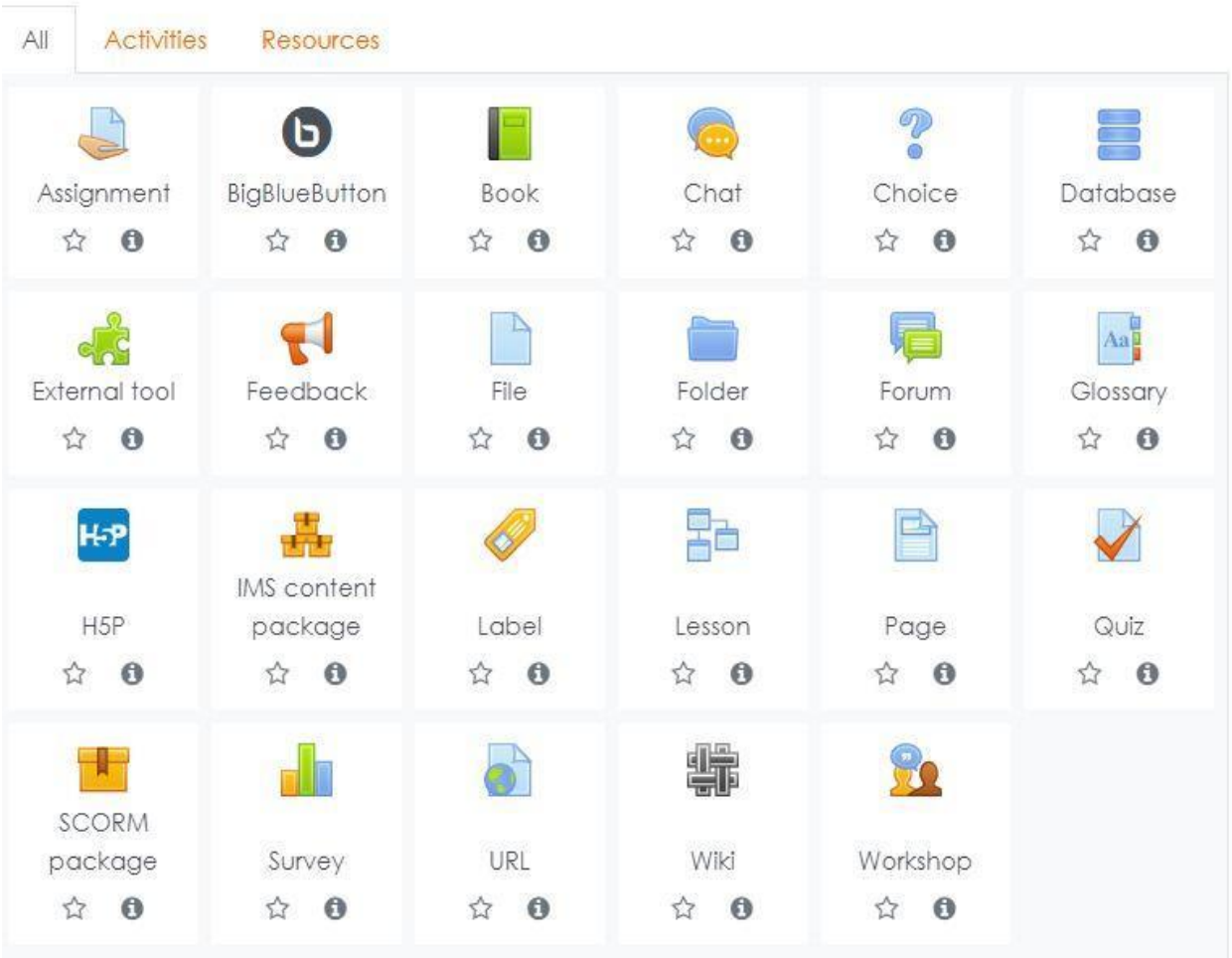


Figure 41 CM1: The system allows to store and manage any type of multimedia content among courses where it can be administered, updated and shared.

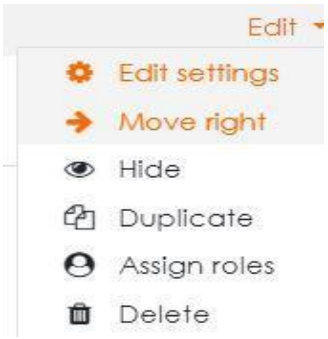


Figure 42 CM2: The system allows users to set different access rights to content.

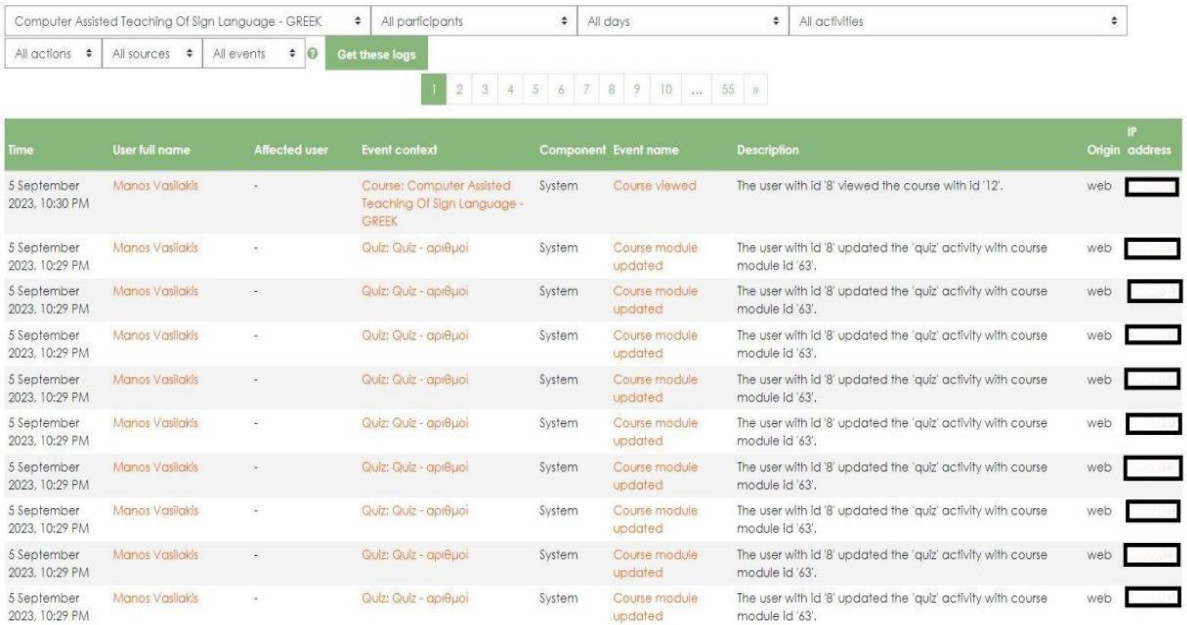


Figure 43 CM3: The system allows to track the history of each file or directory from the content management system, i.e., to check which user accessed it, when and from where.

Figure 44 CM4: Teachers are able to quickly and easily view and manage all the files related to their courses.

3.4. Evaluation

The evaluation part (E) concerns the means for evaluating the students who have already taken a course and the respective evaluation test.

	Description	Example Use Case
E1	Assessment anonymity	Confidential grading
E2	Automatic marking for assignments	Streamlining grading
E3	Generate reports on student progress	Monitoring educational outcomes
E4	Import and export electronic registers	Data sharing with external systems

Table 4: The evaluation functionalities.

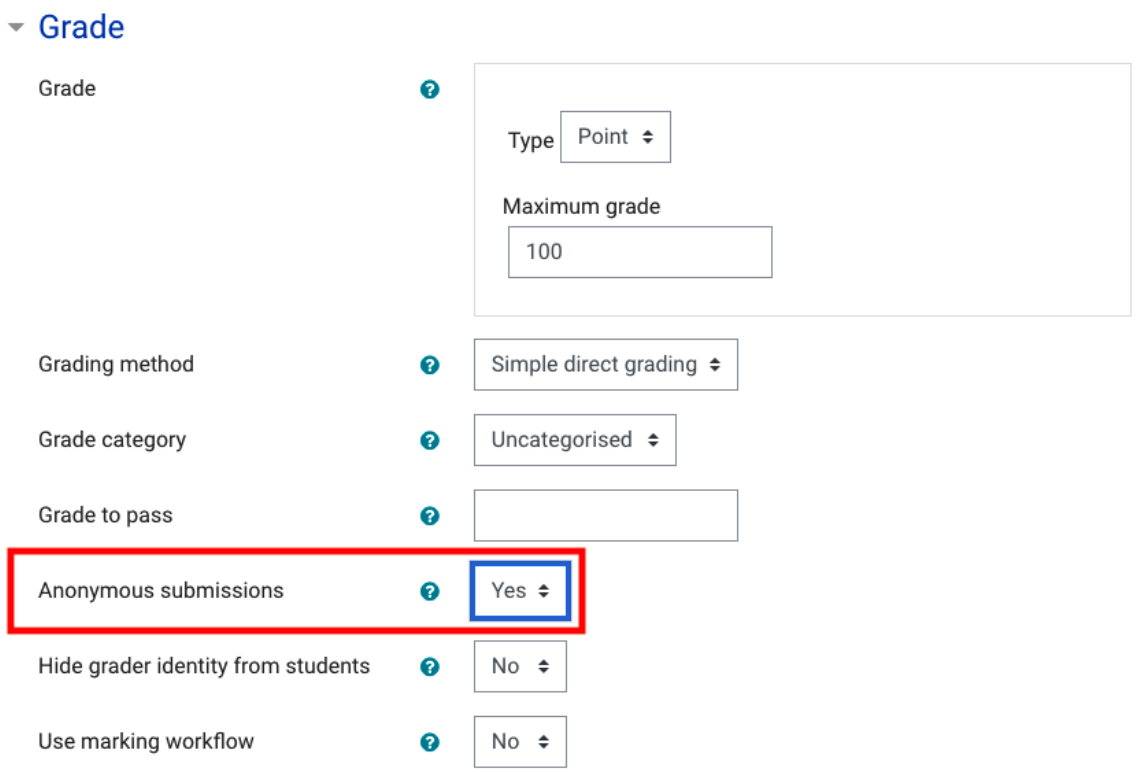


Figure 45 E1: The assessment can be anonymous or not.

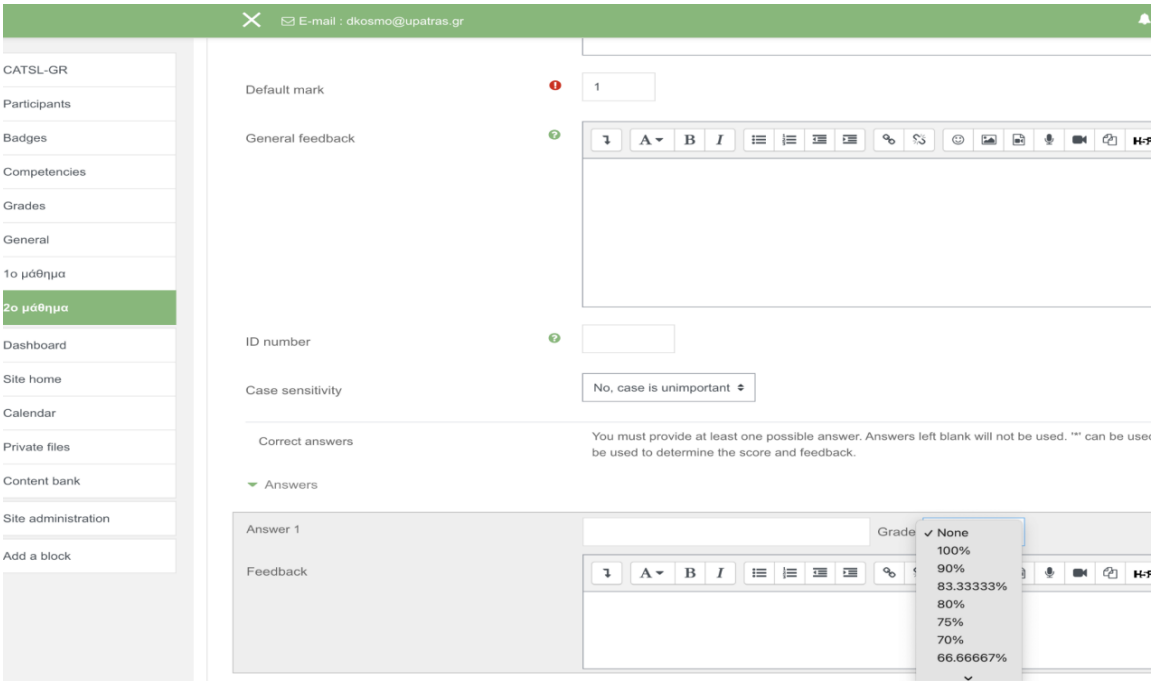


Figure 46 E2: It is possible to automatically add marks for homework, examinations, tests, and other student assignments.

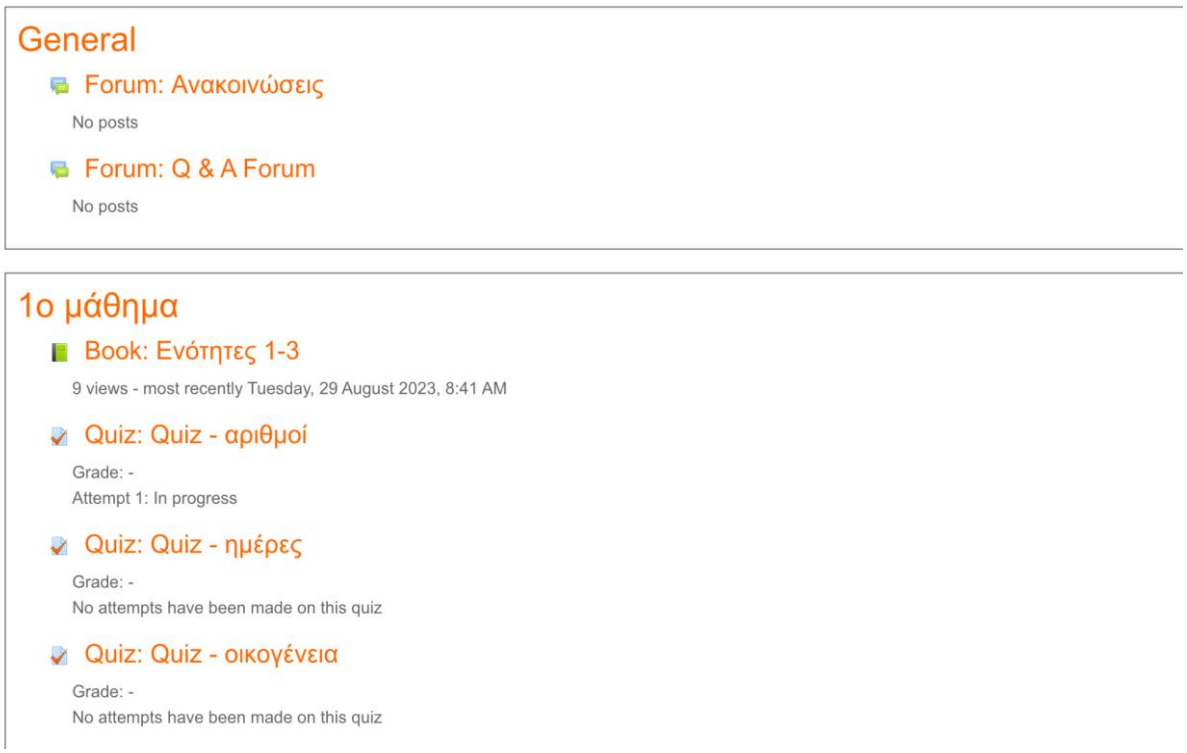


Figure 47 E3: It is possible to generate reports on student progress and analyse the level of acquisition of various elements from the learning material.

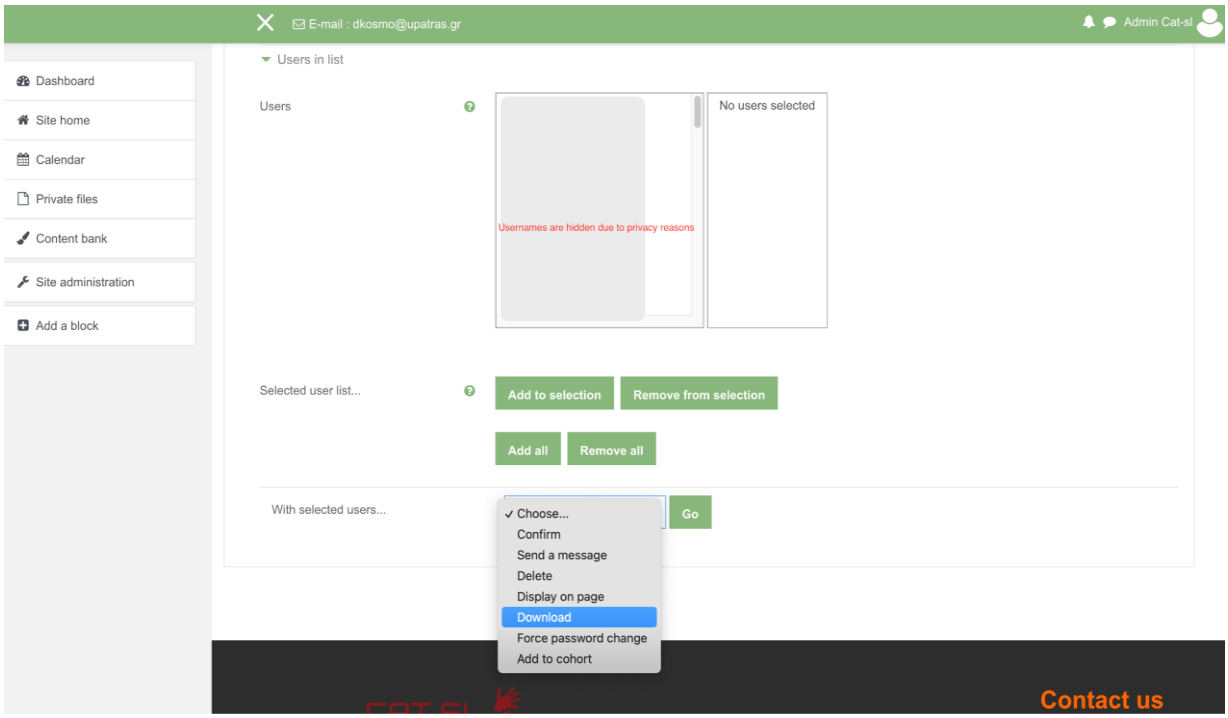


Figure 48 E4: The system allows import and export of the electronic register in xls and csv format for students and teachers.

3.4.1. Type of feedback of CAT-SL to students

The CAT-SL system provides various forms of feedback depending on the type of question asked to the student. In a multiple choice question (e.g. indicate the correct written answer out of a number of given options when a sign video has been presented) could be a simple correct/incorrect feedback, after which a student would be given feedback on what the correct answer should have been in case of an incorrect answer being given.

When asked to imitate a sign that has been presented in the quiz, record and upload the sign video made by the student, CAT-SL returns the uploaded video next to the given target video, with an indication on how successful the student's production has been expressed in percentage correct. The time it takes to upload a video recording and waiting for the system to process it and returning the analysed video was considered to be acceptable.

3.4.2. Assessment of learning effects

Evaluating CAT-SL in a field trial was, unfortunately, beyond the scope of this project.

However, plans have been developed to test the contribution of CAT-SL on the improvement of sign language skills on a representative number of SL-learners and compare results obtained in a control group of sign language learners, who follow sign language classes presented in the traditional way (i.e. without CAT-SL).

To this end randomly selected sign language learners will be given the opportunity to use the CAT-SL system in addition to the regular sign language learning curriculum. The number of times that they will use CAT-SL will be logged, questions will be asked about the user-friendliness of the system, response times, and individual learning curves will be established.

For the control group that will follow the standard curriculum without CAT-SL learning progress curves will also be measured.

In addition teachers will be asked to judge the UI of CAT-SL, focussing on the flexibility of the system, uploading video files and general response times and the ease with which new exercises can be created.

3.5. Body Tracking

The body tracking system (TR) concentrates on identifying the pose of the human parts in each frame acquired by the camera viewing the student. The respective implementations are provided in the following.

	Description	Example Use Case
TR1	The system is able to track both hands in such a way that the hand-shape and hand orientation could be extracted unambiguously.	Student feedback
TR2	The system is able to track the face so that the most common non-manual markers could be extracted unambiguously.	Student feedback

	Description	Example Use Case
TR3	Motion elements like arm motion, head motion, body motion can be extracted.	Student feedback
TR4	Deviations from ground-truth motions can be identified.	Student feedback
TR5	The motion can be calculated in about 10 seconds for small or medium-size sentences.	Student feedback

Table 5 The Body tracking functions.

We implemented all the above functions using the MediaPipe [MediaPipe] library in Python, which is dedicated to human body and hand tracking. In Figure 49 we present a visual example to verify TR1-3.

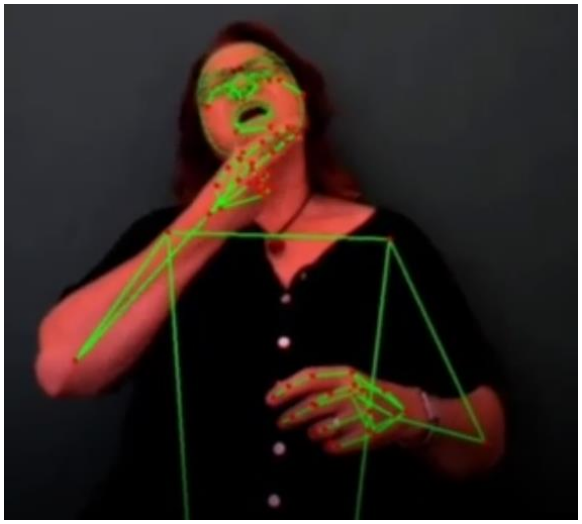


Figure 49 Human body and hand tracking using the MediaPipe functions (TR1-TR2-TR3)

Figure 50 Teacher's and student's hand landmarks (left and right respectively). Student correctly matched the sign for 'a' (middle checkmark) (TR4)

In Figure 50 we present the teacher and student configuration in a form that renders them comparable, so that it is easy to tell the deviations for feedback to the student.

Finally, we have measured the feedback duration for complete signing of short sentences. It is indeed a few seconds in the framework of the e-learning system. It is almost realtime for the stand-alone system, while it is real-time for handshape configurations. That satisfies TR5.

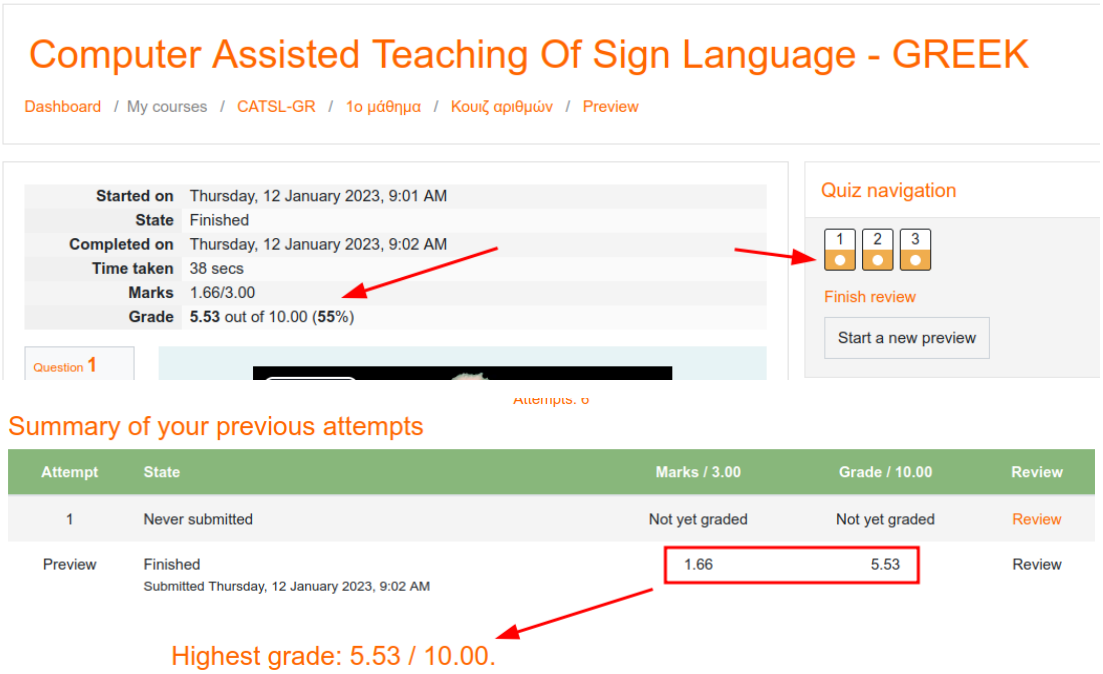


Figure 51 Total score of a quiz attempt on the e-learning platform (RE1-2).

3.6. Sign recognition

The requirements for sign recognition as defined in IO1 are:

RE1: The system is able to tell if the identified motion corresponds to the desired glosses

RE2: The system is able to tell if the identified motion corresponds to the desired punctuation.

Our method uses the hand and body landmarks as a blueprint for comparing and scoring the student's signing against the teacher's. Our method includes custom algorithms that evaluate how closely the student's signing matches that of the teacher. This involves a scoring system that takes into account various factors such as the positioning of the hands, the orientation of the body, and the sequence of movements. This is a very complex task, as we also have to consider variables like lighting conditions, the shape and size of different hands, and the angle at which the signs are performed. Given these complexities, our algorithms are designed to be robust. They can accurately determine how 'acceptable' or close the student's signing is to the teacher's. This is crucial for providing constructive feedback to the student, helping them improve their signing skills over time. To make the process more understandable, we offer visual aids that show the matching and scoring procedure. Example is given in Figure 51.

For longer phrases or sentences in sign language, our method doesn't analyze every single

frame of the video. Instead, it focuses on a selected number of "keyframes." These are frames that our algorithms identify as the most important for the phrase being signed. Transitional movements and pauses, although part of the natural flow of signing, are not as crucial for conveying the core message. This gives an accurate representation of the matching of the entire phrase. This approach to using keyframes ensures that the matching and scoring are focused on what truly matters in sign language communication. It allows us to provide more meaningful and targeted feedback to the student, helping them understand which aspects of their signing need improvement and which are already well-executed.

In addition to our e-learning platform, we also created *a stand-alone application* specifically designed for individual practice of signs. This application uses real-time color-coded scoring for each hand shape, allowing the user to refine their technique with immediate feedback. The scoring system used in the stand-alone app mirrors that of our main platform, allowing easier practice of signs before attempting actual sign language phrases. The idea is to provide a 'training ground' where users can hone their skills before tackling more complex sentences and expressions in sign language. This immediate, color-coded feedback is very useful for users to understand their strengths and areas for improvement in real-time, making the learning process more effective and engaging (Figure 52, Figure 53).

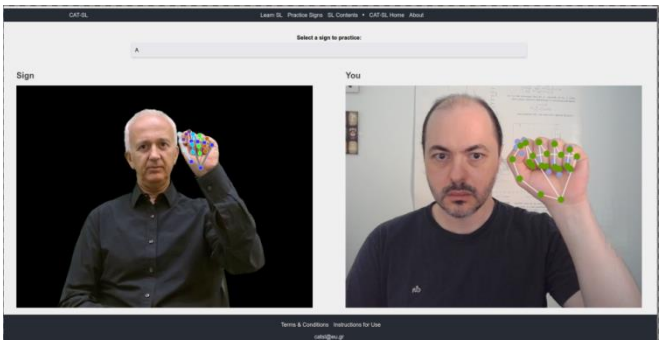


Figure 52 Teacher's and student's hand landmarks in a real-time color-coded scoring practice session.

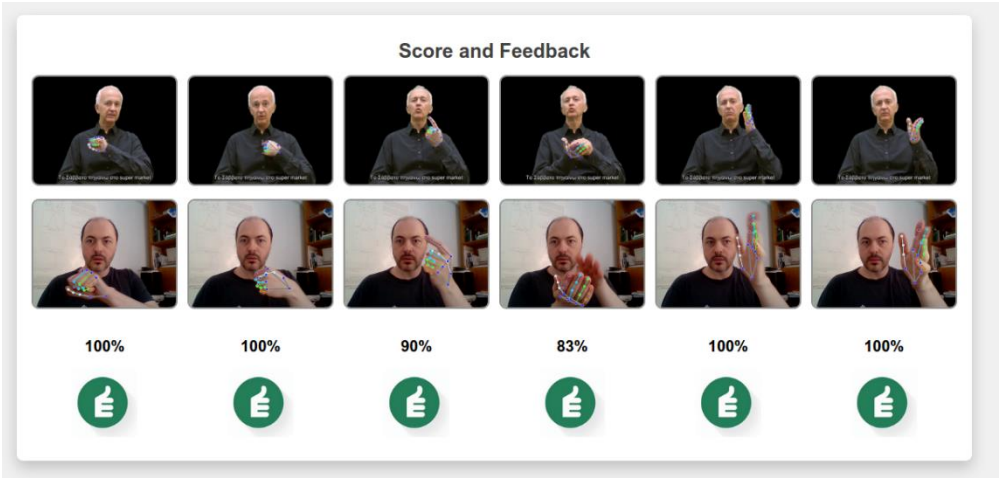


Figure 53 Scoring for individual keyframes of a longer phrase.

This standalone application is available at https://catsl.eelvex.net/. In addition to this, it features a comprehensive overview of the main platform's resources, presenting a bird's-eye view that allows users to navigate and select content easily. Furthermore, the application includes a practice mode that replicates the main platform's environment, ensuring a consistent training experience. This interplay between both platforms enables users to transition smoothly from isolated practice to full sentence constructions within the main system, reinforcing learning and skill acquisition.

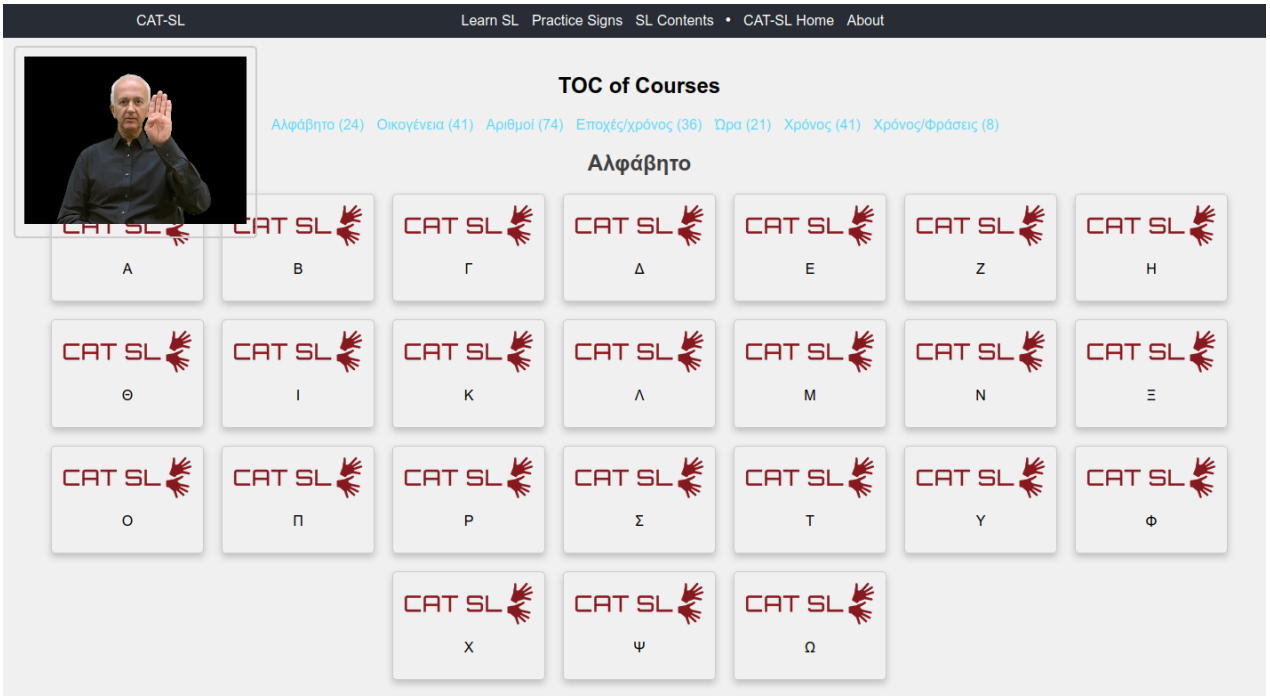


Figure 54 Bird's-eye view of all the content of CAT-SL

3.7. Avatar For Visual Feedback

The avatar, which was mainly developed in the VirtualSign project [Escu2015, VirtualSign] and was adapted for the CAT-SL, is a solution that combines technology and accessibility. It acts as a versatile translator between text and sign languages, with two main functions. It can turn text into sign language performed by a 3D avatar, or it can translate sign language into written text when someone signs in front of a camera with data gloves. This project started from a classroom situation where traditional methods couldn't effectively communicate with students who had hearing disabilities. The CAT-SL project harnesses the power of the VirtualSign technology to provide an innovative sign language course, with a core focus on visual feedback.

In this chapter, we explore the role of the 3D avatar in delivering visual feedback, enabling effective sign language instruction and learning. The 3D avatar acts as a digital sign language model, showcasing signs, expressions, and movements, which are fundamental to the CAT-SL course. The requirements for the avatar are defined in Table 6 and analyzed in the rest of this section.

	Description	Example Use Case
AV1	The avatar is able to display the required hand trajectory.	Content management
AV2	The avatar is able to display the required hand shapes embedded in the trajectory.	Content management
AV3	Motion elements like arm motion, head motion, body motion can be extracted.	Content management
AV4	The avatar is able to display the required body motion.	Content management
AV5	The avatar is able to display the required facial expressions.	Content management

Table 6 List of avatar functions

Figure 55 The 3D avatar

3.7.1. 3D Avatar

The CAT-SL project relies on the VirtualSign 3D avatar to provide visual feedback, ensuring effective sign language instruction and learning. This digital sign language performer serves

as a bridge between teachers and students, delivering an accurate representation of sign language gestures, expressions, and movements.

The 3D avatar used in the CAT-SL project mimics human-like features to enable sign language representation. It's a virtual character capable of performing various sign language signs, creating an immersive learning experience for students. The avatar can convey a wide range of emotions and expressions through its body movements and facial features.

To operate the 3D avatar effectively, it's essential to have sign language experts who understand the intricacies of sign language. They validate the avatar's actions, ensuring the accuracy and cultural relevance of the signs presented.

The 3D avatar is a visual aid that provides immediate feedback on sign language gestures. Students can watch the avatar's movements and compare them to their own, aiding in the learning process.

3.7.2. Anatomy of the 3D avatar

Understanding the anatomy of the avatar is crucial for appreciating its role in teaching sign language effectively. This section provides an overview of the avatar's components and how each contributes to accurate sign language representation (AV1-AV5).

Body: The avatar is built with a humanoid body with a friendly cartoon-game style, allowing for customization and full upper body movement.

Face: The avatar's face is a critical element for conveying emotions and expressions. With a wide range of facial features and expressions.

Arms and Hands: The avatar's arms and hands are customizable, enabling all the necessary hand configurations that are essential for accurate sign language representation. The hands can take on various shapes and positions to perform different signs.

Body Movements: Body movements are a significant part of sign language. The avatar's body can replicate various body movements, including torso twists, head tilts, and other gestures that contribute to sign language grammar.

The facial expressions of the 3D avatar play a vital role in conveying emotions and intensifying the meaning of signs. It can mimic emotions such as happiness, sadness, surprise, curiosity, and even anger. Facial expressions are divided into two categories: eyes and mouth. Users can combine eye expressions from predefined lists with various mouth movements, allowing for a wide range of emotional expressions.

The avatar is designed with sign language in mind. It takes into account the intricacies of sign language, including hand shapes, movements, and facial expressions. Sign language experts work closely with the avatar to ensure that it accurately represents the language's grammar and cultural aspects.

3.7.3. Configuration of Sign Language gestures

In order to generate the avatar videos, the gestures need to be created using the Configurator tool. This section delves into the configuration process, highlighting how sign language gestures are crafted using the 3D avatar.

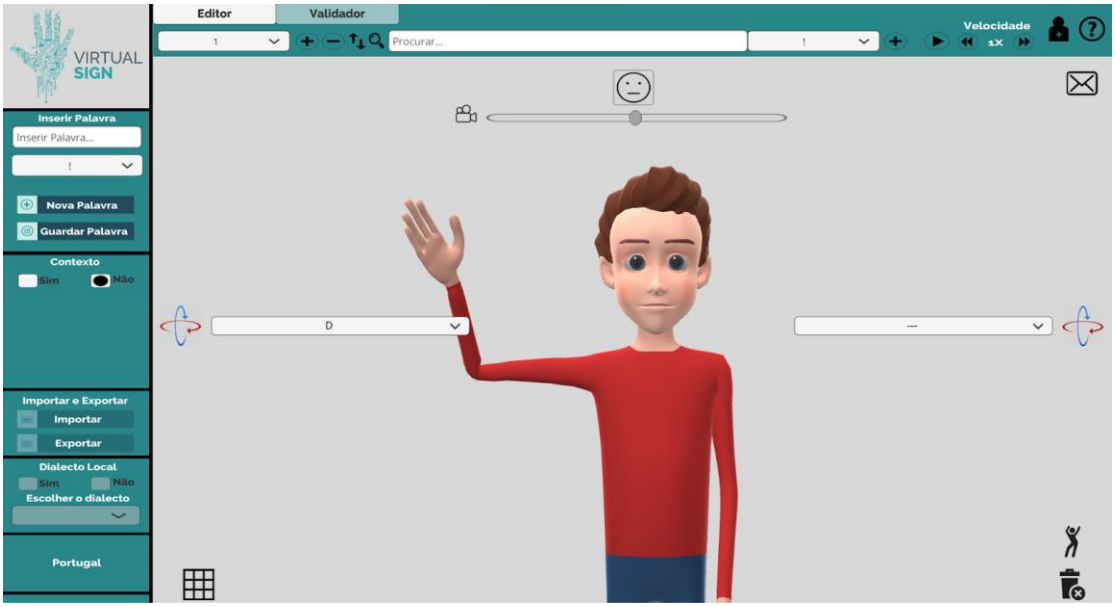


Figure 56 Configuration of an avatar sign

Avatar manipulation: The configuration process starts by manipulating the 3D avatar in a user-friendly interface. The avatar's arms, hands, body, and head are adjusted using an intuitive drag-and-drop style, facilitating precise movements.

Hand Configurations: Hand configurations are a critical aspect of sign language gestures. The 3D avatar allows users to select from a wide array of pre-modeled hand shapes with different finger positions, and movements to match the specific requirements of each sign.

Facial Expressions: Emotions and facial expressions are fundamental in sign language. The 3D avatar can mimic a range of facial expressions, contributing to the nuanced aspects of sign communication. By combining eye and mouth movements, it can easily represent emotions like happiness, sadness, surprise, curiosity, and anger, or even motions such as tongue in cheek.

Keyframes: To create sign language gestures, keyframes are registered to capture different moments within the gesture. Users define the initial and final positions of the avatar, along with keyframes for when hand configurations change during the sign. These keyframes serve as reference points for the avatar's movements.

Interpolation: The software interpolates movements between keyframes to create smooth and natural gestures. This process ensures that the sign language gestures appear fluid and accurate.

Gesture reuse: To streamline the configuration process and save time, users have the option to copy previous keyframes for movements that require repetition. This feature simplifies the creation of multiple signs with similar elements.

Importing gestures: Users can also import keyframes from existing gestures stored in the database. This is particularly useful when configuring signs that share similarities with previously created gestures.

Two main user roles are involved in the configuration of sign language gestures:

Editors: Editors can be any user, with or without expertise in sign language. They use the configurator to mimic sign language movements found in videos and dictionaries. These gestures are saved in the project's database. However, these gestures remain unvalidated and unused in the course material until they are approved by validators.

Validators: Validators are sign language experts, typically individuals with sign language as their primary language or certified sign language interpreters. They are responsible for validating signs created by editors. The configurator provides a built-in messaging system that allows validators to give feedback to the editor responsible for a specific sign. This communication ensures that signs meet the required standards and are accurate representations of sign language.

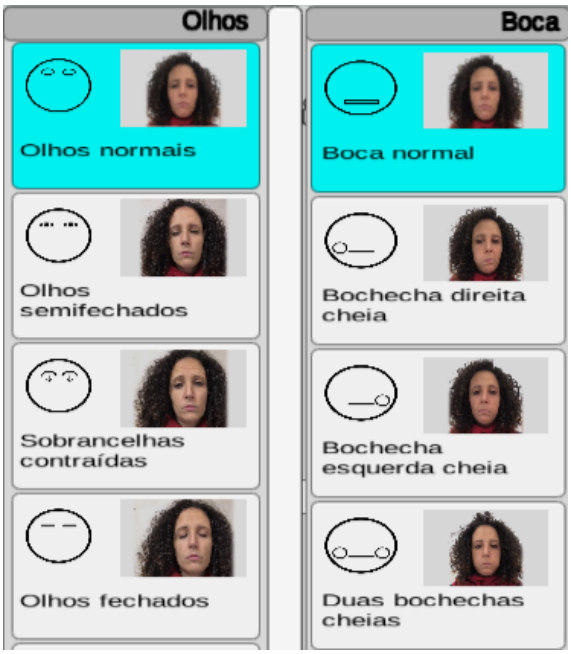


Figure 57 Configuration of expressions

The configuration of sign language gestures is a meticulous process. Users manipulate the 3D avatar, fine-tune hand configurations, facial expressions, and body movements. Keyframes are registered to capture the gesture's moments, and interpolation creates fluid movements. Editors and validators play crucial roles in the validation and creation of these gestures, ensuring they are accurate representations of sign language.

3.7.1. Production of Avatar Videos

As part of the CAT-SL project's commitment to providing an inclusive and accessible sign language course, the efficient creation of avatar videos is essential. To streamline this process, we use the VirtualSign Capture Tool, a program designed to simplify the production of sign language content at scale.

The creation of avatar videos with this program builds upon the foundation of the VirtualSign text-to-sign translator. This tool takes a simple yet powerful approach: it accepts a text file as input, where each line in the file represents a word, sentence, or phrase to be translated into sign language gestures. The program then seamlessly translates the textual content into

corresponding avatar gestures. For each line of text in the input file, the tool records a separate video of the 3D avatar performing the sign language gestures.

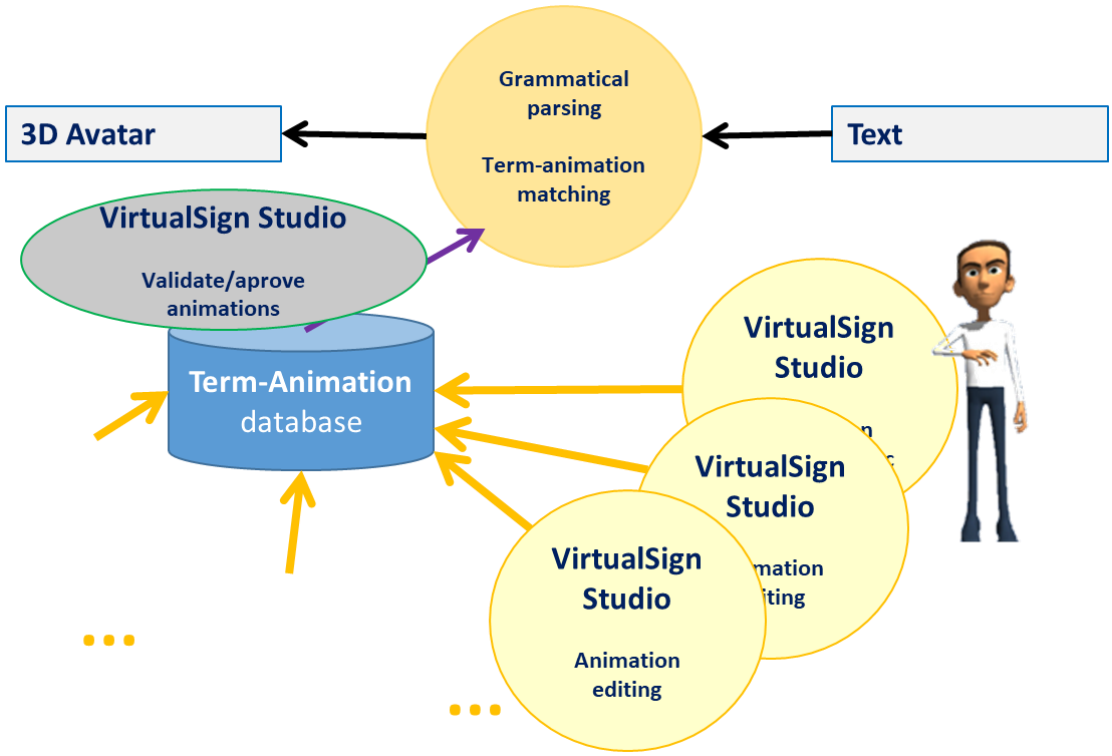


Figure 58 Workflow of the avatar

Figure 59 Playback of the avatar in various contexts

The VirtualSign Capture Tool offers a crucial advantage in the efficiency of sign language content production. By automating the video creation process, it significantly expedites the generation of educational materials. This level of automation ensures that educators and course developers can efficiently produce numerous sign language videos, a particularly

valuable asset for a course like CAT-SL.

Users have the freedom to customize the input text file to include words, sentences, or phrases of varying lengths and complexities. The tool is designed to adapt to these variations, delivering accurate and personalized sign language representations for each word or sentence.

The outcome of the VirtualSign Recorder process is a series of separate video files. Each video corresponds to the 3D avatar performing sign language gestures for the corresponding word or sentence in the input file. These videos are seamlessly integrated into the CAT-SL course materials, contributing to a comprehensive and accessible sign language learning experience.

3.7.2. Training Course: Utilizing the configurator

During the CAT-SL project we have provided a simple training course for using the VirtualSign Configurator. This course is designed to empower educators and course developers, with the skills and knowledge required to effectively utilize the configuration tool.

The primary objectives of this training course are to familiarize participants with the configurator's interface and functionalities. Participants will gain hands-on experience in configuring and validating sign language gestures. The course focuses on providing guidance to both editors and validators in building the sign language lexicon used by the 3D avatar.

Participants in the course undergo a step-by-step process to configure sign language words, offering them valuable hands-on experience with the VirtualSign Configurator. This practical training equips them with the proficiency needed to interact with the tool confidently.

The course serves as an important aspect of the CAT-SL project, empowering individuals to become proficient content contributors. Through the training, they become better prepared to contribute to the development and validation of sign language content, ultimately enhancing the educational experience for learners.

3.7.3. Limitations and Future Improvements

While the CAT-SL project benefits from the integration of the 3D avatar and the configurator, there are some challenges and opportunities for further enhancement, mainly in the facial expression and arm movement areas.

The 3D avatar's facial expressions are still somewhat limited, primarily focusing on eyes and mouth movements. While it provides a large number of combinations, there is room for expanding the range of facial expressions, which can make the signs more expressive.

Some users have provided feedback regarding the avatar's gesture speed and quality. Improvements in animation speed and more detailed gestures, including elbow positions, are areas that will be addressed to enhance the visual quality of the signs.

The process of acquiring data for configuring sign language gestures, especially for new sign languages, can be time-consuming. There's room for streamlining the data acquisition process to make it more efficient.

4. Non-Functional Implementations

The system meets the following quality (NF) implementations:

NF1: Productivity – the system has the capacity to serve a reasonable amount of simultaneous (concurrent) sessions, depending on class size.

NF2: Scalability and flexibility – the system is modular to expand and serve more users, and allows for additional settings, as more users might be interested in using the provided services.

NF3: Compatibility – the system is compatible with current Web standards (HTML, XHTML).

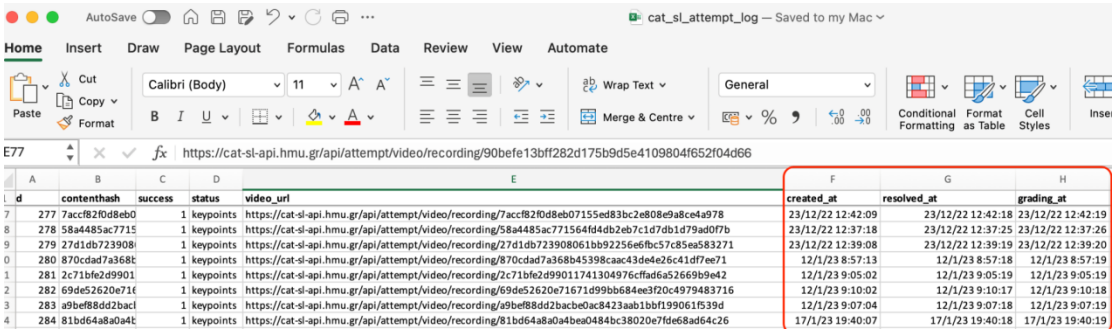
NF4: Accessibility – the system is installed centrally on more than one server and is accessible via http or https over the Internet.

NF5: The system has a web-based user and administrative interface for public and protected sections.

NF6: The system is compatible with the most popular Internet browsers, such as Internet Explorer v.8 and higher, Mozilla Firefox v.3.6 and higher, Safari v.3 and higher and Chrome v.10 and higher.

NF6: Stability – the system guarantees a secure and reliable learning process.

NF7: The automated evaluation is done within 20 seconds for medium size sentences.



d	contenthash	success	status	video_url	created_at	resolved_at	grading_at
277	7accf82f0d8eb0	1	keypoints	https://cat-sl-api.hmu.gr/api/attempt/video/recording/7accf82f0d8eb07155ed83bc2e808e9a8ce4a978	23/12/22 12:42:09	23/12/22 12:42:18	23/12/22 12:42:19
278	58a4485ac7715	1	keypoints	https://cat-sl-api.hmu.gr/api/attempt/video/recording/58a4485ac771564fd4db2eb7c1d7db1d79ad0f7b	23/12/22 12:37:18	23/12/22 12:37:25	23/12/22 12:37:26
279	27d1db723908	1	keypoints	https://cat-sl-api.hmu.gr/api/attempt/video/recording/27d1db723908061bb92256e6fbc57c85ea583271	23/12/22 12:39:08	23/12/22 12:39:19	23/12/22 12:39:20
280	870cdad7a368b	1	keypoints	https://cat-sl-api.hmu.gr/api/attempt/video/recording/870cdad7a368b45398caac43de4e26c41df7ee71	12/1/23 8:57:13	12/1/23 8:57:18	12/1/23 8:57:19
281	2c71bfe2d9901	1	keypoints	https://cat-sl-api.hmu.gr/api/attempt/video/recording/2c71bfe2d990117413049 76cffad6a52669b9e42	12/1/23 9:05:02	12/1/23 9:05:19	12/1/23 9:05:19
282	69de52620e716	1	keypoints	https://cat-sl-api.hmu.gr/api/attempt/video/recording/69de52620e71671d99bb684ee3f20c4979483716	12/1/23 9:10:02	12/1/23 9:10:17	12/1/23 9:10:18
283	a9bef88dd2bacl	1	keypoints	https://cat-sl-api.hmu.gr/api/attempt/video/recording/a9bef88dd2bacbe0ac8423aab1bbf199061f539d	12/1/23 9:07:04	12/1/23 9:07:18	12/1/23 9:07:19
284	81bd64a8a0a4b	1	keypoints	https://cat-sl-api.hmu.gr/api/attempt/video/recording/81bd64a8a0a4bea0484bc38020e7fde68ad64c26	17/1/23 19:40:07	17/1/23 19:40:18	17/1/23 19:40:19

5. Evaluation

During the development of the CAT-SL system a number of evaluation cycles were conducted to validate different system aspects. These aspects can be categorized as follows:

- Functionality and user-friendliness of the user-interface, both from the point of view of the student as well as the teacher.
- Robustness of the system.
- Assessment of learning effects.
- Type of feedback of CAT-SL to the student.
- CAT-SL and the Deaf user.

In the following sections each of these aspects will be separately addressed.

5.1.1. Functionality and user-friendliness of the user-interface

During the development of the CAT-SL system a number of evaluation cycles were run through and presented at the various multiplier events (ME). These events saw participation from 15

to 25 diverse stakeholders each time, encompassing deaf/hard-of-hearing individuals, parents of the deaf/hard-of-hearing, teachers, and more. The first ME was conducted online due to COVID-19 restrictions, while subsequent events were hosted in Porto, 's-Hertogenbosch, and Athens. In these development processes different aspects of CAT-SL were tested and attendees of the multiplier events were asked to give their opinion about them. Starting with the first ME, the initial focus was on gauging interest in a computer-assisted application for learning Sign Language and direct our ongoing implementations. By the time of the last ME, the CAT-SL system had undergone significant refinement.

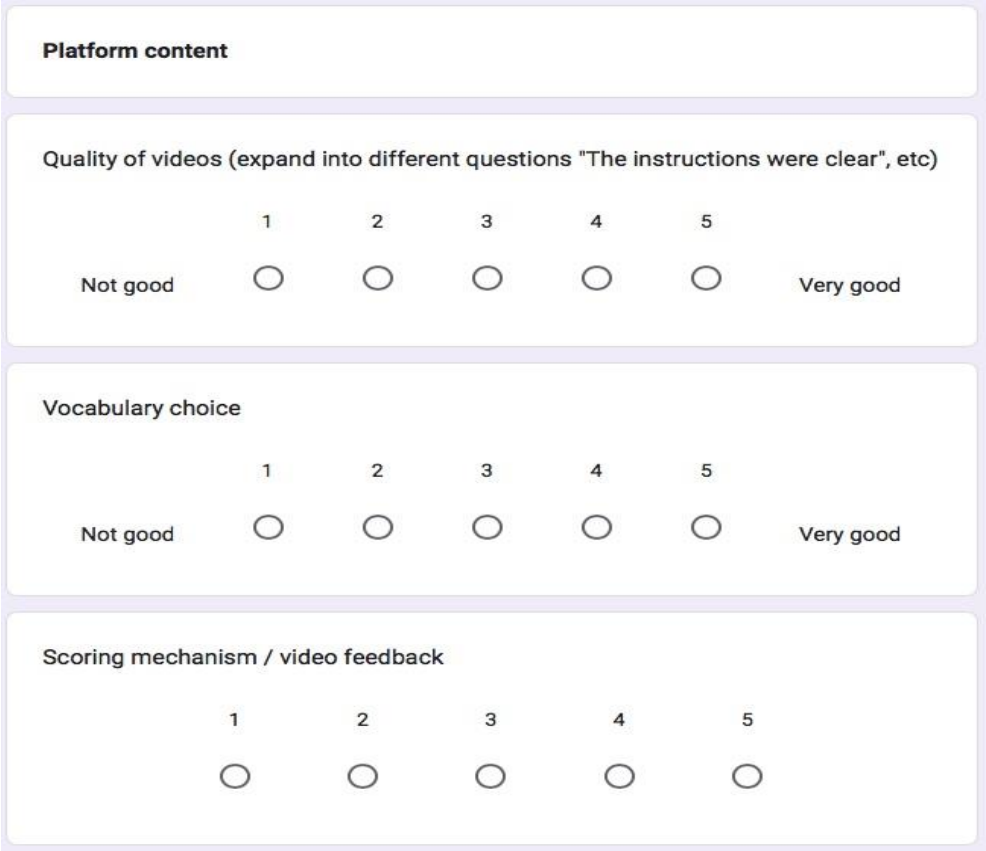


Figure 60 Evaluation Form (Platform)

Whereas the first questionnaire focussed primarily on the possible use of an application for computer-assisted learning of Sign Language, consecutive questionnaires incorporated questions directed at its functionality, UI and user-friendliness.

A first impression of CAT-SL scored 7.7 on a 10-point scale. People valued the innovative aspect of CAT-SL and its state-of-art implementation. It was considered to be of added value in independent learning of Sign Language (7.4). Additionally, it was thought to be an asset which could be very motivating for students, and most importantly, would enable a detailed and repeated look at how a sign should be correctly made. Because of its flexibility, teachers would be able to develop their own educational material, that could be added to an existing curriculum.

Demos that were presented have been given an average value of 7.6 (out of a possible 10). In general, comments were given on the facial expression of the avatar, which needed

improvement.

In subsequent questionnaires (see figures below) 90% of the respondents rated the quality of the signed videos as very good or excellent. They also indicated that in general they were interested in using the CAT-SL application in the future. The most requested improvement is the inclusion of pose and facial expressions.

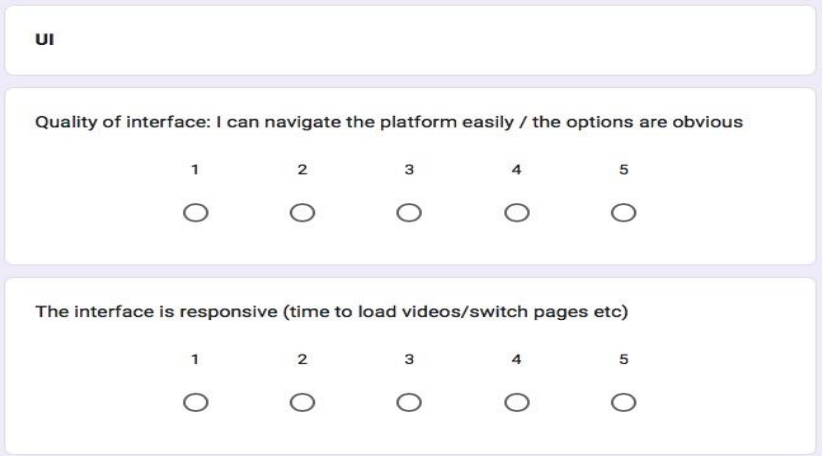


Figure 61 Evaluation Form (UI)

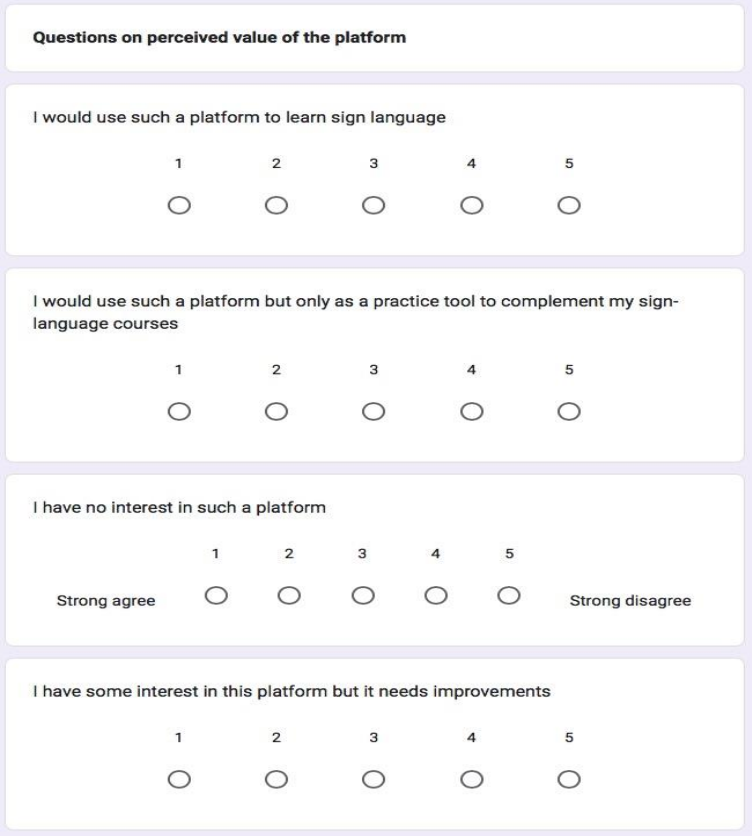


Figure 62 Evaluation Form (Value)

5.1.2. CAT-SL and the Deaf user

Deaf users who have been shown demos of the CAT-SL system were very positive about the possibility of CAT-SL being used for studying sign language, especially by hearing students. In addition deaf users also indicated that the system could possibly be used for learning a second sign language, although for this particular purpose special courseware would have to be developed.

6. Conclusions and Future Steps

We have presented a document that serves as the basis for the development of an automated system for teaching sign languages. It can be used for any such development in the future.

We have presented the user roles and groups. We presented the functional and non-functional implementations and the basic use cases, for use by the developers. We made special mention to quizzes and ways to incorporate them. We finally presented the evaluation criteria for the system.

The CAT-SL project has successfully developed an innovative automated platform for sign language teaching, based on cutting-edge technologies such as computer vision and machine learning. The system has been designed to serve the needs of both teachers and learners, in a diverse range of settings, from higher education institutions to special education schools and independent learners. Importantly, the system has been designed with scalability and adaptability in mind, allowing it to be extended to cover more sign languages and to be adapted for different educational levels and proficiency levels.

The project has demonstrated that the use of advanced technologies can significantly enhance the teaching and learning of sign languages. By providing immediate, visual feedback on signing accuracy, the system offers a unique learning experience that can help students improve their signing skills and gain confidence in their abilities. The stand-alone practice application, with its real-time color-coded scoring, is particularly effective in this regard.

The user-friendly interface of the CAT-SL system, along with its robustness and stability, have been highly rated by users. The system's ability to serve a reasonable number of simultaneous sessions, its compatibility with current web standards and popular Internet browsers, its accessibility over the Internet, and its seamless integration into the Moodle learning management system all contribute to its high user acceptance.

At the same time, the project has identified areas for improvement, notably the need to include pose and facial expressions in the sign recognition process, and the need to improve the facial expression of the avatar. These are complex challenges that require further research and development effort.

6.1.1. Future Steps

Looking ahead, there are several potential directions for further development of the CAT-SL project. Firstly, there is scope for expanding the range of sign languages covered by the system. While the initial implementation focused on Dutch, Greek, and Portuguese sign languages, the same methodology can be applied to many more languages in Europe and beyond.

Secondly, there is potential for developing more courses for different educational levels and

proficiency levels in each sign language. This could include courses aimed at beginners, intermediate learners, and advanced learners, as well as courses focusing on specific topics or domains of use.

Thirdly, there is an opportunity to carry out more extensive field trials to assess the learning effects of the CAT-SL system. This could involve testing the system with a larger number of sign language learners, comparing results with a control group of learners who follow traditional sign language classes, and getting feedback from teachers on the usability of the system.

Fourthly, there could be potential for developing the system for use in learning a second sign language. This would require the development of special courseware and might appeal to a different group of learners.

Finally, ongoing work will be needed to keep the CAT-SL system up to date with advances in technology and pedagogy. This will involve continuous monitoring of developments in areas such as computer vision, machine learning, e-learning, and sign language teaching, and making necessary updates to the system.

In general, while the CAT-SL project has made significant strides in automated sign language teaching, there is a lot of exciting potential for further development and wider application of this innovative approach.

References

[Moodle] The Moodle platform, https://moodle.org

[MediaPipe] Google, https://developers.google.com/mediapipe

[VirtualSign] https://vs.virtualsign.com/en/index.php

[Escu2015] Escudeiro, P., Escudeiro, N., Reis, R., Lopes, J., Norberto, M., Baltasar, A. B., ... & Bidarra, J. (2015). Virtual sign–a real time bidirectional translator of portuguese sign language. Procedia Computer Science, 67, 252-262.

[Luga2019] Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., & Grundmann, M. (2019). Mediapipe: A framework for building perception pipelines. arXiv preprint arXiv:1906.08172.

[Econ2019] Economou, D., Gonzalez Russi, M., Doumanis, I., Bouki, V., Mentzelopoulos, M., Ferguson, J. (2019). Adult learning Sign Language by combining video, interactivity and play in a 3D game platform. https://doi.org/10.3217/978-3-85125-657-4.

[Woll2003] Woll, B. (2003). Modality, universality and the similarities among sign languages: an historical perspective. In A. Baker, B. van den Bogaerde, and O. Crasborn (eds), Cross-linguistic perspectives in Sign Language Research: Selected Papers from TISLR 2000. Hamburg: Signum, 17–30.