Co-funded by the
Erasmus+ Programme
of the European Union

DIVERSASIA

Embracing diversity in ASIA through the adoption of Inclusive Open Practices

618615-EPP-1-2020-1-UKEPPKA2-CBHEJP

WP2 – Development

D2.3 DIVERSASIA portal with integrated mobile app

V1.0 (Final)

Lead WP2 by P1

Author: PhoenixKM & NTU

Contributors: NTU, NARHU, SoftQNR

Table of Contents

1 Introduction

The document at hand provides an overview of the technical setup of the DiversAsia portal with the integrated mobile app.

This is a living document, since functionalities will be added, based on the toolkit content which is being developed in D2.2 DiversAsia Toolkit. These will be described in an updated version at a later stage in 2023.

2 Outline of the portal

The WCAG2.0 compliant portal at https://diversasia-accessible-he.eu/elearning/ is supported by the Moodle platform and an integrated native mobile app for Android and iOS. It contains:

1. The place where DIVERSASIA Toolkit is hosted which will offer up to date information on accessible higher education:
    - Here the main content information captured in D2.2 DiversAsia Toolkit will be integrated.
    - This will be enriched with video material, images, etc. and therefore the H5P content framework will be used, applying interactive rich HTML5 content and applications.
2. HEI staff community
    - Moodle platform functionality offers community functionalities.
3. Outcomes of project, freely available
    - All outcomes of the project will be hosted on the portal, unless third party platforms like Google Play or App Store are being used.
4. A wizard style online database (semantically driven using relevant ontology) of best/good practices hosts the collected practices of WP1. In addition, the MyHub (https://www.inclusion-hub.eu/) practices database has also been integrated into the portal.

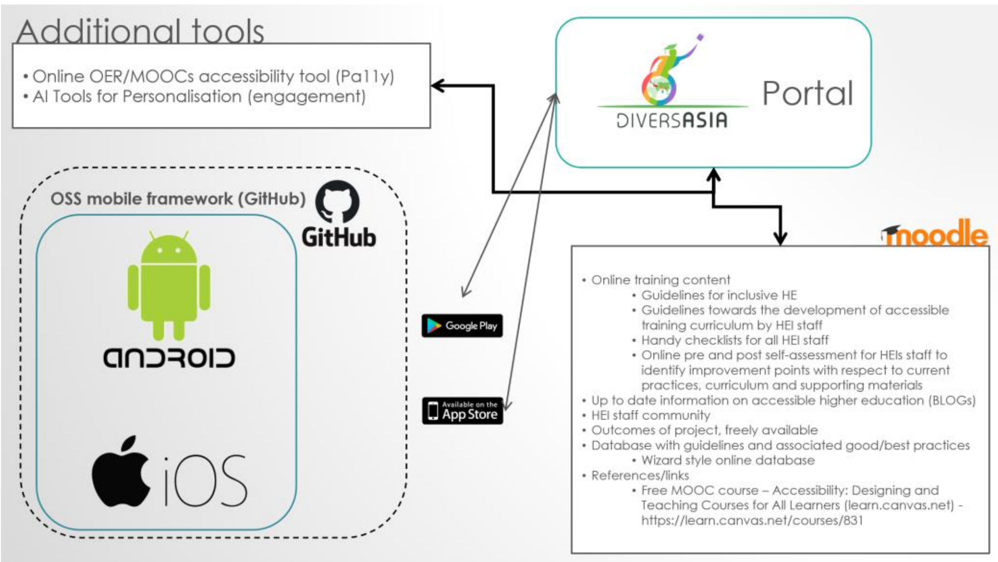


Figure 1 Schematic of DiversAsia portal

2.1 Portal operated by Moodle

We opted for Moodle (https://moodle.com/) for a number of reasons:

- Moodle is a free software, a learning management system providing a platform for e-learning.
- Moodle helps educators considerably in conceptualizing the various courses, course structures and curriculum thus facilitating interaction with online learners.
- It is the most widely used OSS LMS.
- Moodle has a WCAG 2.1 Level AA accreditation.
- There is the availability of an OSS mobile framework so as to develop native Android and iOS mobile applications.

2.2 Newly developed and customised plugins for Moodle

Following plugins are being integrated:

- An in-house self-assessment plugin is ready for integration to support inclusive education readiness self-assessments by higher education institution staff.
- In addition, a customised database plugin was deployed for our portal so that practices can be easily sorted and shared.

We deployed 2 such databases. The screenshot below (figure 2) shows the access to both databases via https://diversasia-accessible-he.eu/elearning/?lang=en

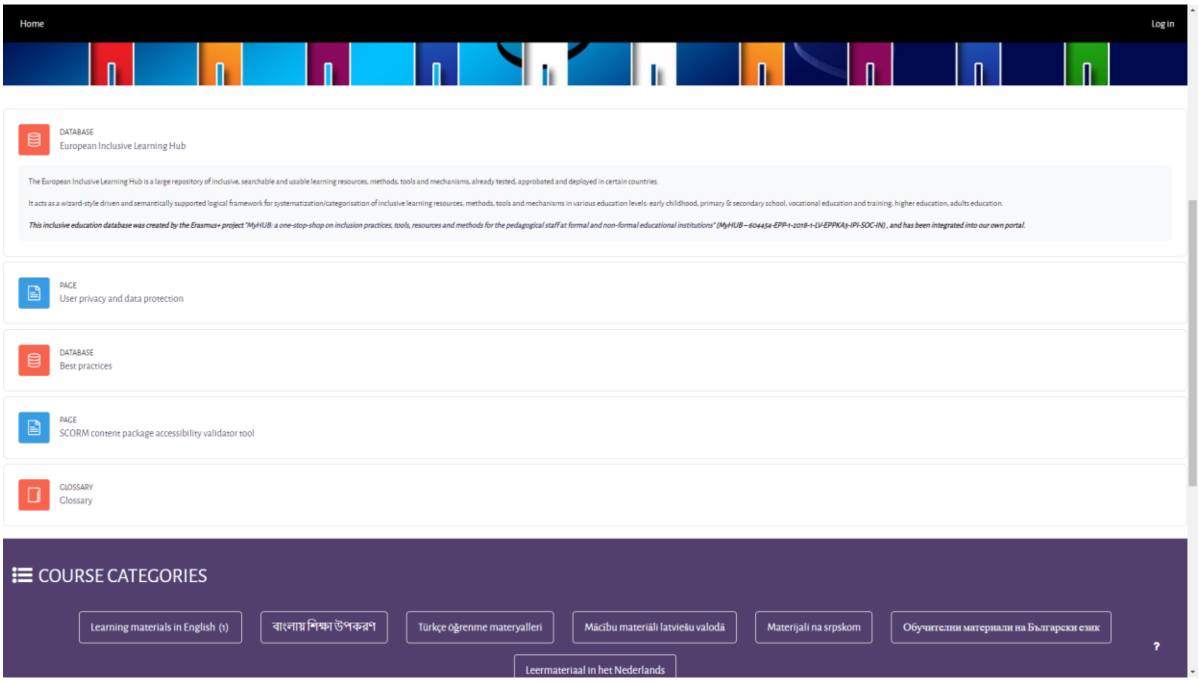


Figure 2 DiversAsia database frontend

2.2.1 DiversAsia best practices database

This captures all the 341 good and best practices as were collected in WP1 in 2021.

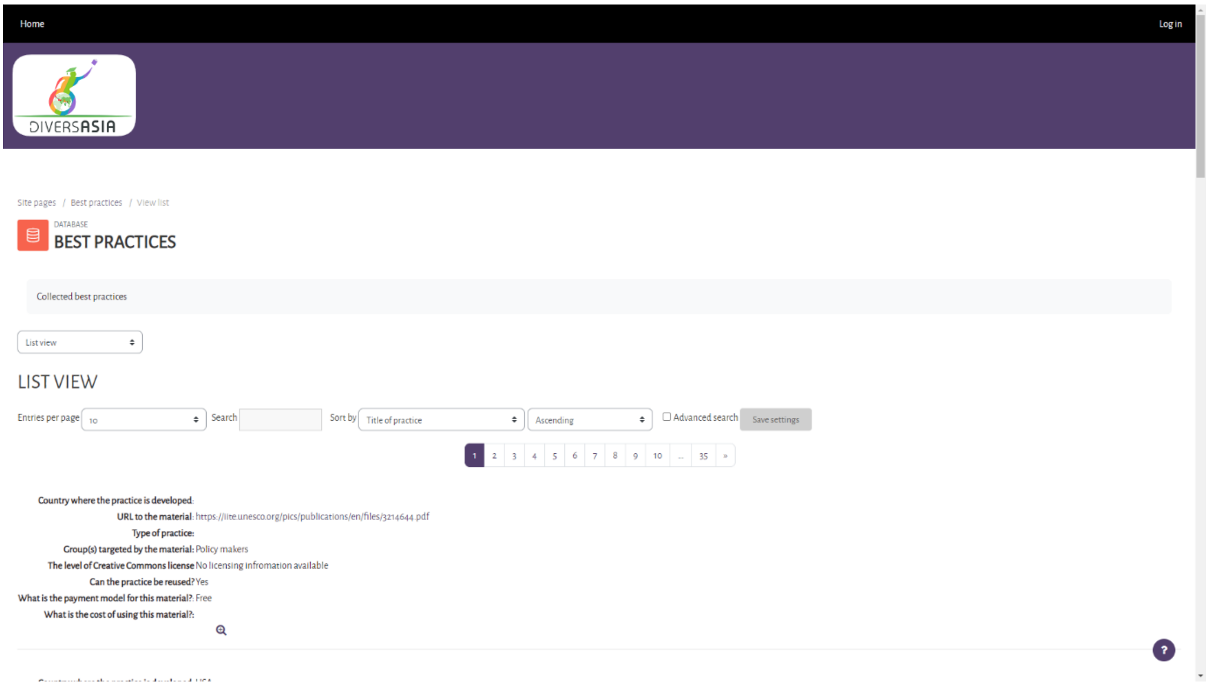


Figure 3: DiversAsia best practices database, integrated in DiversAsia portal

2.2.2 MyHub - European Inclusive Learning Hub

The European Inclusive Learning Hub is a large repository of inclusive, searchable and usable learning resources, methods, tools and mechanisms, already tested, approbated and deployed in certain countries. It acts as a wizard-style driven and semantically supported logical framework for systematization/categorisation of inclusive learning resources, methods, tools and mechanisms in various education levels: early childhood, primary & secondary school, vocational education and training, higher education, adult education.

This inclusive education database was created by the Erasmus+ project "MyHUB: a one-stop-shop on inclusion practices, tools, resources and methods for the pedagogical staff at formal and non-formal educational institutions" (MyHUB – 604454-EPP-1-2018-1-LV-EPPKA3-IPI-SOC-IN, https://www.inclusion-hub.eu/), and has been integrated into our own portal as it offers additional information which can be very useful for our target groups.

226 inclusive educational practices are as such also shared with the portal.

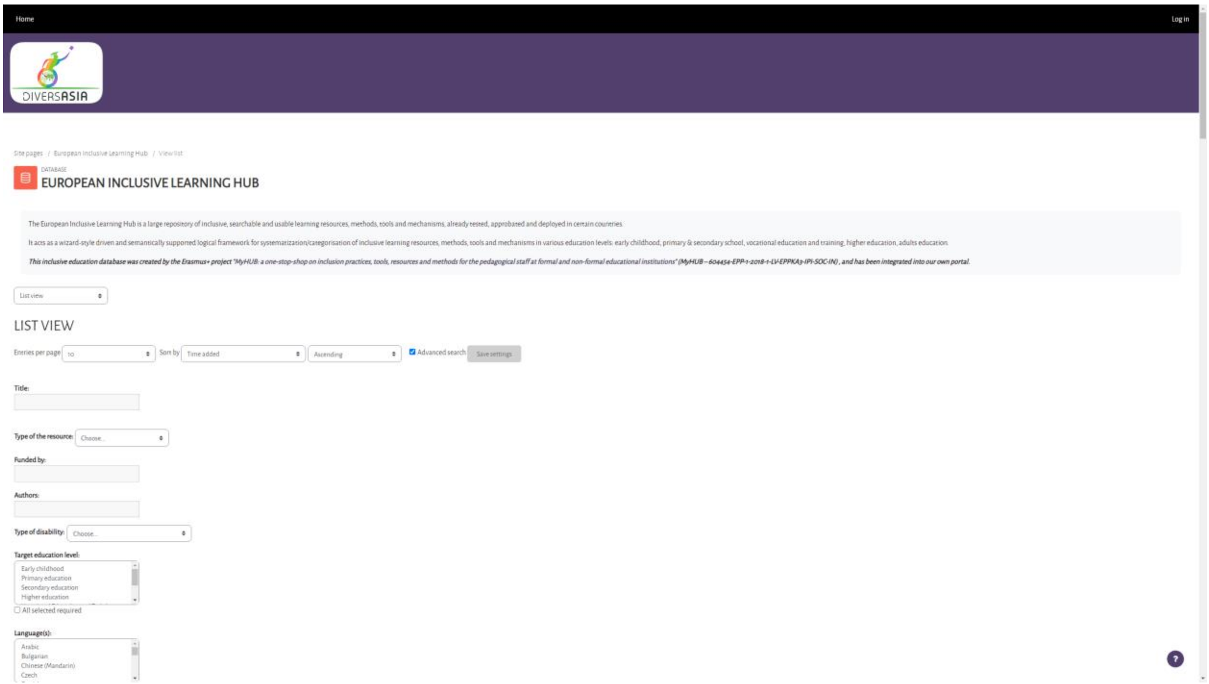


Figure 4: MyHub - European Inclusive Learning Hub, integrated in DiversAsia portal

2.3 Self-assessment

A self-assessment plugin has been developed for Moodle and will host the inclusive education self-assessment checklists for both administrators within a HEI as well as content creators (such as teaching staff), thus facilitating the readiness assessment with regards to inclusive higher education.

These checklists are being prepared at the moment and will form an intrinsic part of the DiversAsia toolkit.

2.4 H5P integration

H5P (https://h5p.org/) is an abbreviation for HTML5 Package. H5P aims to make it easy for anyone to create interactive content such as presentations, videos and other multimedia, questions, quizzes, games and more without the need of programming skills. H5P makes it easy to create interactive content by providing a range of content types for various needs.

Moodle allows for an easy H5P integration by installing the H5P plugin (see https://h5p.org/moodle), after which content creation can go ahead, but this tome providing an engaging interactive learning experience that brings together text, audio, video, etc.

What makes it even more appealing is that once the H5P content is created, you can easily download it and upload/integrate it again to multiple platforms (to another LMS like Opigno, or to CMS's such as WordPress, Drupal, etc.).

The choice for H5P is logical since we want to make the training material as engaging and interactive as possible.

2.5 The DiversAsia Moodle Course Content

The partnership agreed on the following topics that are provided in the toolkit embedded in the portal, and fully described in D2.2:

- Policy
    - Institutional policy and strategy in India and Bangladesh
- Teaching
    - Accessibility tools: authoring, conversion, evaluation
    - Accessibility: Designing and Teaching Courses for All Learners (free MOOC course)
    - Assistive technologies (AT)
    - Co-creation and sharing/personalisation of content, disruptive technologies
    - Creation and validation of accessible OERs and MOOCs
    - Database with guidelines based on disability, teaching content format, type of training
    - Guidelines and AI Tools for Personalisation
    - Handy checklists for all HEI staff
    - Inclusive Higher Education SWOT Analysis
    - Inclusive pedagogical approaches
    - Inclusive teaching practices
    - Make content accessible yet attractive for young students (usage H5P for content delivery)
    - Prepare accessible source material

    - Universal Design for Learning (UDL)
- Administration
    - Accessibility at HEI
    - Audit the physical accessibility of HE premises
    - Audit the physical accessibility of HE premises
    - Inclusion readiness assessment and reflection tool for HE
    - Streamline administrative processes that accommodate the needs of students (and staff) with disabilities
    - Training and professional development
    - Understanding the barriers and difficulties students face
- Technical Staff
    - Assistive technologies (AT)
    - Audit the physical accessibility of HE premises
    - Handy checklists for all HEI staff
    - Inclusion readiness assessment and reflection tool for HE
    - Make content accessible yet attractive for young students (usage H5P for content delivery)
- Librarians
    - Accessibility at HEI
    - Handy checklists for all HEI staff
    - Inclusion readiness assessment and reflection tool for HE
    - Prepare accessible source material
    - Universal Design for Learning (UDL)
- Supporting videos in Tamil
- Glossary
- General Useful Resources

The content is fully described in the Document D2.2.

2.6 Personae Driven Content Access

In previous projects we have successfully employed a personae driven approach which leads us to create some imaginary examples of students with different learning needs which may be present in the context of the work. This seemed a sensible approach here to be able to present information that is relevant to particular target users in Higher Education. Also, it was clear that different staff members within the university would require access to different types of information and tools. This was therefore also considered within the course content development.

Moodle enables a branching hierarchy approach (Figure 5) which gives a tidy route into accessing the information the user actually wants. The content was developed and labelled according to the interest groups by the following criteria:

- University role
    - Staff

        - Teaching
        - Administrative
        - Library
        - Technical
        - Student support
        - Policymaker
    - Student
        - With a disability
            - Vision
            - Hearing
            - Mobility
            - Mental Processing
            - Mental Health
            - Neurodiversity
        - With mental health concerns
        - Interested in accessibility.
    - An interested third party

Courses were then curated that address the particular needs of each branch of the tree covering relevant topics and grouping them by further categories within the courses themselves.

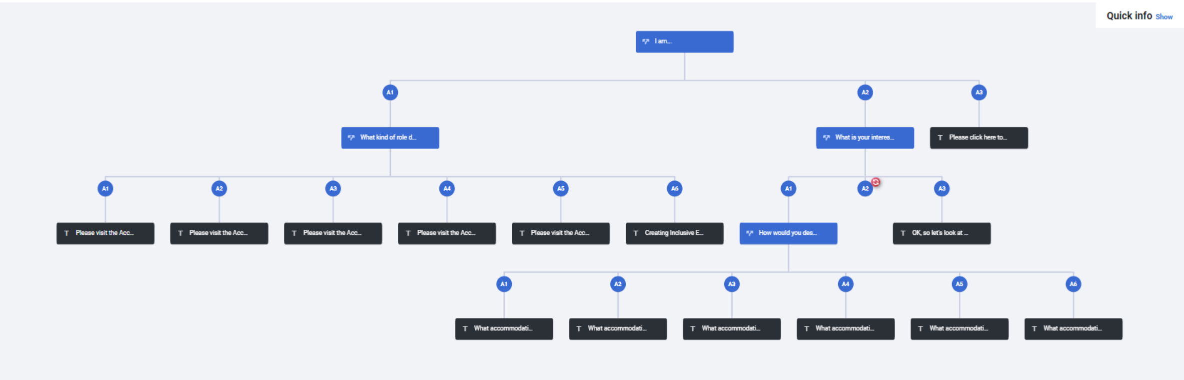


Figure 5 – The branching structure within Moodle

2.7 Accessing the Moodle Courses Online

To access the courses users should navigate to https://diversasia-accessible-he.eu/elearning/ and click ‘Begin’ on the Introducing Accessibility tool (Figure 6).

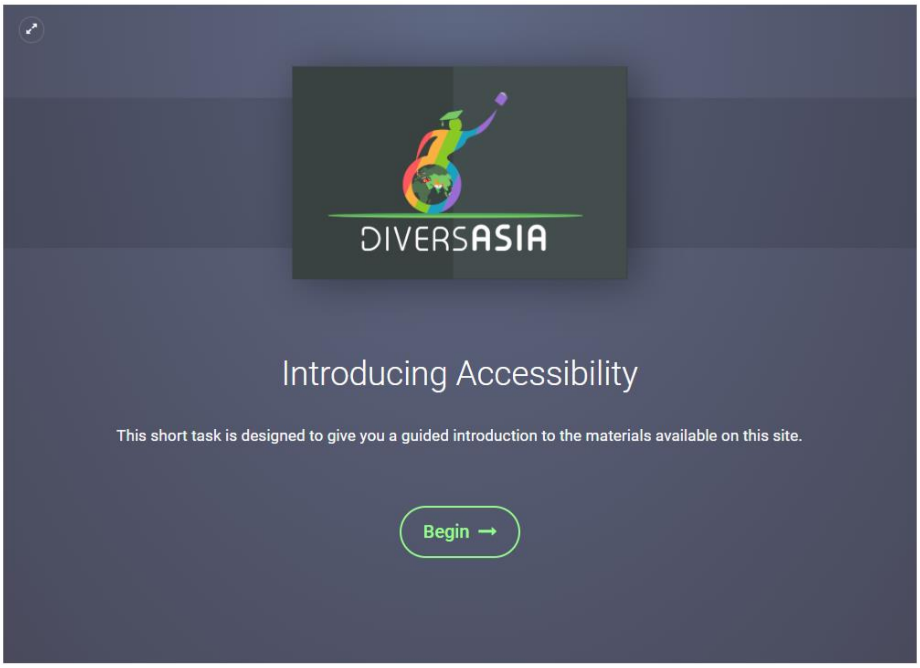


Figure 6 – The Introducing Accessibility tool. The gateway to the DiversAsia learning courses

This tool will ask you about your interests, staff role and reason for accessing the site and guide you towards the best course materials for you.

2.8 Moodle mobile app framework (GitHub)

The Moodle mobile app GitHub framework (https://github.com/moodlehq/moodleapp, we use V3.9.5) allows for the development of the native Android and iOS mobile applications, where we can integrate our own developed plugins (see e.g. the best practices database). Using secure web services that are exposed from the portal, the mobile application then invokes the web services of the platform, and the returned results are presented in the mobile applications. This approach ensures full synchronisation in terms of user profiles, content and services on the portal. Furthermore, the framework used is OSS, and ensures with its large community easy updates and support.

3 DiversAsia tools

3.1 Integrated SCORM accessibility validator Pa11y

Pa11y (https://pa11y.org/) is an open-source tool we integrated to create the accessibility validator for SCORM packages. Pa11y includes and implements standards from 2 checkers, namely aXe and HTML CodeSniffer. Pa11y uses Puppeteer to run its own headless Chrome browser. Axe-core is an automated accessibility testing tool.

Pa11y uses HTML Code Sniffer as its default runner, but it can also run axe-core at the same time if you tell it to. HTML_CodeSniffer comes with standards that enforce the three conformance levels of the Web Content Accessibility Guidelines (WCAG) 2.1, and the web-related components of the U.S. "Section 508" legislation.

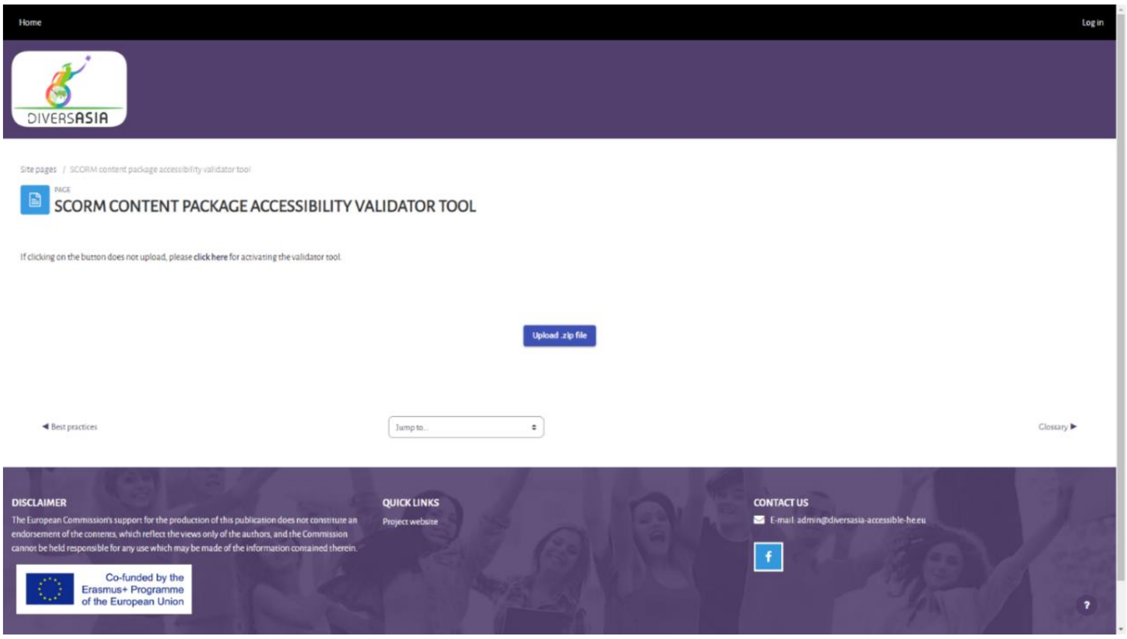
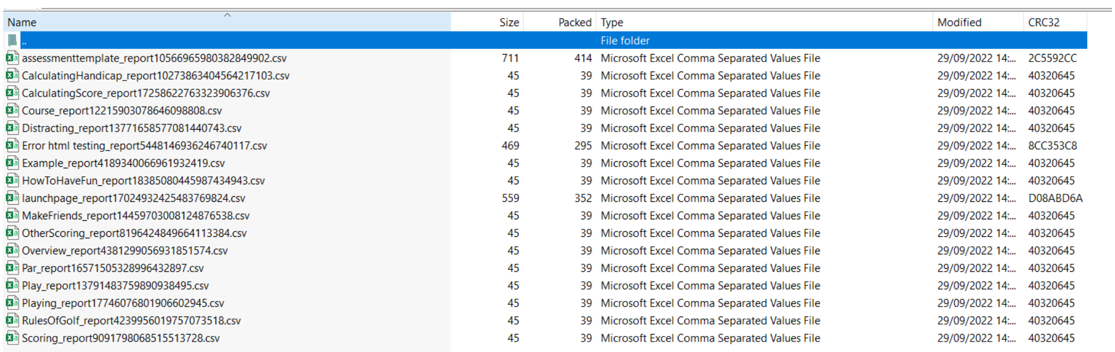


Figure 7: Uploading a package



Name	Size	Packed	Type	Modified	CRC32
..			File folder		
assessmenttemplate_report10566965980382849902.csv	711	414	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	2C5592CC
CalculatingHandicap_report10273863404564217103.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645
CalculatingScore_report17258622763323906376.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645
Course_report12215903078646098808.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645
Distracting_report13771658577081440743.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645
Error html testing_report5448146936246740117.csv	469	295	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	8CC353C8
Example_report4189340066961932419.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645
HowToHaveFun_report18385080445987434943.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645
launchpage_report17024932425483769824.csv	559	352	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	D08ABD6A
MakeFriends_report14459703008124876538.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645
OtherScoring_report8196424849664113384.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645
Overview_report4381299056931851574.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645
Par_report16571505328996432897.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645
Play_report13791483759890938495.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645
Playing_report17746076801906602945.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645
RulesOfGolf_report4239956019757073518.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645
Scoring_report9091798068515513728.csv	45	39	Microsoft Excel Comma Separated Values File	29/09/2022 14:...	40320645

Figure 8: Receive feedback per HTML page in package

	A	B	C	D	E
1	type	code	message	context	selector
2	error	WCAG2AA.Principle2	A title should be provided for the document, using a non-empty title element in the head section.	<head>\n <!--title>How to Have Fun<...</head>	html > head
3	error	WCAG2AA.Principle1	Img element missing an alt attribute. Use the alt attribute to specify a short text alternative.	<img src=\fun.jpg\">"	html > body > img
4					

Figure 9: Detailed analysis of notices, warnings, errors

3.2 DiversAsia native mobile applications for iOS and Android

Using the Moodle mobile framework (https://github.com/moodlehq/moodleapp), we customised this to accommodate the needs of the DiversAsia portal, and realised a native mobile iOS and Android app, which, using secure webservices, is synchronised with the online portal based on Moodle. The advantage is that any content update of these 2 mobile applications can be performed via the WYSIWYG editor functionality of Moodle, thus ensuring no developer or designer is needed whenever content needs to be updated.

These apps are in the process of approval in the App Store (iOS) and Google Play (Android), and will be announced also via the project website https://diversasia-accessible-he.eu/.

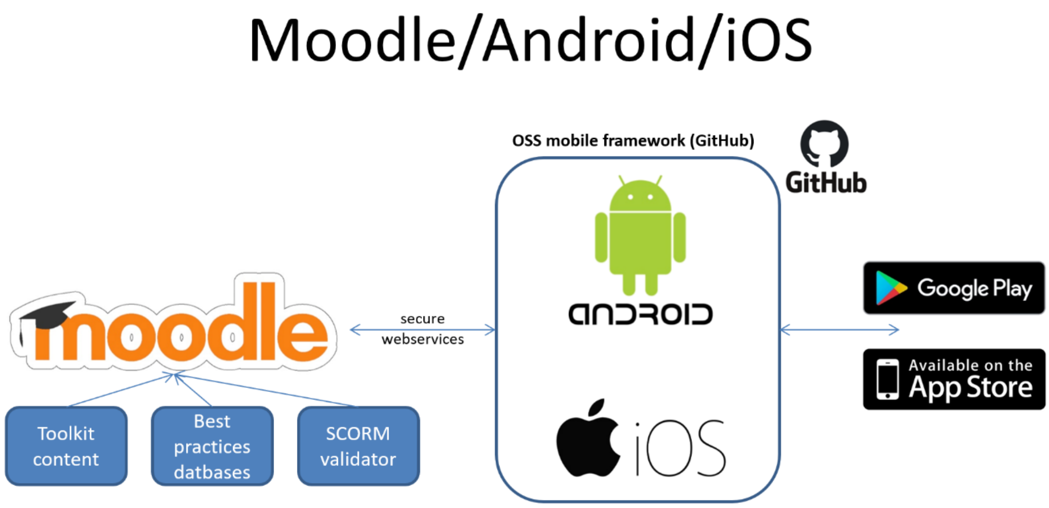


Figure 10: Android/iOS mobile framework deployment

Download link Android app (Google Play):
https://play.google.com/store/apps/details?id=com.diversasia.mobile

Download link iOS app (App Store):
https://apps.apple.com/us/app/diversasia/id6444756230

These links are also integrated onto the project website:

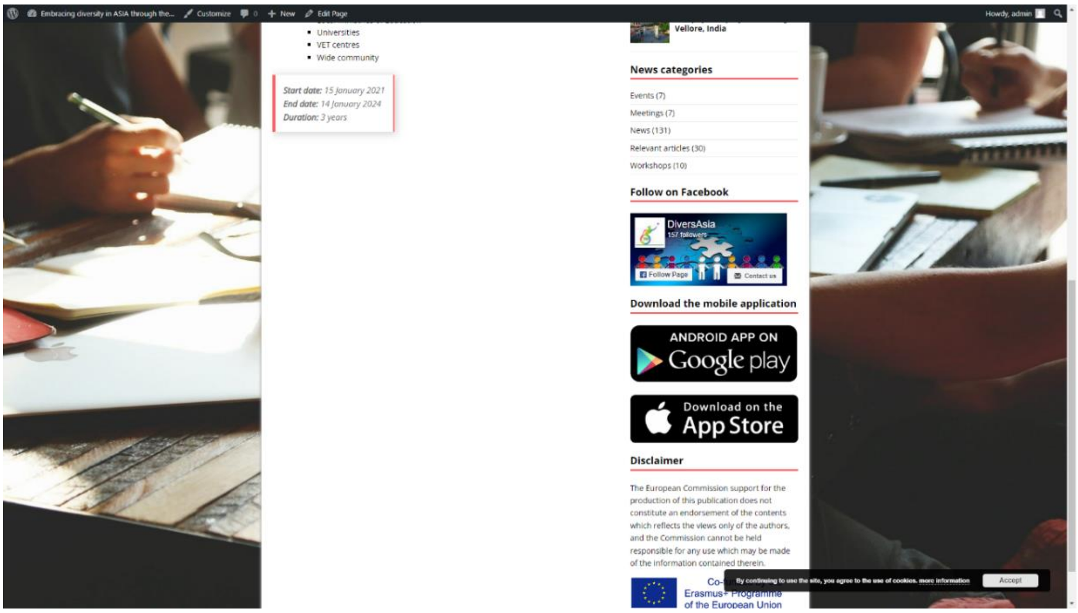


Figure 11: Android/iOS mobile apps download links on project website

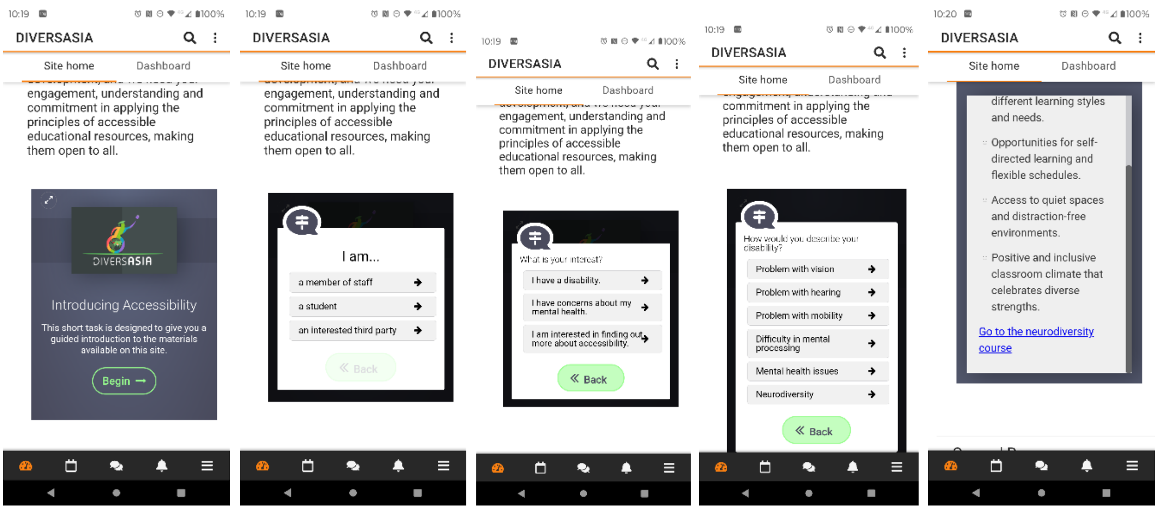


Figure 12 - A typical use of the app to access relevant content (left to right)

Figure 12 shows a set of screen captures from the Android version of the DiversAsia Moodle app. The user starts at the home page (left) and is directed through the content selecting their user type, their reason for interest, their particular disability interest and are then guided to the correct course materials to suit their needs. The course content has been built up from the comprehensive set of H5P modules developed in the project and described fully in D2.2, and each course was created by curation of relevant H5P modules for each use case in the branching tree structure, thus creating course content tailored to specific users, personae and interest as described previously in section 2.6.

3.3 DiversAsia Engagement App - previous work

This tool uses the devices present in a standard laptop (web cam, keyboard and mouse) to assess the engagement of the user seated if front of the device. A USB web camera may be used where such a device is not present.

Engagement is the single most important factor in the education of students with a learning disability and autism – without engagement there is no deep learning or meaningful outcome: Carpenter (2010) states: “Without [engagement], there is no deep learning, effective teaching, meaningful outcome, real attainment or quality progress”.

Teachers commonly use the ‘Engagement Profile and Scale’ (Carpenter et al., 2016) to measure engagement in the classroom.

Our aim is to automate the measurement of engagement. We have developed approaches to this goal in 2 previous EU Projects: the EU H2020 MaTHiSiS project and the EU Erasmus Pathway+ project.

3.3.1.1 The MaTHiSiS Project: adaptive learning for students with disabilities.

Recent reviews have highlighted an explosion of work on the use of artificial intelligence tools for education (AIEd) that detect affective states relevant to learning. The MaTHiSiS ecosystem (Standen et al, 2020) used both affective state and learning achievement to drive the presentation of learning material in an adaptive learning system. The ecosystem comprised mainstream and special schools, and a library of learning materials was developed, according to context, with teachers from the different schools. These learning materials could be displayed on a variety of platform agents - computer screen, an Android tablet, or a NAO robot.

Teachers and trained researchers, familiar with the students, annotated recordings of them working on educational materials on a variety of platforms (desktop, mobile devices and educational robots). These labels were used to train machine learning algorithms for a range of modalities, including for example:

- Facial expressions (pre-trained and subsequently fine-tuned C3D Convolutional Neural Network (CNN) model to infer the emotions from face images).
- Eye gaze estimation (two-stream CNN using 3D gaze vectors).

- Body pose (Speed Relation Preserving Slow Feature Analysis algorithm to extract features classified by a SVM model).
- Also, voice, gesture and interaction tracking.

The performance of each were based on k-fold cross-validation for the majority of the modalities and leave-one-out cross-validation for the interaction parameters classifier.

An equally-weighted late multimodal fusion scheme using the predictions independently inferred by each modality was employed to give an overall understanding of the affective state of each learner.

The materials presented to the learner were adapted in response to changes in affective state, with the goal of keeping the student in a state of engagement, or flow.

3.3.1.2 Pathway+: personalised learning for students with disabilities.

Rather than using teacher labelling of student state of affect to provide training labels, in Pathway+ outcomes from a Continuous Performance Test were used to indicate level of engagement (Taheri et al., 2020).

Here, we need to label the data with objective measures of engagement so that in the future, new unlabelled data can be presented to the model and it can still predict users' levels of engagement. Swanson's Signal Detection Theory (1981) describes how we can label multimodal sensor data (eye gaze, facial expressions, EEG, body posture) using a Continuous Performance Test (CPT). CPT Outcome Measures allow us to segment multimodal data into regions of high and low attention. This segmentation provides the labels by which we can supervise the machine learning method for the data modelling engagement, boredom and frustration. The method used is Random Forest, an ensemble learning method for classification for classification.

Learning challenge can be increased, or learning materials changed if boredom or frustration are detected and displayed using the 'Traffic Light' system.

3.3.1.3 Our current approach to explainable AI for engagement assessment

Building on our previous experience, we have been working on Explainable Multimodal Machine Learning for Engagement Analysis by Continuous Performance Test (Rahman et al., 2022). Here we have developed decision trees from an

Explainable Artificial Intelligence (XAI) model for a continuous performance test and built on the explainable Random Forest for decision making process.

Multi-sensor data and multimodal machine learning, for engagement analysis has been used. We considered body pose, eye gaze, interaction data and facial features by objective labelling of engagement or disengagement for cognitive attention of a Seek-X type task execution. We used decision trees, an XAI algorithm, to visualize the decision which will help us assess the accuracy of the model intuitively and provide us with the explainability of engagement or disengagement for visual interactions.

The accuracy of the model does not give the best possible results, but helps decision making - and it is important that this model is more explainable than the black box-like algorithms of ML.

3.3.1.4 The DiversAsia user engagement detection app

The importance of teaching within the academic ecosystem is truly crucial in allowing students to improve their learning. Students with neurodevelopmental, cognitive, physical and sensory disabilities need special care as their learning needs and requirements differ from those without. Within this context, it is very important that the actual learning situation of these students with disabilities in a classroom can be assessed in real time, so that relevant and appropriate interventions and support can be provided to improve their learning experience, and hence the learning itself. This can be done intelligently using cutting-edge technologies, thanks to the recent developments in artificial intelligence and machine learning as discussed above.

3.3.1.4.1 Real-time Video Analysis for Affective State Detection

The system utilises Attention Mesh (Grishchenko et al., 2020), a lightweight architecture for 3D face mesh prediction that uses attention to semantically meaningful regions. This ground-breaking neural network model is designed for real-time on-device inference and runs at over 50 FPS on a smartphone, such as a Pixel2. The model enables applications like AR makeup, eye tracking and AR puppeteering that rely on highly accurate landmarks for eye and lips regions. Figure 13 shows that different facial landmarks (e.g., eyes, lips, eyebrows, and face) can be clearly detected and identified using this model. The attention mesh model uses a unified network architecture that achieves the same accuracy on facial landmarks as a multistage cascaded approach while being 30 per cent faster, making it suitable for running on commodity portable devices.

Figure 13 Salient contours predicted by Attention Mesh (Grishchenko et al., 2020)

As described in Figure 14, the model will accept a 256 × 256 image as input. This image will be provided by either the face detector or via tracking from a previous frame. After extracting a 64 × 64 feature map, the model splits into several sub-models (see Figure 14). One sub-model predicts all 478 face mesh landmarks in 3D and defines crop bounds for each region of interest (i.e., the face in this case). The remaining sub-models predict region landmarks from the corresponding 24 × 24 feature maps that are obtained via the attention mechanism. We will initially concentrate on three facial regions with key contours: the lips and two eyes. Each eye sub-model predicts the iris as a separate output after reaching the spatial resolution of 4x4. This allows the reuse of eye features while keeping the dynamic iris independent from the more static eye landmarks.

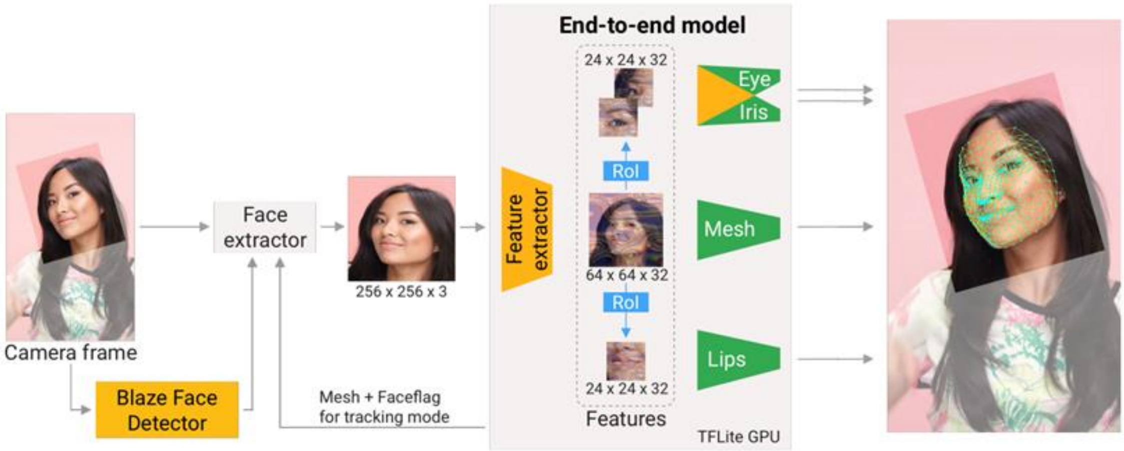


Figure 14 The inference pipeline and the model architecture overview (Grishchenko et al., 2020)

The attention mechanism pulls out visual features of a given region of interest by sampling a grid of 2D points in the feature space and extracting the features under the sampled points. This allows the network architectures to be trained end-to-end and to enrich the features that are used by the attention mechanism. We use Convolutional Neural Networks (CNNs) for the identification of specific features. Though CNNs define an exceptionally powerful class of models, they are still limited by the lack of ability to be spatially invariant to the input data in a computationally and parameter-efficient manner. To overcome this, we specifically use a spatial transformer module, which allows transformation operations like zoom, rotate, translate, and skew on the sampled grid of points.

The classical version of the Spatial Transformer Network (STN) contains three main modules: (i) the localization network, (ii) the grid generator, and (iii) the sampler.

(i) The localization network receives the input image and outputs the parameters that form the transformation matrix. These parameters vary depending on the input image and change during the training because they usually correspond to the weights extracted from a fully connected layer. This module is in charge of detecting the most relevant regions to focus on the image for identifying the class in later layers.

(ii) The grid generator creates a regular grid of a certain size and receives the parameters of the localization network. These are the grid values received from the attention mesh. An affine transformation is applied on the grid, allowing us to perform a mapping that lets us infer a relationship between the coordinates of the pixels in the input image and the expected position of the transformed pixels in the output image.

(iii) Finally, the sampler module applies the spatial transformation by taking the sampling grid and the input image to produce the transformed image.

The transformed version of the mask generated at the output of the sampler passes to the classification network, so this CNN extracts information from the patches or weighted images and identify the Facial Emotion Recognition task, as shown in Figure 15.

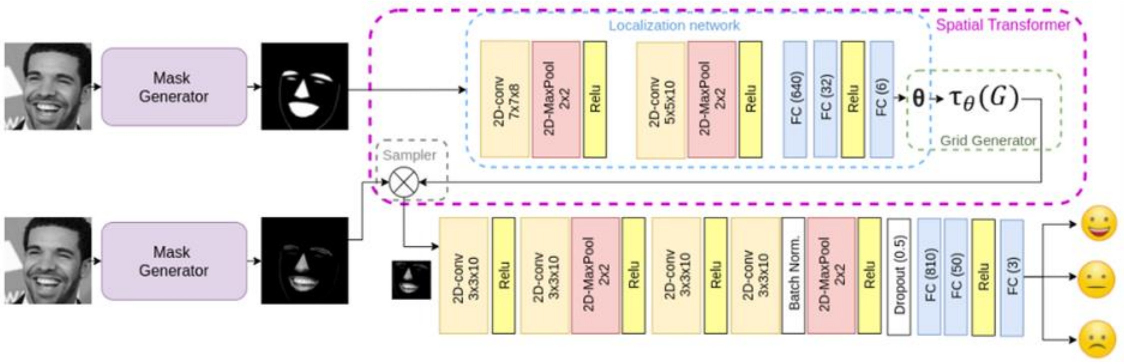


Figure 15 Spatial Transformer CNN architecture. Example for landmarks-based facial patch detection (Luna-Jiménez et al., 2021).

3.3.1.4.2 Real-time Interaction Data Analysis to Detect User Learning Situation

The main source of the user interaction data is the timestamp and the location of the mouse cursor on the learning page, which the learner will be learning from (Miller et al., 2014). These data are analysed to understand students' learning situations. The literature suggests that eye-movement data provides useful information through the duration a learner looks at a paragraph of text and images to indicate what they are thinking about those materials (Zhang et al., 2020). In our system, our assumption is that the longer students remain at a certain vertical position on the page, the longer they think about it. Students' quick scrolling of a content page means a lack of cognitive engagement in the study content. Using research on synchronization between mouse movement and eye movement, we can understand student engagement with the content page by recording the position of the mouse cursor and the corresponding timestamp.

The information related to the facial expression and the time spent on each section of the presented learning material are then fused to get the engagement of the students.

A description of the very latest version of the DiversAsia Engagement App “LEMate” follows.

3.4 The DiversAsia Engagement Detection Tool

3.4.1 Introduction

Learner engagement is a critical factor influencing for teaching and learning effectiveness. Engagement can be assessed using subjective observations or self-reporting, but accurate and real-time insights are near impossible to achieve. In DiversAsia, we developed a system using an Artificial Intelligence Powered Learner Engagement Detection System, called LEMate. The system uses machine learning algorithms to analyse a video stream to detect including facial expressions, eye gaze, heart rate along with interaction patterns with computers (keyboard and mouse), to determine the engagement state of the learner. The system returns predicted engagement states (engaged, bored, and frustrated) and displays this using a traffic light system.

Early trials were held in Bangladesh, where learner engagement states were predicted demonstrating the potential of the system to track user engagement in real time, allowing early interventions to change learning materials or methods to reengage the learners.

The LEMate prototype is designed specifically for low resource classrooms deeming it useful for developing countries such as Bangladesh and India where classrooms are often busy and poorly resourced. The system will continue to be developed and refined in order to improve system accuracy. It will also be tested in other domains, and with other users displaying non-neurotypical behaviours.

The DiversAsia Engagement App system detects engagement using four components which come together as described in:

- interactions with computer (mouse, keyboard and trackpad)
- eye tracking
- emotion detection
- heart rate.

The video feed which is captured through the standard webcam is used to track eye gaze, to detect facial expressions and also to estimate heart rate. The data from these four streams is fused through a rule-base system to accurately detect learner engagement.

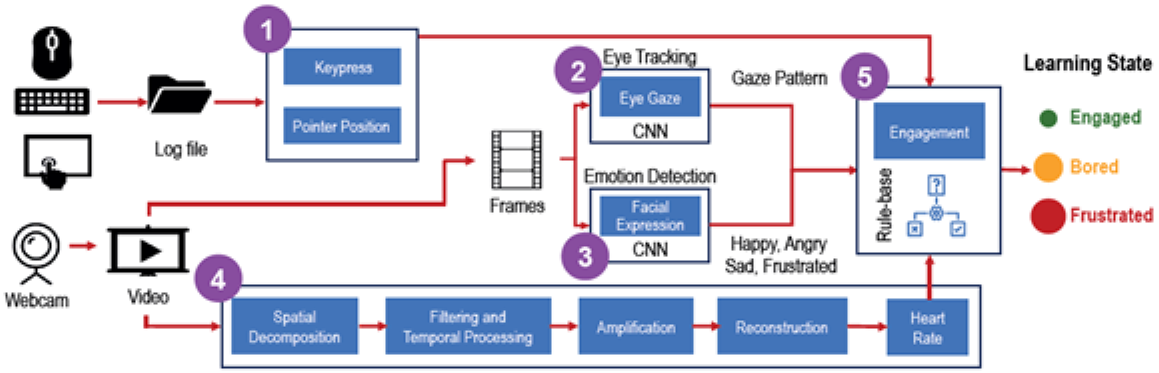


Figure 16 – Flowchart of the Engagement App System

3.4.2 The Data Streams Used

There follows a summary of the component data streams:

- Interaction with the Computer

The data collected for interaction detection is calculated from analysis of mouse movements (or track pad movements) and keyboard key presses. It should be noted that the system does no log key presses, but simply analyses their frequency). The data sources are combined using Entropy to estimate the learner's overall interaction with the computer.

- Eye Tracking

The eye tracking part of the system uses [OpenCV] (an open-source computer vision library) and [Mediapipe] (a Google developed open source project which builds pipelines for processing media-based data). Eye gaze information can be extracted from facial video data feed captured using a standard (or built-in) webcam. Mediapipe contains a built in Convolutional Neural Network (CNN) model specifically for this purpose. The movement of the gaze is tracked by identifying eye positions and estimating the focus. Integration of distance and speed between focus points are integrated in real time giving another measure of estimated engagement with learning materials.

- Emotion Detection

The real-time emotion detection system also uses the webcam video feed. It can extract facial expressions from the feed and classify them into one of seven emotional states as described below:

1. anger,
2. disgust,
3. fear,
4. happiness,

5. sadness,
6. surprise
7. neutral.

Another real-time CNN model is implemented using Python, Keras and OpenCV libraries to identify and predict emotional states from the facial expressions in the video feed as detected using the Haar cascade classifier. The emotional state and changes in the emotional state detected are then is used as an additional predictor for engagement.

- Heart Rate Estimation

This is estimated by detecting minute facial skin colour changes during heartbeats. There are three main steps: identification of the region of interest (i.e. the best exposed skin to use for the colour change detection), pulse signal extraction and heart rate calculation. This system uses Face Mesh for face detection, and extracts the average green channel of a region of the lower face as the pulse signal. You can read more about this process in this paper by Li et al. Changes in heart rate can be strong indicators of arousal and engagement.

- Learner Engagement Data Fusion

We developed an integrated model that then combines knowledge gathered from all of the data streams mentioned above. The learner engagement, or learning state, is conveyed as a traffic light system with green (engaged), amber (bored) and red (frustrated). The circle sizes vary so that colour is not the only indicator. This is an accessibility feature. The learning state is estimated through a set of simple rules based on dynamic threshold values corresponding to the different components of the system across a rolling 5 second window of data. The dynamic threshold values along with the predicted emotion state, form a rule base which is used to estimate the final learning state (i.e., engaged, bored, frustrated).

3.4.3 System Setup and Use

The latest version of the system can be downloaded from this github repository. It should then be extracted and run in the Windows Operating System (10 or 11). Run the LEmate.exe file to start the application's Home Page.

To start a session, you must agree to the terms and conditions of us, and press start (Figure 17), The system will ask the user to input the student ID, student name and class ID. After successful input, the session will start automatically.

In the first 3 minutes, the system is calibrated (Figure 18) according to the user's personalised environment and it then
provides real-time learning states every 5 seconds (Figure 19). During this period, the user should continue their normal learning activities related to the session. To close the system, the user needs to activate the LEmate window and press the escape

(ESC) button. The system will then display the learning states collected to the user (Figure 20).

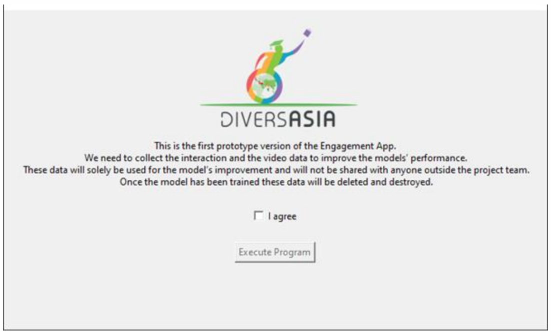


Figure 17 – Splash screen

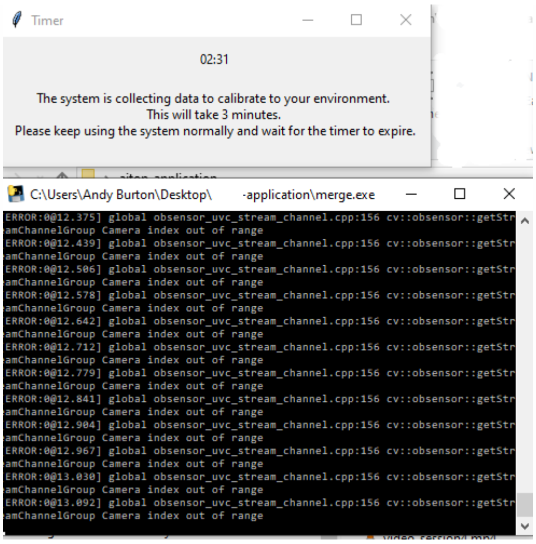


Figure 18 Calibration

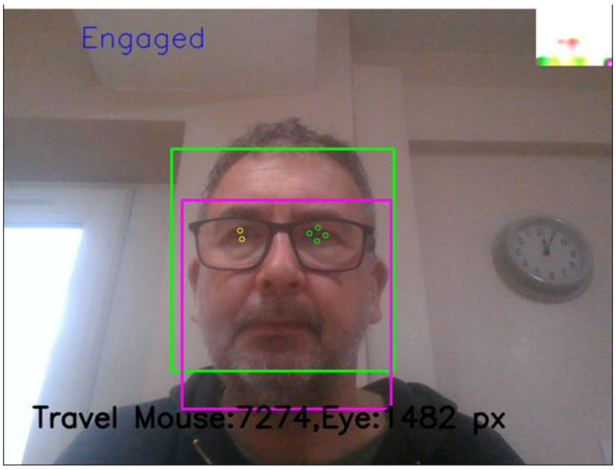


Figure 19 The App running on a Windows Laptop

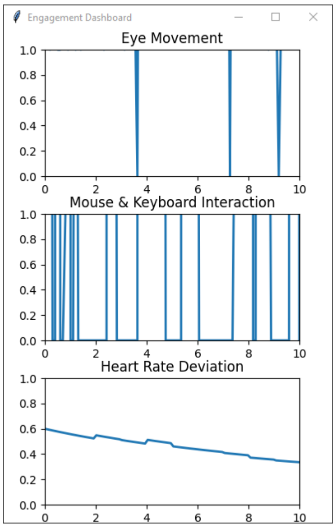


Figure 20 – The Results as displayed following a run of the App

4 Conclusion

The DiversAsia portal was fully launched in 2023, and brings together all developed content as collected in D2.2 DiversAsia Toolkit.

Full learning content and courses are now in place on the technical infrastructure set up earlier in the project, including on the project website, and on the previously described Moodle learning environment accessible both online and through the developed apps for both Android and iPhone.

5 References

Grishchenko, I., Ablavatski, A., Kartynnik, Y., Raveendran, K. and Grundmann, M., 2020. Attention mesh: High-fidelity face mesh prediction in real-time. arXiv preprint arXiv:2006.10962.

Luna-Jiménez, C., Cristóbal-Martín, J., Kleinlein, R., Gil-Martín, M., Moya, J.M., Fernández-Martínez, F., 2021. Guided Spatial Transformers for Facial Expression Recognition. Applied Sciences 11, 7217. https://doi.org/10.3390/app11167217https://doi.org/10.3390/app11167217

Miller B. W., Anderson R. C., Morris J., Lin T., Jadallah M., Sun J. (2014) The effects of reading to prepare for argumentative discussion on cognitive engagement and conceptual growth. Learning and Instruction 33: 67–80. doi:10.1016/j.learninstruc.2014.04.003.

Standen, PJ; David J. Brown; Mohammad Taheri; Maria J. Galvez Trigo; Helen Boulton, Andrew Burton; Madeline J. Hallewell; James G. Lathe; Nicholas Shopland; Maria A. Blanco Gonzalez; Gosia M. Kwiatkowska; Elena Milli; Stefano Cobello; Annaleda Mazzucato; Marco Traversi; Enrique Hortal; An evaluation of an adaptive learning system based on multimodal affect recognition for learners with intellectual disabilities; BJET Volume51, Issue5, September 2020, Pages 1748-1765 First published: 29 July 2020 https://doi.org/10.1111/bjet.13010

Zhang, Z., Li, Z., Liu, H., Cao, T. and Liu, S., 2020. Data-driven online learning engagement detection via facial expression and mouse behavior recognition technology. Journal of Educational Computing Research, 58(1), pp.63-86.