Jorge Miguel da Silva Calado

Licenciado em Engenharia Electrotécnica e de Computadores

A Framework for Students Profile Detection

Dissertação para Obtenção do Grau de Mestre em

Engenharia Electrotécnica e de Computadores

Orientador: Ricardo Gonçalves, Professor Associado com Agregação, FCT/UNL

Co-orientador: João Sarraipa, Professor Auxiliar Convidado, UNINOVA

Júri:

Presidente: Prof. Doutora Maria Helena Silva Fino

Arguentes: Prof. Doutora Anabela Monteiro Gonçalves Pronto

Vogais: Prof. Doutor João Filipe dos Santos Sarraipa

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Março de 2017

A Framework for Students Profile Detection

Dedicated to my family, for the continuous support given throughout this journey.

Acknowledgements

Thank you, Professor João Sarraipa, for the all the invaluable support, mentoring and guidance provided.

A word of appreciation is also due to Fernando Ferreira, Andreia Artifice, and Professor Ricardo Gonçalves for the assistance regarding this dissertation work and its integration in the ACACIA project, and to Educational Psychologist, Dra. Maria de Lourdes Mata, for the expertise given in the consultation of the developed framework.

I would also like to thank my friends, Daniel, Zé, Marta, Telma, Tiago, David, Teresa, Rita and André C., who volunteer as tributes to test and review this project, and to André S. who lent the webcam.

Lastly, and because without their continuous support none of this would be possible, thank you mom and dad for always believing in me, and thank you sister, for enduring me (mostly) all these years, and for lending me your expertise in educational psychology.

Abstract

Some of the biggest problems tackling Higher Education Institutions are students' drop-out and academic disengagement. Physical or psychological disabilities, social-economic or academic marginalization, and emotional and affective problems, are some of the factors that can lead to it.

This problematic is worsened by the shortage of educational resources, that can bridge the communication gap between the faculty staff and the affective needs of these students.

This dissertation focus in the development of a framework, capable of collecting analytic data, from an array of emotions, affects and behaviours, acquired either by human observations, like a teacher in a classroom or a psychologist, or by electronic sensors and automatic analysis software, such as eye tracking devices, emotion detection through facial expression recognition software, automatic gait and posture detection, and others.

The framework establishes the guidance to compile the gathered data in an ontology, to enable the extraction of patterns outliers via machine learning, which assist the profiling of students in critical situations, like disengagement, attention deficit, drop-out, and other sociological issues.

Consequently, it is possible to set real-time alerts when these profiles conditions are detected, so that appropriate experts could verify the situation and employ effective procedures.

The goal is that, by providing insightful real-time cognitive data and facilitating the profiling of the students' problems, a faster personalized response to help the student is enabled, allowing academic performance improvements.

Keywords: Student Engagement; Automatic Emotion Detection; Affective Computing; Eye Tracking; Facial Expression Recognition;

Resumo

Alguns dos maiores problemas que assolam as Instituições de Ensino Superior são as desistências e a falta de motivação por parte dos estudantes. As incapacidades físicas ou psicológicas, marginalização académica ou socioeconómica, e problemas emocionais e afectivos, são alguns factores que podem contribuir para esses problemas.

Esta problemática é agravada pela falta de recursos educacionais, que consigam ligar as falhas na relação entre os professores e as necessidades afectivas destes estudantes.

Esta dissertação centra-se no desenvolvimento de uma plataforma, capaz de capturar dados analíticos, de uma selecção de emoções, afectos e comportamentos, adquiridos por observação humana, como por exemplo por um professor numa sala de aula ou por um psicólogo, ou por sensores electrónicos e software de análise automática, tais como dispositivos de *eye tracking*, software de detecção de emoções através de reconhecimento de expressões faciais, detecção automática do andar e da postura, entre outros.

A plataforma estabelece uma orientação para compilar a informação recolhida numa ontologia, para permitir a detecção de *outliers* dos padrões, através de técnicas de *machine learning*, para assistir à identificação de perfis de estudantes em situações críticas, tais como falta de motivação, falta de atenção, desistências e outros problemas sociológicos.

Consequentemente, é possível despoletar alertas em tempo real, quando estas condições dos perfis forem detectadas, para que os especialistas adequados possam verificar a situação e utilizar procedimentos eficientes.

O objectivo é, através da disponibilização de dados cognitivos em tempo real e da facilitação da detecção dos problemas dos estudantes, que seja possível uma resposta rápida e personalizada na ajuda ao estudante, possibilitando melhorias no desempenho escolar.

Palavras-chave: Motivação Académica; Detecção Automática de Emoções; Computação Afectiva; *Eye Tracking*; Reconhecimento de Expressões Faciais;

Contents Index

Figures Index

Tables Index

List of Abbreviations

3D – Three Dimensions

ACACIA – Apoyo, Refugio, Comunicación, Adaptación, Innovación y Enriquecimiento (Support, Refuge, Communication, Adaption, Innovation and Enrichment)

AdeLE – Adaptive e-Learning with Eye tracking

API – Application Programming Interface

BROMP – Baker Rodrigo Ocumpaugh Monitoring Protocol

CADEP – Professional Educational Development and Supporting Centers

DSM-V – Diagnostic and Statistical Manual of Mental Disorders

ECG – Electrocardiogram

EEG – Electroencephalogram

EMG – Electromyography

EML or EmotionML – Emotional Markup Language

GSR – Galvanic Skin Response

GUI – Graphical User Interface

HART – Human Affect Recording Tool

HCI – Human Computer Interaction

HEI – Higher Education Institutions

HTTP – Hypertext Transfer Protocol

ICA – Index of Cognitive Activity

ICT – Information and Communication Technology

IoT – Internet-of-Things

ISPA – *Instituto Superior de Psicologia Aplicada*

MSc – Master of Science

OWL – Web Ontology Language

POG – Point of Gaze

RDF – Resource Description Framework

RDFS – RDF Schema

RGB – Red Green Blue

ROC – Receiver Operating Characteristic

SDK – Software Development Kit

SQL – Structured Query Language

TCP – Transmission Control Protocol

URI – Uniform Resource Identifier

USB – Universal Serial Bus

XML – Extensible Markup Language

1 Introduction

A growing problem in Higher Education Institutions (HEI) is the amount of student dropouts due to disabilities, emotional factors, and economic, social or academic marginalization. This is worsened by the lack of educational resources, among faculty staff, to meet the requirements affecting vulnerable students, and by the gap in communication among staff, educators and administrators.

The students' cultural and social background is a strong influence in their educational process, and can often lead to social exclusion and marginalization, in learning activities and community life, further reducing students' will to learn and self-grow.

An adapted educational system which makes use of technology to aid teachers and students, allied with inclusive educational planning and policy-making, may help mitigate these problems.

In the recent decades, society shifted from a passive consumerism of information to an active producer and gatherer of knowledge, mainly due to the globalization of the Information and Communication Technologies (ICT). This shift in paradigm is more so evident in our education systems, where the integration and use of new technologies has been stimulated by governments worldwide [1].

In the United States, the National Education Technology Plan sets a series of recommendations for teachers, educators, faculty staff, researchers, policymakers, technology developers, community members and organizations, and learners and their families, in order to create a collaborative environment for technology enabled learning [2].

A report to the European Commission on new modes of learning and teaching in higher education sets the recommendation to personalize individual ways of learning, centered in the student, using technology in the classroom to combine all the different needs and to offer a personalized experience for each student, which improves the effectiveness of learning and can also have significant effect in reducing drop-outs [3].

Technology-enabled environments can help improve the students learning procedure, facilitating the access to information and knowledge, adapting its content to the individual needs of each student, increasing the stimulus for learning, and easing the communication between the teacher and the learner, providing a scenario of equal opportunity for every student, open to a collaborative process.

This ICT integration in the learning environment has been accomplished in many ways, the most recognized being the virtual or online learning platforms, where numerous courses and degrees are available. Other examples are the use of electronic grade books, real-time feedback on teacher and student performance and learning games [4].

1.1 Dissertation Integration – ACACIA

The project ACACIA stands for *Apoyo, Refugio, Comunicación, Adaptación, Innovación y Enriquecimiento* (Support, Refuge, Communication, Adaption, Innovation and Enrichment) and is designed around the integration of experiences, resources, teams, problems and solutions in HEI.

The main goals of the project are the detection, prevention and eradication of several education problems in the HEI in South and Central America, such as school dropouts due to emotional factors, discrimination and marginalization by disparity or inequality, learning disabilities or communication gaps between involved members.

In order to achieve such goals, the intra- and inter-institutional cooperation in the development and production of educational and technological resources is needed. Using innovative technologies to complement educational processes, the development of social and emotional strategies for the population risking academic exclusion, is required. This is accomplished with new institutional organizations: the Professional Educational Development and Supporting Centers, named CADEP.

CADEP centers make use of an integrated system of 5 modules (*Empodera*, *Innova*, *Cultiva*, *Apoya* and *Convoca*), in order to articulate with the educational community for comprehensive support with an individual approach to each student. The CADEP modules function together to track students at risk, to support and train technical, academic and administrative staff of the institution, explore new teaching strategies and to utilize innovative information and communication technologies in didactic classes, stimulating the entrepreneurship between students and professors.

This dissertation work is integrated in the ACACIA project, more specifically in the *Apoya* module. This module has two components, the first one is technological and the second one is human.

The goal of the technological component is to implement an automated emotion detection system that allows to monitor and support students by providing automatic recommendations.

The human component seeks to educate, inform and divulge throughout the academic community, guidelines for the recognition and handling of persons in situations of exclusion (attention disorder, including disability, cultural differences, extreme emotional situations, etc.).

This module directly relates to all other development modules, providing valuable information to *Empodera*, *Innova*, *Cultiva* and *Convoca*. In the final design the module will be integrated in CADEP, and will communicate with *Disemina*. *Disemina* provides guidelines for the detection and execution of situations that may generate social exclusion attending the broadcasting actions to raise awareness, for apprehension and for the action.

Apoya has a set of tasks that are described as follows:

Task 1 – Automatic detection of affective states. The work focuses on the implementation of a system for detecting emotions, also defining a methodology for data collection and methodology for emotional labeling and system management.

Task 2 – Tracking and recommendations automation. This work deals with the automation of monitoring the student and the system that generates recommendations to help improve their academic level.

Task 3 – Promotion of multiculturalism and diversity. The objective of this task is to generate guidelines for the detection and recognition of people at risk of social exclusion, a guide for action (treatment, activities, visibility) and a system of courses to train university and technical personnel (teachers, administrative staff and technical supporters included).

Task 4 – Integration strategies. The objective of this task is to define the protocol operation and the internal articulation of *Apoya* module, define the operation of the laboratory module, define the diffusion and disclosure system and the system of internal evaluation of the module itself.

1.2 Objectives of the Dissertation

The work presented in this dissertation focused in the creation and development of tasks 1 and 2. It begins with a study of the human emotions and their role in education systems, followed by possible application scenarios, the design, development and implementation of a framework, and respective aiding tools, that allow the gathering of students' emotions and behaviours, and analyses the collected data to extrapolate suggestions and alerts, of problematic student profiles.

1.3 Research Method

The research method used in this dissertation follows the Classical Method, detailed in [5], with the resulting Research Question and Hypothesis:

1.3.1 Research Question

Can a technological solution capable of identifying users' emotional state, which then through the help of the teachers be used to promote student's motivation and consequently prevent school dropouts?

1.3.2 Hypothesis

If a platform that integrates a set of automatic emotion detection systems, with real-time alerts to the teachers, is developed, then the identification of student's behaviour changes is facilitated, leading to the potential prevention of dropout intentions.

1.4 Dissertation Outline

The outline of this dissertation document tries to follow the steps of the classical research method, beginning with a more theoretical overview and then shifting to a more practical one.

- Chapter 1 – Introduction: In this introductory section, the motivation background and framework of the dissertation is presented, as well as the underlying Research Question and the formulated Hypothesis;
- Chapter 2 – State of the Art: In this section, a summary is made of the theory and technology available in the same research area, starting with an overview of the theories of Emotion and methods of observation, followed by an analysis of the eye tracking, gait recognition and facial expression detection technologies;
- Chapter 3 –Students' Monitoring Knowledge Base: This section presents the application scenario proposed, consisting of four case studies, and details the proposed ontology model, to serve as a knowledge base for the management of the students' data;
- Chapter 4 –Framework for Students' Profile Detection: In this section, a framework for monitoring and managing the students affects is proposed, and the architecture developed, for the implementation of such framework, is presented;
- Chapter 5 – Proof-of-Concept Experiment: This section describes the method deployed and the implementation of the experiment, used to test the hypothesis. An analysis of the results is provided as well as the validation of the implemented work;
- Chapter 6 – Conclusions: To finalize this document, the concluding remarks of the dissertation work are made, and the future work proposals are anticipated;

1.5 Original Contributions

The ontology model, the framework for student's profile detection and its architecture, the software implementation of the ontology, user interfaces and processors, presented in section 3.2 and chapter 4, are all original contributions from this dissertation.

Additionally, the application scenario, described in section 3.1, was developed in a collaborative process with the ACACIA project.

The device technologies used in the architecture implementation, such as the GP3 Eye Tracker, which includes the GazePoint Control software and Application Programming Interface (API), and the Affectiva emotion recognition Software Development Kit (SDK), were developed respectively by Gazepoint and Affectiva. However, the data treatment process was an original contribution as well.

2 State of the Art

In this chapter, a review of the state of the art is made, regarding the main topics analysed in this dissertation. A general overview of the current understanding of emotions, and how influential they can be, is presented first, followed by a review of available contactless emotion detection systems and their usage, namely eye tracking, skeleton tracking and facial expression recognition.

2.1 Emotions

In our everyday life, emotions play a very important role in every action we make: in nonverbal communication, in the way we process information, in our attention, and our biases towards information [6]. This can all be influenced by our affective state, consciously or unconsciously [7].

Emotions are also one of the central keys in ensuring our survival, fear being one of the most common and recognizable emotion among us human and many other animals.

There is no scientific consensus on a single, clear definition of emotion, as many different theories use the term, either in a broader or narrower sense.

Sometimes the concepts of emotions, moods and feelings are confused and used interchangeably with no regard to their meaning variances.

In the narrow sense, an emotion is more commonly observed as a short, intensive, mental activity [8].

According to [9] emotions arise from internal or external events, as cognitive data, and are attributed to concepts and states, feelings are the subjective experience of an emotion or set of emotions, and moods are an overall state of emotion, which is sustained over longer periods of time, is less intense, and is more consistent than emotions themselves, who can change rapidly.

It seems there is consensus, however, that some emotions (Anger, Fear, Sadness, Disgust and Enjoyment) are universal to all humans, despite the sociological or environmental background [10].

2.1.1 Components of Emotions

Scherer defines a core set of emotion components [11] that are commonly accepted, despite the lack of consensus on a model for what causes emotions, how they are classified, and how they are described (categorically or dimensionally).

Appraisals are the cognitive component that causes cognitive processes to evaluate events and situations in the genesis of emotions.

The physiological component accounts for the variation responses in the central and peripheral systems, and is responsible in the modulation of expressive features and in appearance changes (posture, skin colour).

Action tendencies are the behavioural component associated with the motor responses caused by appraisals.

The expression component covers the facial and vocal expressions, as well as body posture and gestures, generated by emotional responses.

Feelings are the subjective component, which are part of the complex phenomenon of the emotion.

Emotions Dimensions

Several models concur in defining emotions in different dimensions [12]–[14], as an attempt to conceptualize graphical representations of emotions' taxonomies.

One of the more recent studies, proposes a factor structure with four dimensions, namely:

- Valance, that ranges from Joy to Sadness (pleasant to unpleasant);
- Arousal, that ranges from Attention/Love to Anger/Hostile (passive to active energy);
- Dominance, that ranges from Interest/Self-Assurance to Fear/Anxiety (in control to lack of control);
- Unpredictability, that ranges from Surprise to Disgust (preference for to aversion for unpredictable)

2.1.2 Emotion's Recognition Devices

There are multiple platforms that make use of different bio-sensors, in order to analyse and determine a person emotional state. This is usually accomplished via invasive sensors like Electrocardiogram (ECG), Electroencephalogram (EEG), Electromyography (EMG), Galvanic Skin Response (GSR), Temperature and Respiration sensors.

Alternatively, some tools use facial recognition software to analyse basic and complex emotions. For example, the iMotions Platform (which also integrates other sensors data) uses two similar modules to achieve this. The first uses Affectiva's AFFDEX algorithm, the second uses Emotient's FACET algorithm, and both analyse facial expressions to determine human emotional reactions [15].

However, regarding an analysis environment, like a classroom, one fundamental problem can arise: If the student is aware that he/she is being monitored, his behaviour is going to reflect that, possibly leading to anxiety, insecurity, pretence or other states that disguise his regular emotional state, leading to inaccurate readings.

Thus, the need to conceal the measuring device from the student's attention, or at least cause the smallest amount of discomfort and distraction, is desirable.

This means that, sensors that require contact with the user, can contribute to undesired disturbances in the measured emotions, as well as disrupt the performed activity, thus the choice of non-invasive technologies for this type of application.

2.1.3 Human Observation and Coding of Emotions

Usually two methods are used when assessing emotions analogically: either the person being analysed self-reports their belief of their current emotional state, or an independent trained observer reports the emotional changes observed.

While the self-evaluation of one's current emotional state is subjective to the time it is made and by the emotional state itself, a biased observation can occur. On the other hand, an external observer can only register significant emotional changes, portrayed by the student, thus missing any internal feelings not expressed or possibly misread.

Baker Rodrigo Ocumpaugh Monitoring Protocol

An example of a quantifiable observation method is the Baker Rodrigo Ocumpaugh Monitoring Protocol (BROMP) [16], where different coding schemes can be programed in order to record students behaviour and affect in a classroom.

The observations recorded using this method are often used to obtain ground truth labels for Educational Data Mining research.

Emotional Markup Language

Similar to BROMP observation method, an Emotional Markup Language (EmotionML or EML) [17] has been defined by the W3C Emotion Incubator Group, to serve as a general emotion representation language, usable in three different setups: the manual annotation of observed emotion data; automatic recognition of emotions; and in the generation of emotion related system responses.

2.2 Eye Tracking

An eye tracker is a device that measures the activity of a person's eyes (position and movement), and is usually comprised of a head-mounted or a remote (desk mounted) apparatus.

Eye trackers are commonly composed by a light source (usually infrared or near-infrared) and one or more cameras pointed at the person eyes.

The use of an infrared light source is preferred due to the fact that it helps reduce the negative effects of ambient lighting conditions.

The light reflection on the cornea is then used to extrapolate the direction of the gaze, as it is the brightest spot in the pupil [18].

Other biometrics can also be measured by the eye tracker, such as:

- gaze fixation: the location where the gaze is maintained, for necessary amount of time for the eye to focus, when the visual input occurs;
- gaze duration: the amount of time for which a gaze is maintained;
- scan path: the vectors set that represent the path of gaze;
- areas of interest: refers to designated gaze locations and the collected gaze statistics for each area;
- pupil size: variations occurring in the shape and diameter of the pupil;

And various types of eye movements as well:

- saccades: the quick eye movement between two fixations;

- smooth pursuit: the eye movement while following a moving object;
- vergence: the simultaneous movement of both eyes in opposite directions to obtain or maintain single binocular vision;
- vestibulo-ocular movements: reflex eye movement, in the opposing direction of the head movement, that preserves the image in the centre of the visual field.

2.2.1 Commercial Eye Tracking Devices

As eye tracking technology becomes cheaper over the years, solving acquisitions costs in under-developed markets, more applications are now possible as a result.

There are now many low-cost devices, such as the Visual Interaction's myGaze tracker (499€) [19], GazePoint's GP3 tracker (461€) [20], The Eye Tribe's Pro tracker (185€) [21], or the Tobii's EyeX controller (119€) [22]. These products announced features and specifications are presented in Table 2-1.

Table 2-1 Low-cost eye trackers technical specifications and features

Company	Visual Interaction myGaze	Gazepoint	The Eye Tribe	Tobbi Tech
Product	myGaze Eye Tracker	GP3 Eye Tracker	EYE TRIBE TRACKER PRO	Tobii EyeX
Website	http://www.mygaze.com/products/mygaze-eye-tracker/	http://www.gazept.com/product/gazepoint-gp3-eye-tracker/	https://theeyetribe.com/products/	http://www.tobii.com/xperience/products/
Prices	€645.41	€461	€185	€119
Main Technical Specifications:				
Sampling Rate	30Hz	60Hz	30Hz to 75Hz	60Hz
Gaze position accuracy	0.5°	0.5° – 1°	0.5° – 1°	N.A.
Precision	0.1°	N.A.	N.A.	N.A.
Operating Distance	50cm - 75cm	N.A.	45cm – 75cm	45-100cm
Tracking Range/Head motion (Width X Height at Distance):	32cm x 21cm at 60cm	25cm x 11cm at 65cm	50cm x 30cm at 65cm	40cm x 30cm at 65cm
Latency	< 50ms	N.A.	< 16ms	15ms +/- 5ms
Tracking Technique	Binocular	N.A.	Binocular	N.A.
Calibration Mode	0/1/5 points	5 or 9 points	6, 9 or 12 points	N.A.

The main differences between these low-cost sensors and the high-end tier ones (which usually sell for around 10000€ and upwards) is the frequency of the frame rate, the accuracy and the precision of the measurements, as would be expected. While the lower tier trackers usually announce an average accuracy and precision, respectively of 1º and 0,1º, the high-end tier provides values at least 10 times better.

Each tracker provides its own software and/or an API or SDK, to help analyse the metrics given by the trackers raw data.

The basic metrics include the 3D position of the eye and the Point of Gaze (POG). However, some trackers like the Eye Tribe don’t yet support the measurement of the pupil diameter due to the impossibility to evaluate the person distance to the sensor.

A pre-print, comparing the EyeLink 1000 with the Eye Tribe tracker, concluded that the velocity and trajectory of the saccadic movement cannot be accurately measured by the Eye Tribe tracker due to its lower frequency sampling [23].

Other devices like the myGaze (and possibly others) are not optimized for high resolution quantitative measurements of saccadic movements and fixations due to the camera’s low frame rate (30Hz).

In [24] the GazePoint GP3 proved to be unreliable in some users wearing glasses, also with small eye size persons, persons with glossy hair, or even direct sunlight. It was also concluded that the GP3 was not good for detecting character reading, but otherwise well suited for eye tracking studies.

The Gazepoint GP3 eye tracker was chosen to be used in this dissertation, because of the reported frame rate value of 60Hz, and because it was one of the few low-cost devices to report the pupil size.

2.2.2 Emotion and Affect Analysis

In the last decade, the potential of eye tracking has also been employed to detect emotional state and cognitive activity.

The EyeTracking, Inc. Workload Software uses a patented algorithm, the Index of Cognitive Activity (ICA), to analyse the pupillometry data and extract the cognitive activity signal.

The ICA measures pupil changes, filtering light reflexes that cause circular muscle activity, thus inferring the mental effort, thru radial muscles activity [25].

These two methods benefit from being non-invasive and allowing the person to have more freedom of movements and less discomfort while being tested.

2.3 Skeleton Biometrics Using Kinect and Emotional State Correlation

It is well known that human emotions can affect the person's gait and locomotion [26]–[28].

In [29], 4 emotions were studied (anger, sadness, fear, happiness), and the resulting impact in locomotion was evaluated in intersegmental coordination patterns, amplitude of thigh elevation angles compared from those in neutral locomotion, intersegmental plane different orientations, and speed gait, concluding that emotions can be successfully recognized from the changes in locomotion.

In the recent years, the 3D mapping of the human body has been widely improved with the availability of low cost tracking devices (e.g. Microsoft Kinect).

Traditionally in computer vision, two distinct approaches can be used in the analysis of a person skeleton key points: marker-based and marker-less tracking.

While the use of markers, attached to the person's body, and the use of multiple camera views, provides the more accurate results, a marker-less solution also yields reasonable performance values [30].

One marker-less solution is the use of different camera views to extract the person silhouette, so that a skeletal model can be fitted in. Another one can be achieved using Microsoft's Kinect 3D depth sensors, which use an infrared camera to gage the distance from the camera to the tracking points.

Many scientific studies have demonstrated the feasibility of using the Kinect to measure spatial and gait parameters, such as walking speed, step length, gait symmetry among others [30]–[35].

In [35] the authors compared the measurement of walking speed, stride time and length against the well-established Vicon marker-based motion capture system, concluding that decent performance values can be achieved.

In [33] the authors successfully attempt to use the Kinect to recognize a person from skeleton data, using automatic detection of half gait cycle and feature calculation of every gait cycle. This study included novel features such as, area of upper and lower body parts, distances between upper body centroid and the centroids derived from different joints of upper and lower limbs, and feature selection and classification using Adaptive Neural Networks.

In [31] three classifiers were studied (Naïve Bayes, 1R, and C45), regarding their performance in gait recognition, concluding with 91% accuracy, using the Naïve Bayes classifier and requiring only 4 features to correctly identify a person: height, leg length, torso length and arm length.

2.4 Facial Recognition Software and Emotion Recognition

The human computer interaction has increased exponentially in the last decades, becoming more and more important in all activity sectors.

In order to improve these interactions, computers can now detect the user emotional state, which directly influences their actions and decision making.

This advance in computer vision was made possible by the evolution in automatic facial recognition systems and the developments in neural networks, in particular with the use of deep learning to make classifications and pattern analysis. Thus, emotion detection by facial expression recognition is now possible through image processing.

There are currently several algorithms based in the detection of image characteristics, that deliver, with reasonable accuracy, the analysis of the primary set of emotions (defined by Paul Ekman in the 1970s as anger, sadness, disgust, joy, surprise and fear [36], [37]), a few complex ones (neutral, contempt, puzzlement, confusion and frustration), and even some additional characteristics (valence, engagement, attention, expressiveness, negative and positive mood).

The majority of the algorithms analysed can be divided in three distinct stages: face acquisition, extraction of the desired facial features, and classification of the facial expression analysed.

2.4.1 Face Detection

Face detection algorithms are generally used beforehand to identify the region containing the face, being Robust Real-Time Face Detection algorithm [38], or an adapted version of it, the most commonly used.

Because pose angle and distance, scale changes and uneven illumination can be a problem, some algorithms require some sort of face normalization.

Some algorithms also process background separation and facial features separation before applying a face segmentation technique.

2.4.2 Facial Feature Extraction

For the Facial Feature Extraction process two distinct methods are used: Deformation Extraction and Motion Extraction.

Deformation extraction requires a baseline or a neutral face from which to compare against the facial actions, whereas Motion Extraction records the changes in facial expressions.

Several methods can to accomplish this feature extraction, such as neural networks, Gabor wavelets, differential images, 3D motion models, point distribution models, labelled graphs, feature tracking, parametric motion models, high gradient components, among others [39].

2.4.3 Emotion Classification

Usually two different approaches can be used in order to describe the identified emotion: facial affect detection and facial action unit detection [40].

The more common descriptors used in facial affect detection are the six basic emotions defined by Ekman (anger, sadness, disgust, joy, surprise and fear). For the facial action unit detection, one of the most popular methods for coding the facial actions is the Facial Action Coding System [37].

2.4.4 Example of an Algorithm

The study presented in [41], used a set of 313 sequences of images, with 640x480pixels in 8-bit grayscale. The face detection algorithm, a cascade of classifiers (each with a subset of Haar Basis functions), used scan patches of 24x24 pixels, classifying each as face or non-face. Between 2 and 200 subsets of filters were selected using a feature based on Adaboost. The best performance single-feature classifier was selected.

The weights of the weak classifiers in the neural network were then adjusted using Adaboost rules. These adjustments were repeated until strong classifiers had the desired performance. The images were then converted into a Gabor magnitude representation, using a bank of Gabor filters at 8 orientations and 5 spatial frequencies.

The facial expression classification was based on support vector machines (SVM). A combination approach, in which the Gabor Features chosen by Adaboost were used as a reduced representation for training SVM's (AdaSVM's) outperformed Adaboost by 3.8 percent points, and higher spatial frequency Gabors and higher resolution images, also improves performance.

White dots indicate locations of all selected Gabors. In figure Figure 2-1 each expression is a linear combination of the real part of the first 5 Adaboost features selected for that expression. Faces shown are a mean of 10 individuals.

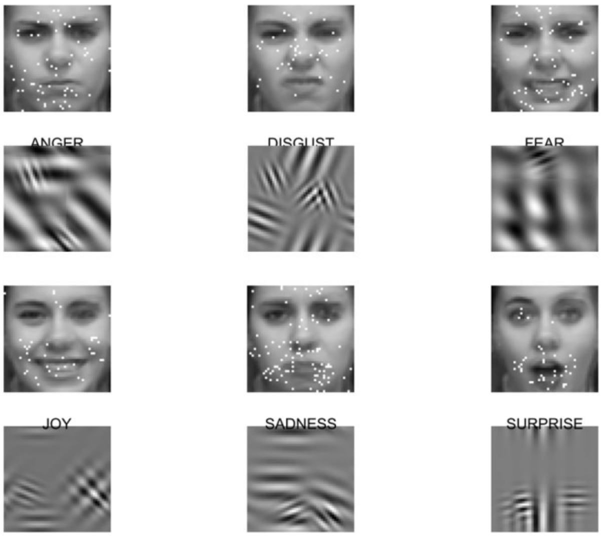

Figure 2-1 Gabors selected by Adaboost for each expression [41].

2.4.5 Commercial Emotion Detection API's and SDK's

Commercial emotion detection algorithms are difficult to obtain so, in order to make an overview, one must present the results provided by the companies.

iMotion is a biometric research laboratory where human behaviour is one of the major study areas. They developed an image processing platform, in cooperation with Emotient, capable of detecting human emotions, and connected this feature with other biometric sensors in order to improve the accuracy of their results. The company shares their findings to Affectiva which develops products based on iMotion's developments [15].

Affectiva's product "Emotion as a Service" provides advanced emotion analysis in order to increase the consumer experience with gaming, product testing, learning, entertainment, and research, among others. Their software is powered by iMotion's software, and the combination with "Affdex" (another product of Affectiva) provides a bigger solution towards emotion classification [42].

Emotient's product "FACET technology" intends to gain profound knowledge on human emotions reactions via facial expressions. The facial expression detection algorithm gives valence, action units and emotion channels. The company has since been bought by Apple. As such, information about this product is no longer disclosed.

Noldus FaceReader is a product able to do facial expressions analysis. The software had more than 10.000 images used for training and is capable of detecting the six basic facial expressions as well as neutral and contempt. It can also calculate gaze direction, head orientation and person traits. The software is used by more than 400 universities, research institutes and companies in many markets, such as consumer behaviour research, and psychology, among others [43].

nViso SA is a company that is also doing research in facial emotion recognition. The EmotionAdvisor is their main product and provides a scalable, robust and accurate artificial intelligence solution to measure emotional reactions. The company recently won the "IBM Bestseller Award 2015" during the "Success Made with Partners" conference, confirming the success combination between nVisio EmotionAdvisor and IBM Cloud Service [44].

Microsoft has its own Cognitive Service's API. The Emotion API, as Microsoft named it, is capable of detecting faces and analyse a range of feelings. The API uses images as input and returns the confidence across a set of emotions, as well as boxes of the detected faces in the image [45]. The API is free to use up to 30.000 images per month.

Table 2-2 Comparison Table of Commercial Emotion Detection APIs and SDKs

	NAME	AFFECTIVA AFFDEX	EMOVU EYERIS	NVISO	NOLDUS	SIGHTCORP INSIGHT SDK	SIGHTCORP F.A.C.E.	IMOTIONS	CROWDEMOTION	MICROSOFT COGNITIVE SERVICES EMOTION API
	WEBSITE	http://www.affectiva.com/solutions/apis-sdks/	http://emovu.com/e/developers/api/	http://www.nviso.ch/index.html	http://www.noldus.com/human-behavior-research/products/facereader	http://sightcorp.com/insight/	https://face.sightcorp.com/	https://imotions.com/api/	http://www.crowdemotion.co.uk/solutions.html	https://www.microsoft.com/cognitive-services/en-us/emotion-api
#EMOTIONS DETECTED	anger	✓	✓	✓	✓		✓	✓	✓	✓
	sadness	✓	✓	✓	✓	✓	✓	✓	✓	✓
	disgust	✓	✓	✓	✓	✓	✓	✓	✓	✓
	joy/happy	✓	✓	✓	✓	✓	✓	✓	✓	✓
	surprise	✓	✓	✓	✓	✓	✓	✓	✓	✓
	fear/scared	✓	✓	✓	✓	✓	✓	✓	✓	✓
	contempt	✓			✓			✓		✓
	neutral		✓	✓	✓	✓	✓			✓
	Puzzled					✓				
	Confusion							✓		
	Frustration							✓		
ADDITIONAL FEATURE DETECTION	valence	✓	✓					✓		
	engagement	✓								
	Attention		✓							
	Expressiveness		✓							
	Negative Mood		✓							
	Positive Mood		✓							

The different APIs detect different emotions. Table 2-2 provides a comparison between the software applications mentioned before and others less known. The table reveals the number of emotions that each application can recognize as well as some additional features.

The Affectiva SDK was chosen to be used in this dissertation, because of the higher number of emotional metrics reported, and because of it is free for individual research use.

2.5 Existing Technologies for Student's Monitoring in the Learning Environment

One of the biggest challenges that arise from the technology assisted learning environment, or e-learning, is the inability to create an emotional connection between the student and the teacher.

In order to assess the student level of understanding, a test needs to be performed afterwards. Contrasting with in-person learning, in which that appraisal can be done in real-time, by the teacher, enabling the detection of cognitive stress, comprehension difficulties, tiredness, among other problems [46].

By tracking eye and gaze position, fixation time and location, and blink rate, it is possible to correlate these metrics to the student "level of attention, stress, relaxation, problem solving, successfulness in learning, tiredness or emotions" [18].

The e5Learning platform uses pupil information, fixation duration and blink rate to recognize the user emotion, namely "high workload/non understanding" and "tiredness" [47].

Higher cognitive load has been correlated to occur when a student performs a search, with an increase in pupil size, and lower saccades and blinks. Pupil size is also larger, after highly arousing positive and negative stimuli, than after medium arousal with neutral stimuli [47].

Eye tracking is one particular technology that has been used directly in e-learning systems, in order to merge the gaps in Human Computer Interactions (HCI).

One of the first projects to include this technology was “Adaptive e-Learning with Eye tracking” (AdeLE) [48], which analysed eye-movement patterns during learning and tried to link those patterns with cognitive processes.

In [49], the eye tracker is used to adapt presented content to the learner by following his topics of interest.

In [50], a empathic software agent interface was developed using an eye tracker, to infer the focus of attention and motivational status of the learner, responding with instructional behaviours and display of emotion.

The e5Learning learning environment recognizes basic emotion via an eye tracker, assessing “high workload”, “non-understanding” and “tiredness” situations in order to adapt content presentation in real-time [51].

As the cost of these eye tracking devices lowers, it is expected a more generalized use with platforms that require the input of the learner’s attention and level of understanding, in order to adapt and personalise their content to the student best interest.

3 Student's Monitoring Knowledge Base

The prevention of HEI dropouts and other learning related problems, using a technological approach, is now possible with the multitude of emotion analysis systems available. Using the Internet of Things (IOT) paradigm, integrating several devices such as biomedical sensors, eye trackers and others, to collect information from the students, can support teachers in the identification and management of students' emotional state, during classroom lessons.

In this chapter, a scenario for the monitoring of students' behaviour and emotional state, using a technological support platform, is proposed, and an ontological model, designed to gather and manage student's information as a knowledge base, is presented.

3.1 Application Scenario

The proposed scenario consists in evaluating the students' emotions, behaviours and affective states, aggregating the information, using a framework capable of assessing problematic situations. The cases studies, represented in Figure 3-1, comprise three biometric acquisition devices and one integration platform, for information analysis and reporting system, as a proposed solution.

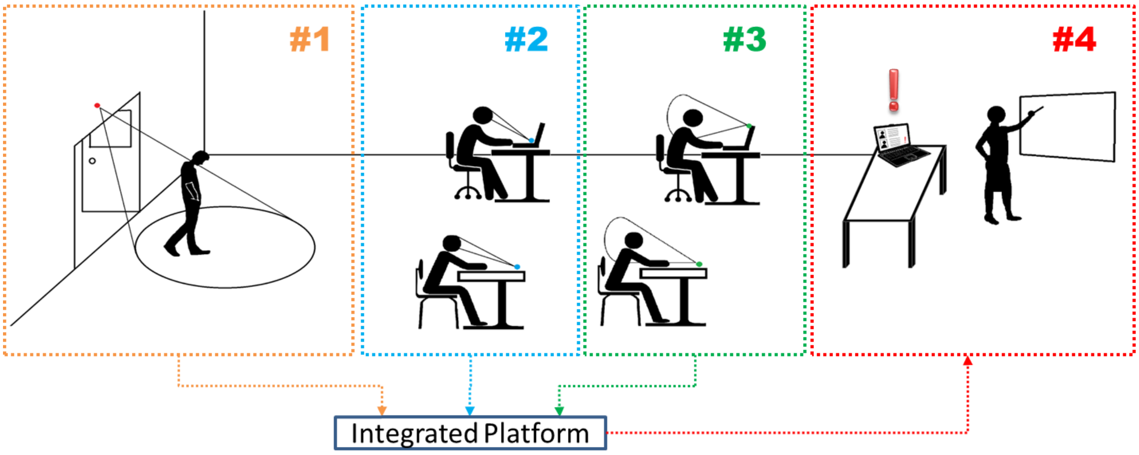

Figure 3-1 Scenario for student affective management

3.1.1 Biometric Data Acquisition

For this scenario three different capturing devices can be used: a Kinect, an Eye Tracker and a Camera, each described in the corresponding case study.

Case #1: Gait and Posture Analysis while entering or leaving the classroom.

This case study implementation will use a 3D motion capture device, composed by an RGB camera and an infrared depth sensor (i.e. the Kinect device), to track the student skeleton key points, in order to analyse its posture and gait, while entering and leaving the classroom.

It is expected with this observation, that a change in one student regular gait can be flagged as a signal for changes in determined emotional states.

Case #2: Eye Tracking while sitting at a desk.

This case study will use a remote Eye Tracking device, which is generally composed by an infrared light source, an infrared or near-infrared camera and an RGB camera. The Eye Tracker sits in the desk, directed at the student, and tracks head movements as well as eye movements (saccades, smooth pursuit, vergence and vestibulo-ocular movements, gaze fixations and eye blinks), and pupil's size and dilation changes (proven to be an indicator of cognitive activity).

The analysis of these biometrics can enable the recognition of the student different affective states, during the learning process, and can also be used to measure the engagement level of the student.

Case #3: Automatic Emotion Detection through Facial Expressions, while sitting at a desk.

This case study will use a regular RGB camera, pointed at the student's face, in order to record the changes of the facial expressions.

An algorithm will detect and isolate the face, then extract the facial features collected, analyses the classified facial expressions, and determines the correspondent emotion or emotions.

Both eye tracking and facial expression emotion detection case studies can be deployed to a student using a computer, alongside with the sensorial device, or without the computer and just the device.

3.1.2 Acquired Data Processing and Integrated Warning Platform

All the data recorded in the previous cases is to be collected and processed in real-time, and will be compiled creating an emotional map, that can be analysed to flag potential problematic situations of disengagement, attention disorder, learning difficulties, emotional stress, HEI dropouts, and possibly other situations.

Case #4: Integration Platform

An integration platform will be created to manage each student information and historic track record, and to provide a real-time early-alert system when a deviation from the regular patterns is detected. This platform is intended to be available, in real time, to the teacher, to aid in the class management, so that he or she may act accordingly to each situation and/or adapt the course to meet the needs of the students. It should also be possible that, in the end of each class, an automatic report is drafted and made available to the interested parts, namely the teacher, the student, the parents and the school.

3.2 Ontology

An ontology consists in a formal representation of concepts in a domain of discourse, called classes, the properties of this entities, and the relationships between them, called restrictions.

Alongside with the individual instances of classes, the ontology forms a knowledge database. Thus, it is used to manage all the knowledge of the proposed scenario.

The Web Ontology Language (OWL) [52] is a language schema, designed to represent this relational database, storing this concepts, their meaning, and how they relate to each other, in tables. It works as a layer above Resource Description Framework (RDF) [53] and RDF Schema (RDFS) [54], which in turn are a layer above Triples (the codification of semantic data in the form of subject–predicate–object) and Uniform Resource Identifiers (URIs).

Additionally, it allows reasoning over the ontology to infer implicit knowledge, even if it has not been specified.

3.2.1 Design

The ontology used was designed with Protégé 3.5 [55], [56], an older version that, although the lack of support for OWL 2 [57] and some outdated features, is much simpler to learn and to use. Additionally, it is always possible at any moment, to use it in newer versions, enabling it to all the existent capabilities. For the specific future usage of it, such evolution could be necessary.

Figure 3-2 is a diagram representation of the ontology model used to build the framework. The graphical representation used is based on a graphical representation tool of OWL ontologies, called VisioOWL [58].

Some classes have data type properties, represented in green, that help define each class individual.

Relationships and rules between classes are represented as object properties, in blue, and they provide the meaning between this classes.

Finally, each type of property is defined by cardinality restrictions, represented by the arrows with the respective cardinality written as a mathematical restriction, in order to ensure the appropriate data and individual relationship.

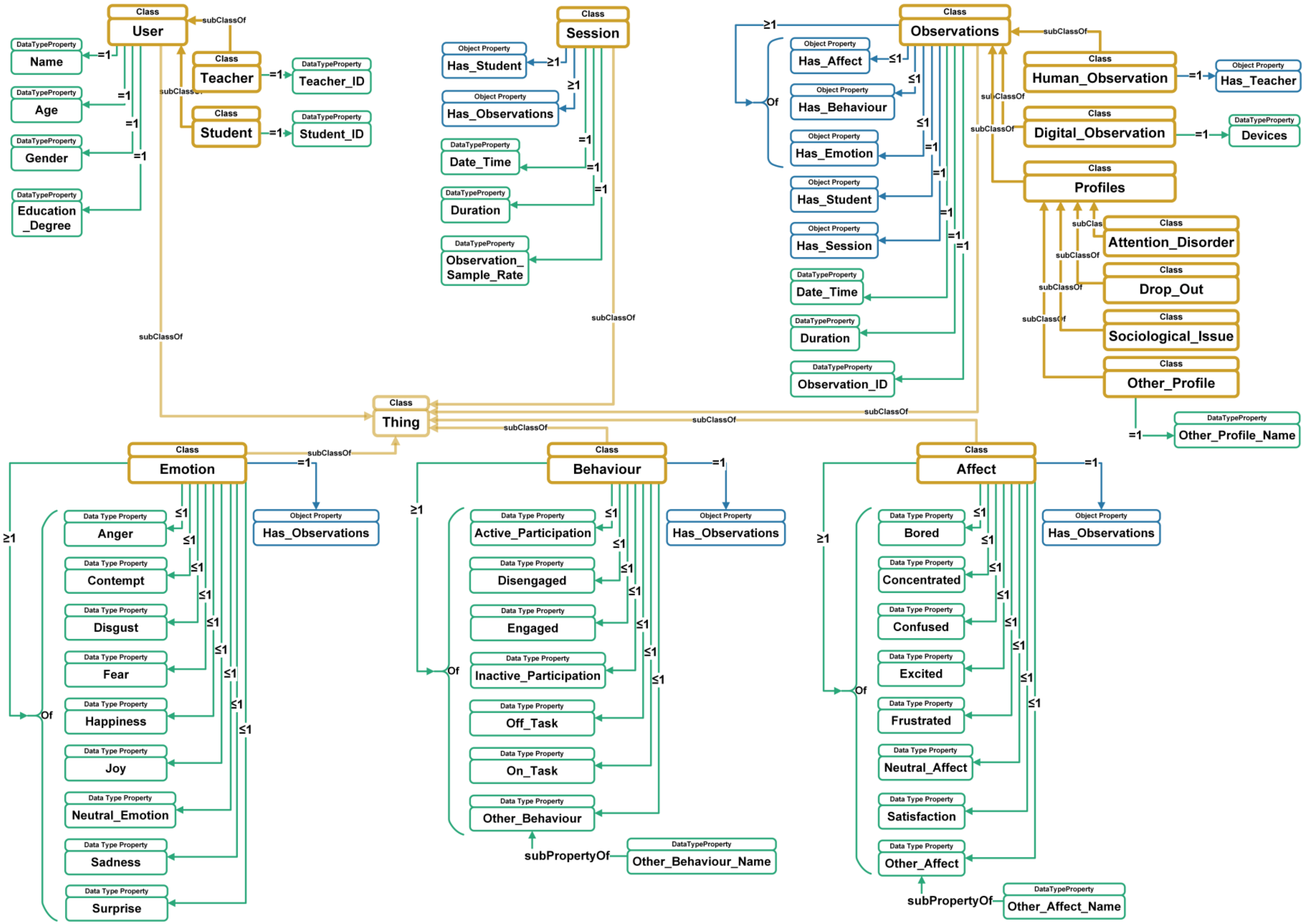

Figure 3-2 Ontology Diagram

There are six primary classes that represent the focus of the framework data input: User, Session, Observation, Affect, Emotion and Behaviour.

The proposed workflow of the data input in the framework follows the following principles:

– The first ontology individual to be created should to be of the class User, specifically a Student, given the inferred rule that any of the remaining primary classes require a relationship with at least one Student.

– The User class serves the purpose of storing the user personal details, and there can be two different types of Users: subclass Student and subclass Teacher, the Student being the target to be analysed by the framework, and the Teacher being a possible means of collecting observation data from a Student.

– Next a Session can be created, requiring at least one Student individual to be associated. It is also associated with one or more Observations individuals that are recorded during the Session duration, meaning that during one session, multiple observations can be recorded. This class represents a timeline in which the Student Observations were recorded, such as a classroom during a lecture, or a e-learning lecture.

– Then Observations can be made. This class requires exactly a Student and a Session to exist and that exactly one of each is associated with the observation. It is also required that, at least one of three classes (Affect, Behaviour or Emotion) and at the most one of each is associated.

Three types of Observations direct subclasses can be created: Human_Observation, Digital_Observation and Profiles.

Human_Observations represents the observations collected via the teacher or an expert input, and therefor require the link with exactly one Teacher individual.

Digital_Observations on the other hand, represents the observations collected via hardware or software devices that target the student.

The Profiles subclass and its subclasses (Attention_Disorder, Drop_Out, Sociological_Issue, Other_Profile) are intended to represent the various observation requirements necessary to trigger an alarm, flagging the situation corresponding each profile. In section 3.2.3, a more detailed explanation of each profile is presented, along with the proposed conditions necessary to trigger each profile alert.

– Finally, the triple twin classes, Affect, Behaviour and Emotion, each require to be connected to an existing Observations individual, and are the representation of the key input data, collected from either a digital device or a teacher, required to assess and profile the student.

Each of these classes have a set of defined data properties that are used to store the float values of the respective property, and it is required that at least one and at the most one of each exists for the respective class individual to be considered a valid input.

A more detailed explanation into each of these three twin classes and their properties is provided in 3.2.2.

3.2.2 Affect, Behaviour, and Emotion Classes Properties

It is important to note that, since there is no scientific consensus in the definitions of emotion and affect, with both terms being used, in some cases with the same connotation, with exchanging meaning, or with different purposes, in the presented work a clear distinction is made between emotion and affect: it is considered that an emotion is a person's internal state of being, while affect is considered to be the body physical reaction to an emotion or emotional event.

The initial idea for affect and behaviour properties was to provide a similar tool to BROMP, Human Affect Recording Tool (HART) [16], which is a "momentary time sampling method in which trained, certified observers record students' behaviour and affect individually". BROMP utilizes several affect and behaviour coding schemes, related to the students' engagement and their relationship with their learning environments, and its more common use is to obtain ground truth labels for Educational Data Mining research.

Just like the HART tool, the developed framework also allows to record these observations simultaneously, due to the apparent independency between affect and behaviour.

However, due to the particularities of each application scenario, described in section 3.1, it was decided to create two new coding schemes, for affect and behaviour, based on some of the coding schemes already used in BROMP. The coding schemes used in this framework, for the three twin classes, ought to be subject to change or adaptation, per the needs of any new application scenario.

Affect Coding Scheme

Bored: when a student appears to feel tedious/lacking interest about the activity he/she is participating in.

Concentrated: the student is engaged, paying attention and focusing in a task.

Confused: when a student appears to have difficulty understanding a concept of a received information or an activity he/she needs to perform.

Excited: the student is in a restless or agitated state, in the execution of the current task.

Frustrated: if a student feels distressed, annoyed or incapable to accomplish the assigned task.

Neutral_Affect: when the student shows no signs of affect that may influence his/her engagement in the current activity.

Satisfaction: the student shows enthusiasm in the execution of the current task, indicating a pleasurable experience.

Other_Affect: used to describe an affect not listed in this coding scheme.

Behaviour Coding Scheme

Off_Task: the student is not working on the activity assigned by the teacher.

On_Task: the student is doing the task he/she is supposed to be doing.

Active_Participation: the student is actively involved in the activity, showing work in its execution. The student needs to be On_Task to be able to be in Active_Participation.

Inactive_Participation: the student is involved in the activity, but shows no clear signs of his/her contribution. The student needs to be On_Task to be able to be in Inactive_Participation.

Engaged: the student is willingly working/learning, despite any difficulties in doing so. For the student to be Engaged he/she needs to be On_Task and can also be either in Active_Participation or in Inactive_Participation.

Disengaged: despite participating in the task, the student shows a clear disconnection to it, is uninterested in its execution and/or is in unwillingly to learn from it. For the student to be Disengaged he/she needs to be On_Task and can also be either in Active_Participation or in Inactive_Participation.

Other_Behaviour: used to describe a behaviour not listed in this coding scheme.

Emotion Coding Scheme

The coding scheme used for emotions followed the range of emotions reported by the various softwares of automatic emotion recognition through facial expressions. Thus, and to ensure the framework remains compatible with any automatic emotion recognition software, the definition of usage, of each of the emotions categorised, is relegated to the provider of the referred emotions' classification (either the specific software or the human observer).

The list of coded emotions is the following: Anger, Contempt, Disgust, Fear, Happiness, Joy, Neutral_Emotion, Sadness and Surprise.

3.2.3 Profiles

The four profile classes presented, named Attention_Disorder, Sociological_Issue, Drop_Out and Other_Profile, are intended to be used as the knowledge variables, representing the conditions necessary to trigger the respective profile alert.

The Other_Profile is a class open to expand the existing profiles to future work developed using this framework.

Attention_Disorder

Attention deficit hyperactivity disorder is usually regarded as having three main characteristics: hyperactivity, impulsivity and inattention. The fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V) defines it as a "a persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development" [59]. It also states that in adults (17 years and older) five or more symptoms (of the total 18 listed) are required to be observed for at least 6 months.

Sociological_Issue

Oppositional defiant disorder is defined by the DSM-V as "a pattern of angry/irritable mood, argumentative/defiant behaviour, or vindictiveness" [59]. Again, at least four symptoms from any of these categories (angry and irritable mood; argumentative and defiant behaviour; or vindictiveness) must be observed for at least 6 months.

Drop_Out

Because there are many factors that can lead to school drop outs and there seems to be no consensus in a clear diagnose, the observations used for this profile were proposed using common sense. The same process was used for the sociological issues (which can cover a broad range of particular situations).

Proposed Profile Detection Conditions

Since the objective with these observations is to assist teachers and faculty staff with the early detection of possible student's problems, and not with the diagnosis of said problems, a set of conditions are proposed in Table 3-1, to serve as minimum requirements to create alerts, related to the possibility of the student having the described profiles.

Table 3-1 Proposed Profile Alert Conditions

Profiles	**Emotions values above 0.5**	**Behaviours values above 0.5**	**Affects values above 0.5**	**Minimum number of occurrences, in the previous 30 minutes**
Attention Disorder	Sadness	Disengaged Off_Task	Confused	10
Drop Out	Anger Contempt	Disengaged Off_Task	Bored	25
Sociological Issue	Fear	Disengaged Off_Task	Excited Frustrated	10

The emotions, affects and behaviours, proposed for each profile alert, were based on the feedback received from an expert in educational psychology, and the values were assigned based on common sense, and will be subject to proper scientific validation in future work. This initial set of conditions can be used in a neural network analysis, as the set of initial weights for the neurons, to be developed in a future work, as part of an integration with machine learning algorithms, to be used for in correlation with the observations made by the teachers or psychologists, to attempt to predict the diagnose of the mentioned profiles.

4 Framework for Student’s Profile Detection

In this chapter, a framework is proposed, to streamline student’s data acquisition from an array of devices, to generate and manage a knowledge base, defined in section 3.2, intended to assist in the detection of profiles, related to student’s problems.

The architecture used for the implementation of the framework, as well as each module description, is also presented.

4.1 Framework Design

The proposed framework, illustrated in Figure 4-1, consists in a multi-level platform, encompassing hardware sensors, their respective middleware, SDK’s or API’s, data processors, an ontology, a file database and a graphical user interface (GUI).

The framework main goal is the storage standardization of the data, collected from the student, enabling the compatibility with any device, supporting multiple devices in real-time.

The secondary goal is to provide a platform that analyses and processes the collected information, extracting knowledge from the student data, to extrapolate possible profiles of student’s problems, offering real-time warnings and reports.

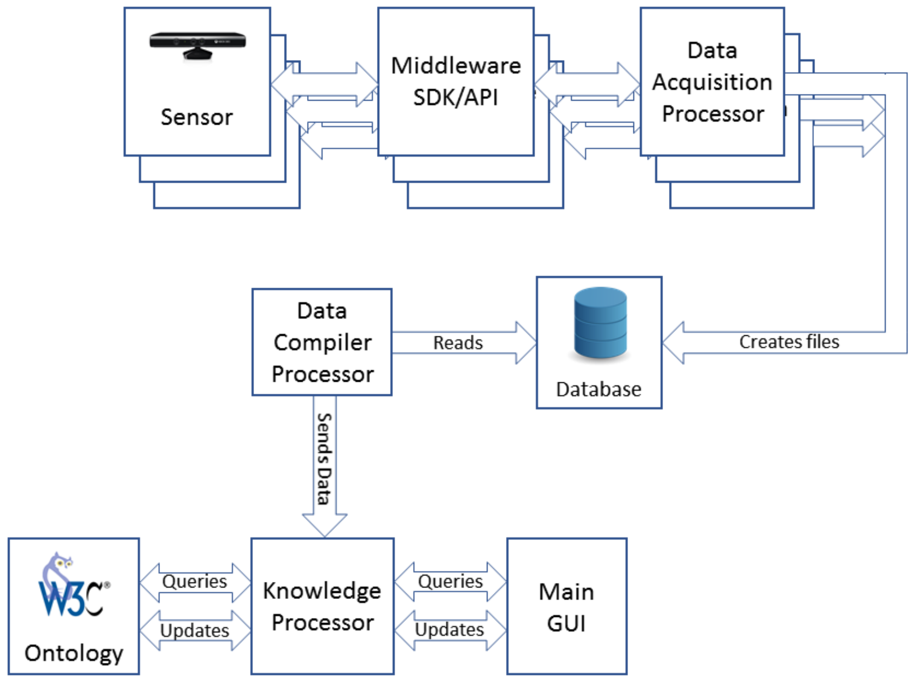

Figure 4-1 Framework Diagram

The first data entry point is through a sensor, collecting the student biometric data. Usually the sensor is accompanied by a custom software, an SDK or an API, that is used to connect to the sensor and convert its signals into biometric data.

This biometric data will then be processed in a Data Acquisition Processor, that compiles it into the appropriate format and saves it to files in a database.

A Data Compiler Processor will scan the database for new files and, for each file, compile the data average results into a single set of knowledge, which is sent to the Knowledge Processor.

The Knowledge Processor is responsible for the handling of all the knowledge being queried and updated into the Ontology, whether it comes from the Sensor-Data Compiler Processor chain, or from the second data entry point, the Main GUI. It is also responsible for the activation of the profiling alerts in the Main GUI.

The Main GUI presents the user (e.g. the teacher) with a set of tools to insert new information in the Ontology, to consult the existing knowledge and to manage the Data Processor's settings.

The Ontology is used as a relational knowledge database and its design is presented in more detail in section 3.2.

4.2 Architecture

This section presents the architecture developed, as the proposed implementation of the framework described in section 4.1.

The conceptual model presented offers a structural overview of the system components and the behavioural relationship between them, as shown in Figure 4-2.

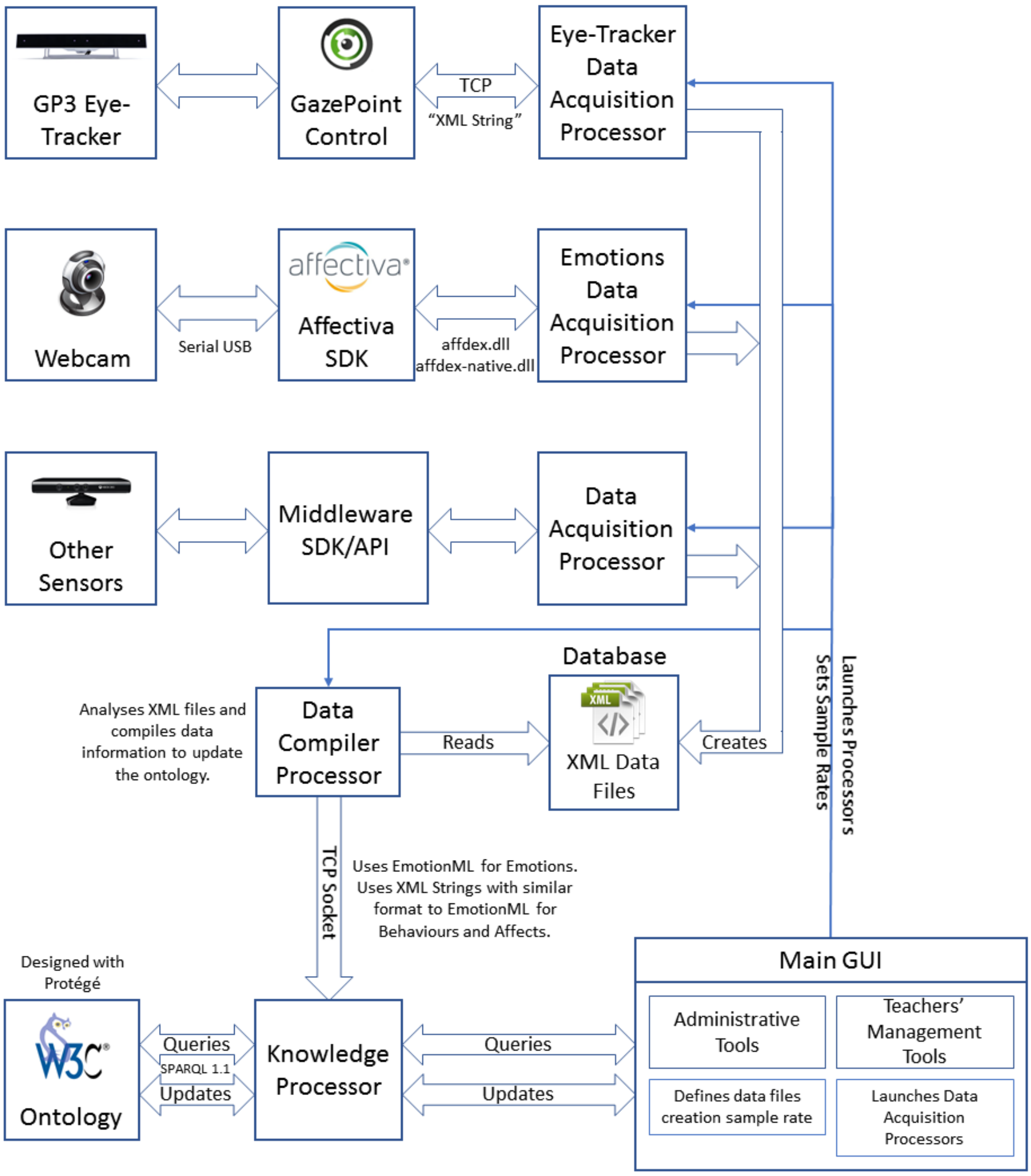

Figure 4-2 Architecture Diagram of the Student's Profile Detection Platform

Unfortunately, due to timing constrains, the implementation of a module containing the case study #1, proposed in section 3.1.1, was not feasible, and will be an option for future work.

4.2.1 Ontology

For the implementation of the platform ontology, a variety of .Net C# library API's were analysed: RDFSharp [61], OwlDotNetApi [62], LinqToRdf [63], OWL API .NET [64], and dotNetRDF [65].

A .Net C# library API was chosen to manage the ontology, contrary to the more traditional/common approach of using the Java framework Apache Jena [60], to keep all the coded programs in the same language, and to enable the seamless integration across applications and processes.

DotNetRDF was chosen because it provides almost the same functionalities as Jena (except OWL reasoning, for now), it is one of the most commonly used RDF/Triple Store libraries in C#, and it has the more detailed documentation of all.

The dotNetRDF API allows to manage, parse, read and write OWL and RDF files, as well as Graphs and Triple Stores, offers several engines for reasoning RDFS, Simple Knowledge Organization System (SKOS) [66], N3 Rules [67] and Triple Stores, and also several processors for SPARQL queries [68] and for SPARQL 1.1 updates [69], [70].

Two important functions were created to handle SPARQL queries (SparqlQuery) and SPARQL updates (SparqlUpdate), respectively.

Each with similar behaviour: receive the ontology graph and a command string containing the SPARQL query or update, as input parameters. The SparqlQuery then returns a string array with the query results and the SpaqlUpdate updates the ontology graph, while also saving it to a predefined RDF file.

Since many query and update commands are used more than once, all the commands used in the platform are set as constant strings or as static string returning functions with string(s) input(s), for commands that have fields with variable names. This allows the possibility to create an easy access library with these functions (OntoFunctions.cs).

SPARQL Query Example

The following is a Sparql query example, used to list the Student individual, linked to a given Observation sub-class individual, "Observation_Ind", by the object property "Has_Student":

```
PREFIX onto: <http://www.owl-ontologies.com/Ontology1480873742.owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

SELECT ?z
WHERE {
?y rdfs:subClassOf* onto:Observations .
?x rdf:type ?y .
FILTER regex(str(?x),'Observation_Ind','i') .
?x onto:Has_Student ?z .
}
```

All query commands start with a set of prefixes for the namespaces used by the queries.

Following the prefixes, most queries start with the “SELECT” statement, which serves a similar purpose to its Structured Query Language (SQL) counterpart: defining which variables are returned by the query (“?z” in this example).

The “WHERE” clause is where the values and variables being queried are defined.

In the example, the line “?y rdfs:subClassOf* onto:Observations .” is a triple, formed by a subject, “?y”, a predicate, “rdfs:subClassOf*”, and an object, “onto:Observations”.

If any of the triple elements is a variable, it is preceded by a “?”.

As such, the query result of this line would be all the sub-classes of the class Observations, defined in the ontology used, and would be returned to the variable “y”.

The following line, “?x rdf:type ?y .”, would return all the individuals of each sub-class in the variable “y”, to the variable “x”.

The “FILTER regex(str(?x),'Observation_Ind','i') .” would then filter the individuals in the variable “x”, leaving only the individual with the name matching “Observation_Ind”, in the variable “x”.

Finally, “?x onto:Has_Student ?z .” would return the Student individual, listed in the Has_Student object property, of the “Observation_Ind” individual, to the variable “z”.

SPARQL Update Example

The insertion of new knowledge in the ontology is accomplished using SPARQL updates, such as the following example, which inserts a new student individual and the respective properties in the ontology:

```
PREFIX onto: <http://www.owl-ontologies.com/Ontology1480873742.owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

INSERT DATA {
```

```
onto:Student_id rdf:type onto:Student .
onto:Student_id onto:Name "Name"^^xsd:string .
onto:Student_id onto:Age "Age"^^xsd:int . "
onto:Student_id onto:Gender "Gender"^^xsd:string .
onto:Student_id onto:Education_Degree "Education_degree"^^xsd:string .
onto:Student_id onto:Student_ID "Student_id"^^xsd:int .
}
```

Just like in a SPARQL query command, an update command also is preceded with the set prefixes to be used.

The “INSERT DATA” operation contains the triples to be added to the ontology graph.

In this example, the Student individual triple is inserted first, followed by triples of the values of the data properties: Name (string), Age (integer), Gender (string), Education_Degree (string) and Student_ID (integer).

4.2.2 Knowledge Processor

The platform input data can be received from any type of sensorial device, through Transmission Control Protocol (TCP) socket communication.

TCP socket programming was chosen over Windows Communication Foundation, Hypertext Transfer Protocol (HTTP) Web Services and named pipes, due to the easiness of using asynchronous socket programming (using Microsoft .NET Framework 4.5), for client-server systems.

The TCP socket client-server communication follows the stated protocol:

The Data Compiler Processor must establish a TCP socket connection to the specified IP and port of the machine running the platform, which must be in the same LAN of the device.

After the connection is established successfully, a string with the following parameters must be sent: “Connect_Request”.

This prompts the platform Main GUI, to pop-up a window, asking the user (e.g. teacher) to select or create the observed student and the session in which to record the data to be received next.

If a valid student and session are selected, the platform replies the received “Connect_Request” with the string “Connection_Established”, after which the device can start sending the sensor data, creating an Observations individual, of the subclass Digital_Observation, per each data record received.

The data sent by the Data Compiler Processor must follow the Emotion Markup Language [17] schema, using the FSRE [14], [71] category set, for reporting student's emotions, along with the starting time of the recording and the respective duration.

Because this schema offers a very clear, understandable syntax, that uses Extensible Markup Language (XML) serialization, enabling an easier computer reading, an adapted version of EML is also used to report student's behaviours and affects.

The name of the device used to record the observation is also passed as an <info> element, before the single EML emotion annotation.

The following code is an example of a hypothetical communication stream, of one Observation, sent from the Data Compiler Processor (with the compiled emotion's, behaviour's and affect's data), and the Knowledge Processor:

```
<info>DeviceName</info>

<emotion xmlns="http://www.w3.org/2009/10/emotionml" category-
set="http://www.w3.org/TR/emotion-voc/xml#fsre-categories" start="1268647200000"
duration="130">
        <category name="anger" value="0.00"/>
        <category name="contempt" value="0.08"/>
        <category name="disgust" value="0.43"/>
        <category name="fear" value="0.00"/>
        <category name="joy" value="0.00"/>
        <category name="sadness" value="0.02"/>
</emotion>

<BEHAVIOUR start="1268647200000" duration="130">
        <category name="Active_Participation" value="1.0"/>
        <category name="Disengaged" value="0.7"/>
        <category name="Engaged" value="0.3"/>
        <category name="Inactive_Participation" value="0.0"/>
        <category name="Off_Task" value="0.0"/>
        <category name="On_Task" value="1.0"/>
        <category name="Other_Behaviour_Name" value=""/>
        <category name="Other_Behaviour" value="0.0"/>
</BEHAVIOUR>

<AFFECT start="1268647200000" duration="130">
        <category name="Bored" value="0.8"/>
        <category name="Concentrated" value="0.0"/>
        <category name="Confused" value="0.3"/>
        <category name="Excited" value="0.2"/>
        <category name="Frustrated" value="0.5"/>
        <category name="Neutral_Affect" value="0.0"/>
        <category name="Satisfaction" value="0.0"/>
        <category name="Other_Affect_Name" value=""/>
        <category name="Other_Affect" value="0.0"/>
</AFFECT>
```

The Knowledge Processor will analyse this data, and insert the Observation records in the Ontology in real-time.

As a proof-of-concept of a real-time profile detection system, an alert system was implemented, for the detection of a mock-up case of a student with attention problems.

Every time there is a new Observation update, a query is made, searching every Observation made in the 30 minutes prior to the current one, for the value of the data property "Off_Task", of the class Behaviour, higher than "0.5", and if the returning number of Observations is higher than 10, an update is made to the Main GUI to pop-up an alert window, warning the GUI user (e.g. the teacher) that the Student is exhibiting attention problems.

4.2.3 Data Compiler Processor

This module works as a watcher of new files in the database, as a data compiler and as an element of communication of the recorded Observations data and the Knowledge Processor.

It is instantiated by the Main GUI and, on launch, immediately starts monitoring the folder where the database files are stored, and connects to the Knowledge Processor as a TCP socket client.

When a new XML data file is created in the database directory, if it is no longer in use by another process, the Data Compiler reads its content and analyses its data, compiling all that file data into a single set of knowledge (by averaging the data values), formatted as an EML element, or in case of Affects and Behaviours, as XML elements with the identical formatting to EML. It then sends this knowledge to the Knowledge Processor TCP server, following the protocol described in section 4.2.2.

In theory, there could be one instance of the Data Compiler Processor for each sensor, analysing and compiling the data files respective to each device. In practice, only one Data Compiler Processor was implemented, in order to save time and hardware resources, and currently it can analyse and process data from two sensors, simultaneously: the Gazepoint GP3 Eye Tracker biometrics and the Affectiva's emotion and behaviour data.

Since the database files, from the Emotions Data Acquisition Processor, are already formatted using the stated protocol (using EML), this processor takes all the recorded measurements in one file and calculates the average for each category value, normalized to 0.000-1.000 range, thus compiling the results as knowledge of one Observation.

As for the database files of the Eye Tracker Data Processor, because they contain the raw biometrical data collected, a deeper analysis and correlation of each metric is required.

Due to time and availability constrains, only a measure of on/off task properties (described in section 3.2.3) was deployed. This was accomplished by assessing the point of gaze, as a percentage of the distance, to centre of the primary task area (the centre of the screen).

4.2.4 GP3 Eye Tracker, Gazepoint Control and Eye Tracker Data Acquisition Processor

The GP3 Eye Tracker, Gazepoint Control and Eye Tracker Data Acquisition Processor modules represent the data influx, collected from the student gaze, to the creation of files in the Database.

GP3 Eye Tracker

The Eye Tracker used was the Gazepoint GP3 Eye Tracker [20].

This eye tracker has a sample rate of 60Hz and provides the following metrics:

- Fixation POG;
- Left and right eyes POG;
- Best POG, which is the average POG data from the left and right eyes if both are available, or just either one, depending on which one is valid;
- Left and right eyes pupil coordinates and diameters, in pixels and in meters;
- Blink duration and count per minute;

It connects to a computer through Universal Serial Bus (USB) and requires two powered ports to work.

Due to the delay in the provision of this sensor, the basic metrics it provides, and the low accuracy results, the study and development of cognitive and affective metrics was not feasible.

Gazepoint Control

Gazepoint Control is a proprietary middleware that is required to collect the captured data from the GP3 Eye Tracker.

It allows TCP client to connect (using port 4242), and has a set of commands it can receive, to configure the eye tracker data and control its usage.

The Gazepoint Control GUI, shown in Figure 4-3, offers the video overview being captured by the eye tracker, an option to select the screen in use, and a calibration option, where a 5 or 9 point calibration is possible. However, the eye gaze video, captured by the eye tracker, is not recorded.

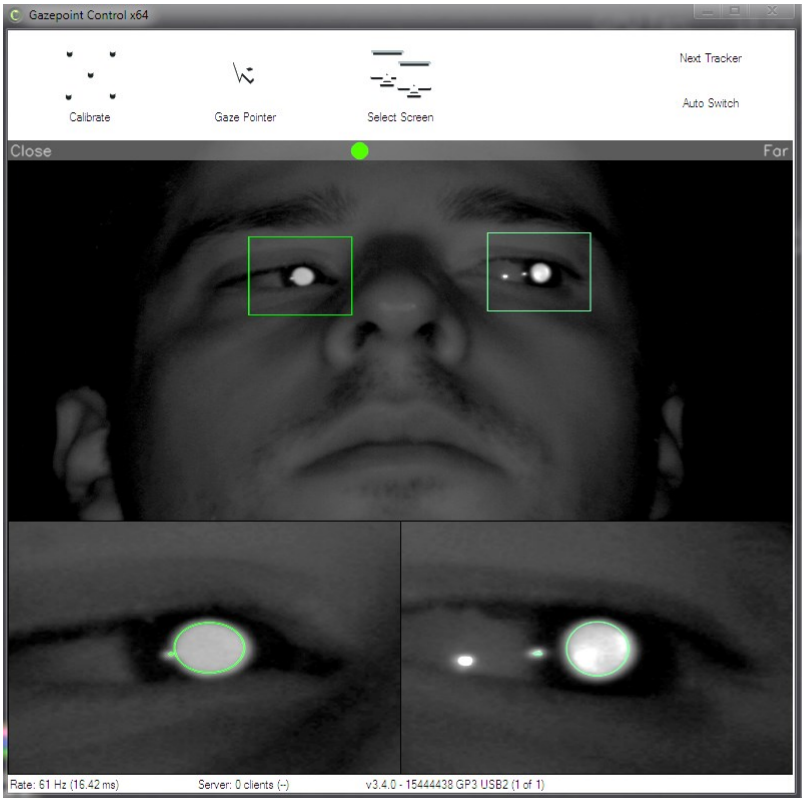

Figure 4-3 Gazepoint Control

Eye Tracker Data Acquisition Processor

The Eye Tracker Data Acquisition Processor is the module used to connect with the Gazepoint Control, and record the eye tracker data in the Database as XML files.

This module offers the option to configure which categories of data records are received from the Gazepoint Control TCP server, to connect/disconnect from the server and to start/stop the data stream.

The TCP stream is received as a string in XML format, and contains the metrics provided by the eye tracker as well as a record counter, the time (in computer ticks and in seconds), the cursor position and optional user data.

As each stream string is received, it is written in a XML file, named with the device name ("GP3") and with the starting date and time of the Observation, which is stored in a "Records" directory representing the Database.

The strings keep being written in the file until it contains all the records made during the time specified by the file creation sample rate variable. This sample rate variable is predefined to 60 seconds but can be changed as a launch argument when the processor starts.

4.2.5 Webcam, Affectiva SDK and Emotions Data Acquisition Processor

The Webcam, Affectiva SDK and Emotions Data Acquisition Processor modules represent the data influx, collected from the analysis of facial expressions recorded by a webcam, to the creation of files in the Database.

The Emotions Data Acquisition Processor integrates the Affectiva SDK, through the use of the affdex.dll and affdex-native.dll libraries.

At launch, this module connects to a computer webcam, and hands over the video stream to the Affectiva algorithm. This algorithm locates the human face in the video stream, analyses the facial points extracting facial expressions, and uses deep learning to correlate all data to the emotions displayed by the recorded person. To avoid privacy concerns, the captured webcam video stream is not recorded.

The emotions returned by Affectiva are: Anger, Contempt, Disgust, Fear, Joy, Sadness and Surprise. These metrics are stated to have a Receiver Operating Characteristic (ROC) of 0.8 or above [72]. A measurement of Engagement and Attention are also provided, although the resulting accuracy for these two metrics is not disclosed, and there for they should be considered with the appropriate cautions.

For each frame that is processed by the Affectiva algorithm, the returned set of metrics is saved to a XML file, following the EML document schema, effectively creating an EML document. This file is named with the algorithm name ("Affectiva") and with the starting date and time of the Observation, and is stored in a "Records" directory representing the Database.

Just like in the Eye Tracker Data Acquisition Processor, each file contains all the records made during the time specified by the file creation sample rate variable. This sample rate variable is predefined to 60 seconds but can be changed as a launch argument when the processor starts.

4.2.6 Main GUI

The platform application offers a user interface, that enables the manual querying of the data stored in the ontology, as well as the option to manually insert data into the same ontology, through the Knowledge Processor.

Using the Main GUI, the user can change the Ontology file in use to a different one, select or insert a new Session, User (either a Student or a Teacher), and select Observations (either Human_Observation or Digital_ Observation), or insert a new Human_Observation, along with an Emotion, Behaviour and/or Affect.

It can also be used to launch the Data Acquisition and Data Compiler processors, and to set the "File Creation Sample Rate", which is the variable that defines the time window, containing the recordings made during that time, that each file can have. Along with the launch of the Data Acquisition Processors it is also launched the middleware for the respective device (if there is any).

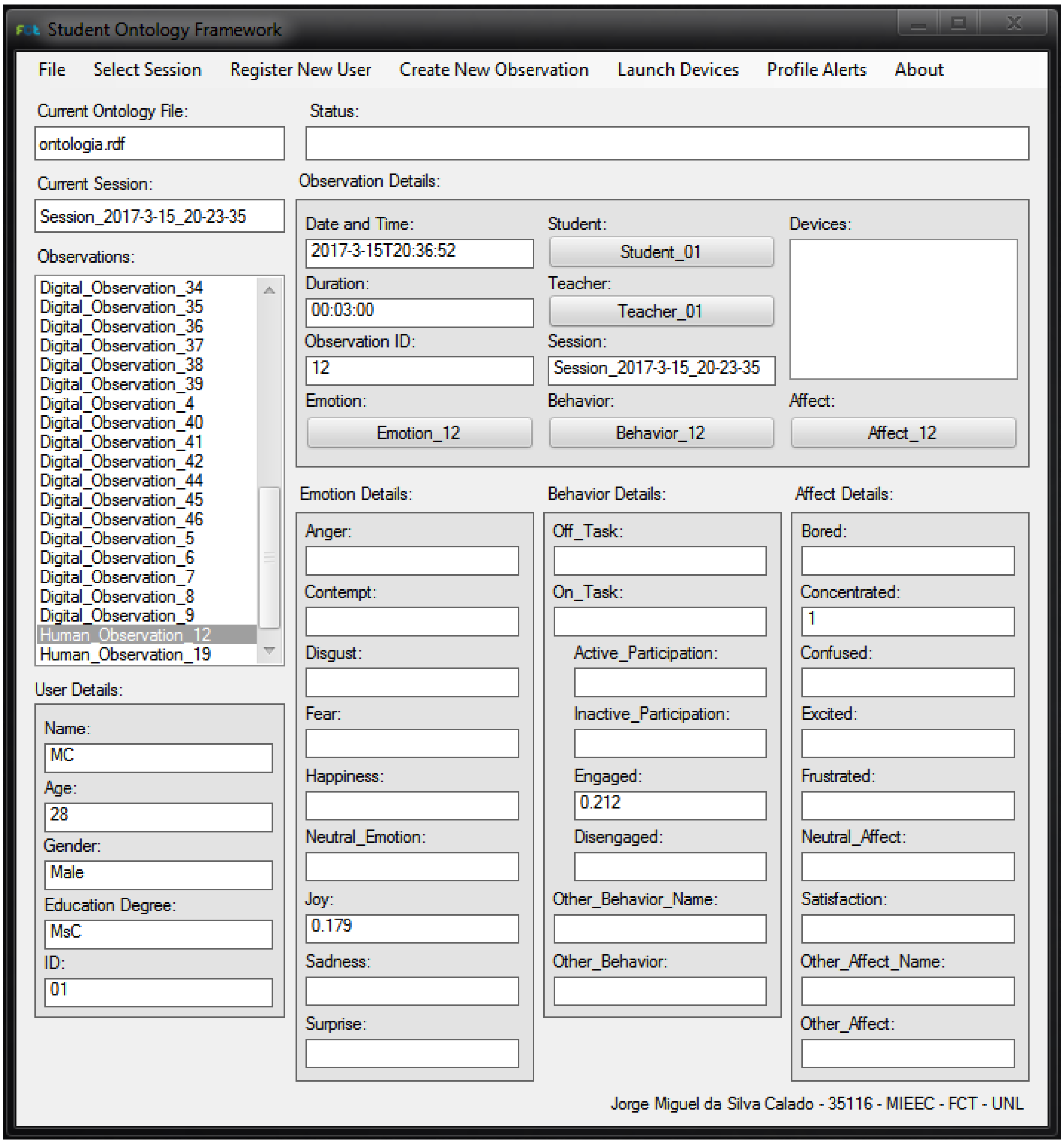

Figure 4-4 Main GUI

Figure 4-4 is a screen caption of the Main GUI window, where it is possible to see some of the options available to manipulate the Ontology, as well as to Launch the sensorial devices and set the Profile alerts. It is also shown the list of Observations, and the properties available for consultation: the Observation properties, the User (student) data and the Emotion, Behaviour and Affect values.

4.2.7 Software Licenses and Hardware Requirements

The Gazepoint software is available under a Creative Commons Zero license [73].

The dotNetRDF framework is available under the MIT license [74].

The Affectiva SDK, the Affdex SDK and the Affdex Code used in this architecture falls under the Affectiva free license agreement that is available at [75].

The current use of the implemented architecture, as part of a dissertation work, respects all the licenses previously mentioned.

The minimum requirements, combined for all components implemented by this architecture, are the following:

- Quad-core i5 or equivalent processor;
- 8 Giga Bytes of Ramdom-access Memory;
- 2 powered USB ports;
- 1 webcam;
- Windows 7 (or above) 64bits operating system;
- .Net 4.5 Framework;

4.2.8 Application Demonstration

The Main GUI window, illustrated in Figure 4-4, is the first point of contact between the Teacher, or Expert, and the developed platform. Through this GUI the quick and easy access to each student information is facilitated, as well as the availability of tools to register new information, such as Student and Teacher data, Sessions, Human_Observations, Emotions, Behaviours and Affects, as shown in Figure 4-5, Figure 4-6, Figure 4-7 and Figure 4-8.

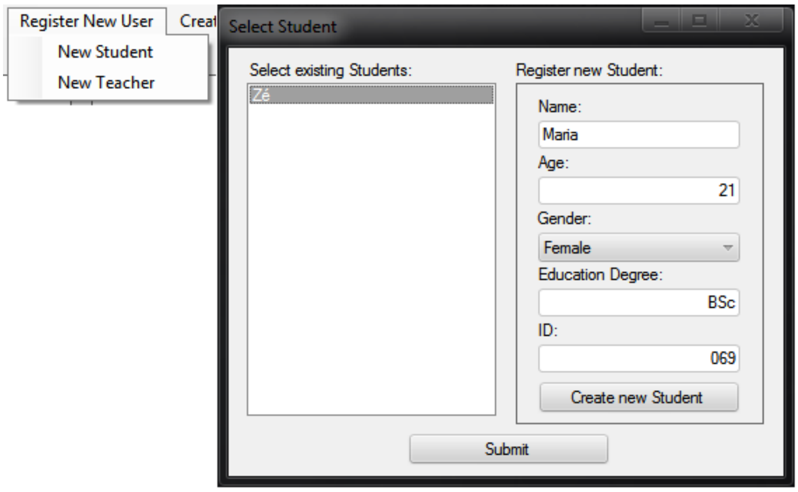

Figure 4-5 User Selection/Creation

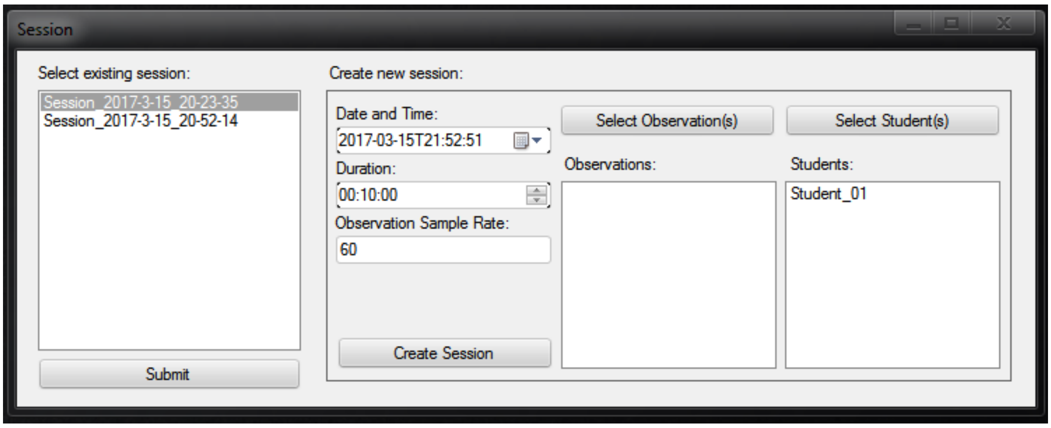

Figure 4-6 Session Selection/Creation

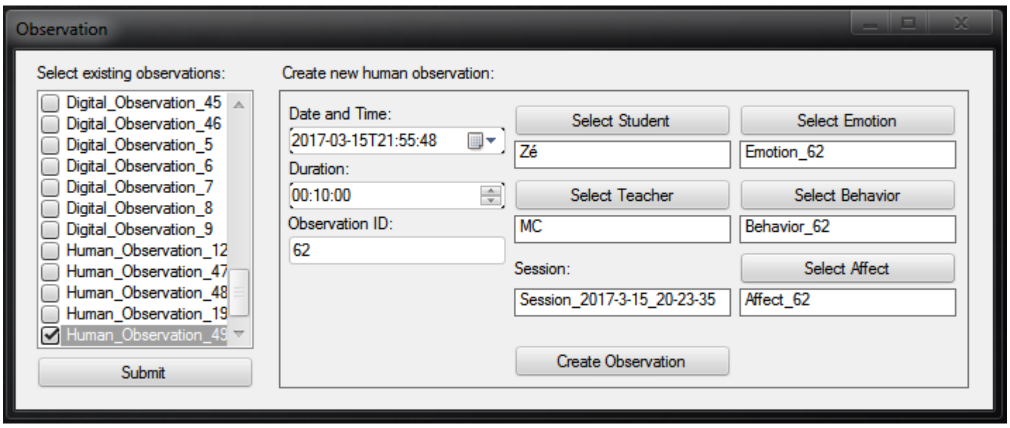

Figure 4-7 Observation Selection/Creation

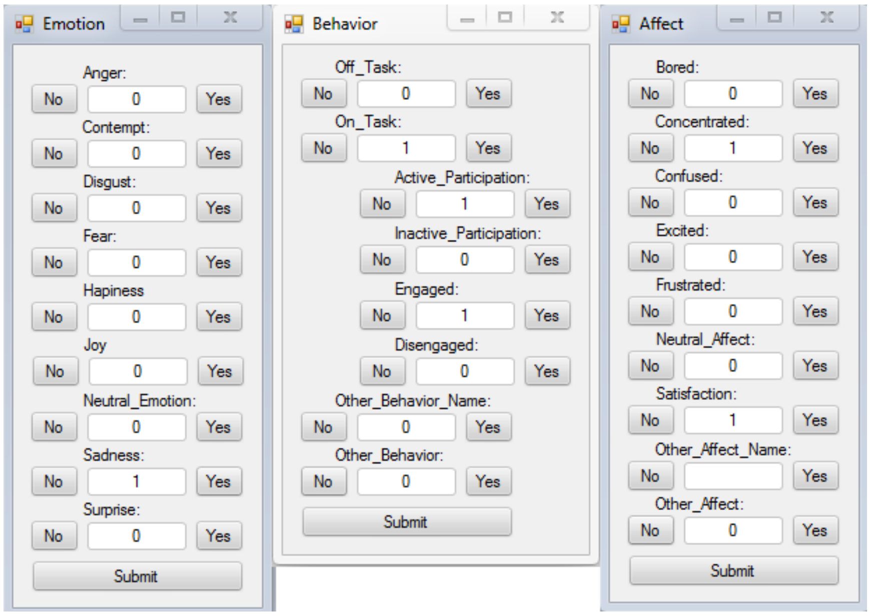

Figure 4-8 Manual Registration of Emotion, Behaviour and Affect Values

Also in the Main GUI is possible to launch the “Data Processor” modules, displayed in Figure 4-9, used to automatically acquire the behavioural, emotional and affective data from the student.

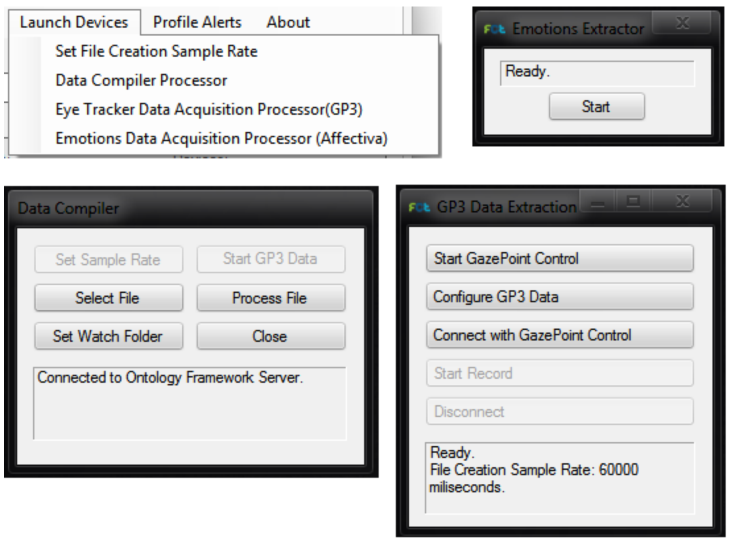

Figure 4-9 Data Processor Modules

Each module can run simultaneously, along with the registration of new Observations in the Ontology, making the student data available practically in real-time.

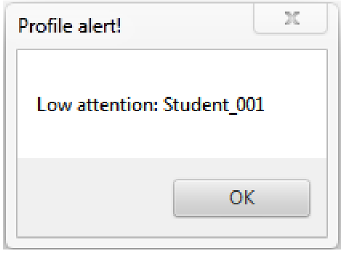

Figure 4-10 Profile Alert

Since with every new Observation made, the conditions used to set the Profiles are checked, when the evaluation of those conditions is authenticated, a pop-up window displays the alert message, informing the teacher for the occurrence of a potential problematic situation, as illustrated in Figure 4-10.

5 Proof-of-Concept Prototype

This chapter describes the experiment, used to test the proof-of-concept prototype, which intended to validate the framework proposed in section 4.1, and its architecture implementation, presented in section 4.2. The experiment included a learning environment simulation, as proposed in the application scenario, in section 3.1, including case studies 2, 3 and 4, applied to single student sessions.

5.1 Method

The experiment consisted in a simulated test of an e-learning environment, where the participants were asked to read two excerpts of this document (from the preceding chapters), while their emotions, behaviours and affects were being recorded, by the sensors and by an expert.

Each of the two excerpt readings was timed, and in one of them (the first one to half the participants, and the second one for the other half), external stimuli were introduced in the experiment, to induce distraction and disrupt the person's task, in order to record a situation resembling the conditions set for the Attention_Disorder Profile.

At the end of each excerpt reading, the participant was asked to answer five questions about its content, to assess the results of their attention level, during the reading.

5.1.1 Material

To perform the experience in optimal conditions, two computers were used (one desktop and one laptop). One was used by the participant to interact with the testing environment of the assigned task, while the other one was used by an expert, for simultaneous recording of the observations made.

The GP3 Eye Tracker was placed under the desktop monitor, pointing out directly to the participant eyes, and connected to the laptop, using two USB 2.0 ports.

A USB webcam was placed on top of the desktop monitor, pointed at the participant face, and connected to a USB port in the laptop.

5.1.2 Participants

The experiment participants were 3 male and 3 female individuals, aged between 28 and 30, all of them having a Master of Science (MSc) degree, in different engineering areas. 5 out of 6 participants wore glasses. All the participants reported a level of understanding of written English, equal or above C1, in the Common European Framework of Reference for Languages [76].

One expert, Margarida Calado, who has a MSc degree in Educational Psychology, was present during each experiment, observing and recording the behaviour and affective state of the tested participants, in order to provide ground-truth label Observations, named Human_Observations.

5.1.3 Apparatus

Before the experiment, the participants were asked to sit in a desk, facing the desktop monitor, in a comfortable but straight position. For each participant, the webcam and eye tracker orientation angles were adjusted, to be more accurately pointed at the participant's face.

The eye tracker was calibrated per each individual, using the Gazepoint Control software, using a 5-points setup. After confirming the validity of the calibration, the expert sat in a discrete position near the participant, facing him/her, unobtrusively.

Figure 5-1 shows the apparatus setup used in this experiment, comprising the GP3 Eye Tracker (highlighted in blue) and a webcam (highlighted in green), attached to the screen used by the participant to perform the tasks, and the expert laptop, used to control the experiment and annotate the expert observations

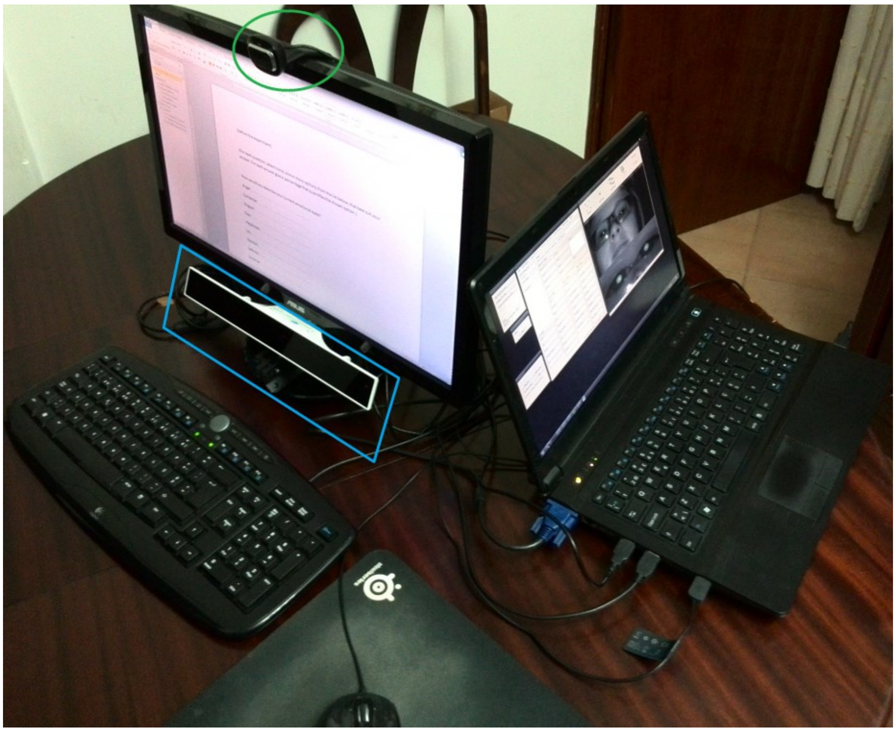

Figure 5-1 Experiment Apparatus Setup

5.1.4 Procedure

The participants were asked if they preferred anonymity in the treatment of their personal data, like the name, age, gender, education degree and recorded Observations.

It was explained that the experiment consisted in two excerpt readings of a dissertation document, using the computer, each timed with 10 minutes, and after each one, there would be a small quiz about the content of the excerpt. Text editing tools were also made available to the participants, in case they wanted to take annotations.

Out of the two excerpts given to each participant, one was selected to be targeted with external stimuli (to half the participants it was the first one, and for the other half it was the second). The external stimuli included diverting the participant attention with a side conversation, turning on loud music, showing engaging videos close to the participant screen, among other distractions. During the other excerpt reading, the opposite environment was attempted to be provided, compelling to concentration and attention to the task given.

Throughout each 10-minute reading, the expert observed and recorded any changes in the participant emotions, behaviours and affects, using the platform Main GUI, running in the laptop.

In the end of each excerpt reading, the participants were asked a five-question quiz, related to specific details of the excerpt read.

For the first excerpt (excerpt A), the text chosen had 2488 words, and included the Abstract, the Introduction, and section 2.1, of the State of the Art of this document. The five questions for this excerpt were the following:

- Q1 – "What does ICT stand for?"
- Q2 – "What are the primary goals of the ACACIA project?"
- Q3 – "In which module of the ACACIA project, is the work of this dissertation integrated?"
- Q4 – "What is the research method used in this dissertation?"
- Q5 – "Enumerate the sensorial devices mentioned that can be used to detect emotions. (as many as you can)"

For the second excerpt (excerpt B), the text chosen had 2652 words, and included sections 2.2, 2.3, 2.4 and 2.5, of the State of the Art of this document. The five questions for this excerpt were the following:

- Q6 – "What does an eye tracker measures?"
- Q7 – "What is smooth pursuit eye movement?"
- Q8 – "What is approximately the difference in precision from low-end and high-end eye trackers?"
- Q9 – "What device is described to be used in gait recognition?"
- Q10 – "What are the three stages, described in the document, deployed by the emotion detection algorithms?"

All participants read the excerpt A first, and excerpt B secondly.

5.1.5 Results

Since the participants were divided in half, one half to receive distractive stimuli only during the reading of excerpt A, and the other half to receive the same distractive stimuli only during the reading of excerpt B, the first half is hereinafter called "group 1", and the second half called "group 2".

Using the word count number of the text read by the participants and the evaluation of the answers given about each respective excerpt, detailed in Table 5-1, it is possible to confirm that every participant read significantly fewer words and overall performed poorly, when reading the excerpt subjected to stimuli, thus validating the premise that during the stimulated excerpt their attention was noticeably lower. This fact was also confirmed by the expert observations.

Table 5-1 Evaluation of Participants Results

Partici-pant	Excerpt	Used Stimuli	Total Words Read	Answers Quotation (%)									
				Excerpt A					Excerpt B				
				Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
1	A	No	1182	100	40	100	0	0	-	-	-	-	-
	B	Yes	933	-	-	-	-	-	100	100	0	0	0
2	A	Yes	1304	0	90	50	100	0	-	-	-	-	-
	B	No	1918	-	-	-	-	-	100	0	100	100	10
3	A	No	2006	0	100	100	0	0	-	-	-	-	-
	B	Yes	1258	-	-	-	-	-	100	100	0	0	0
4	A	Yes	962	100	40	0	0	100	-	-	-	-	-
	B	No	1916	-	-	-	-	-	100	100	100	100	0
5	A	No	2488	5	60	10	0	100	-	-	-	-	-
	B	Yes	1940	-	-	-	-	-	20	0	0	0	0
6	A	Yes	674	0	100	0	0	50	-	-	-	-	-
	B	No	1868	-	-	-	-	-	100	20	10	100	0

On a curious note, participants of group 2 performed higher than group 1, in the quiz evaluation of the non-stimulated excerpt.

The analysis of the measures detected by the eye tracker (Off_Task), and by the emotion detection algorithm (Anger, Contempt, Disgust, Fear, Joy, Sadness, Surprise and Engagement), are detailed in Table 5-2. The analysis of the observations made by the expert are detailed in Table 5-3, and although in some emotions the average value is always zero, it only means that the expert did not observe any of those emotions during all the experiments.

The analysis of the Behaviour properties shows a good indicator for the detection of attention problems, being Off_Task, reported by the eye tracker, a suitable property to be used as a threshold for setting this Profile condition. The difference in the Off_Task average value, comparing the stimulated excerpt to the non-stimulated one, ranged between 8.8% (participant 3) and 25.9%. (participant 4).

This conclusion falls in line with the assumptions made for the Attention_Disorder Profile conditions, proposed in section 3.2.3, and with the Human_Observations made by the expert. However, the average Engaged values were not consistent with the Observations made by the expert. As stated in section 4.2.5, because no information is provided about this measure accuracy, which proved to be inaccurate during these tests, the Engaged property will be disregarded from the analysis.

Table 5-2 Average Values Measured by Digital Observations

		Average Values Measured (0-1 range)											
Participant		**1**		**2**		**3**		**4**		**5**		**6**	
Excerpt		A	B	A	B	A	B	A	B	A	B	A	B
Stimuli		No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No
Emotion	Anger	0.0000	0.0000	0.0098	0.0165	0.0001	0.0000	0.0000	0.0000	0.0000	0.0005	0.0001	0.0000
	Contempt	0.0147	0.0224	0.0612	0.0026	0.0653	0.0971	0.0479	0.0142	0.0030	0.0379	0.0103	0.0023
	Disgust	0.0062	0.0063	0.0066	0.0040	0.0041	0.0066	0.0040	0.0042	0.0060	0.0198	0.0115	0.0042
	Fear	0.0012	0.0050	0.0006	0.0008	0.0000	0.0004	0.0005	0.0001	0.0002	0.0003	0.0014	0.0000
	Joy	0.0208	0.1522	0.0316	0.0000	0.0000	0.1220	0.0621	0.0269	0.0016	0.0958	0.1712	0.0000
	Sadness	0.0006	0.0007	0.0088	0.0138	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0014	0.0000
	Surprise	0.0476	0.1162	0.0169	0.0018	0.0024	0.0329	0.0276	0.0081	0.0092	0.0364	0.0553	0.0030
Behaviour	Off_Task	0.2146	0.3510	0.3900	0.2949	0.2650	0.3536	0.4230	0.1632	0.2535	0.3748	0.4116	0.2944
	Engaged	0.1197	0.3310	0.0829	0.0305	0.0258	0.1850	0.1134	0.0391	0.0289	0.2102	0.2358	0.0036

Table 5-3 Average Values Observed by the Expert

		Average Values Observed (0-1 range)											
Participant		**1**		**2**		**3**		**4**		**5**		**6**	
Excerpt		A	B	A	B	A	B	A	B	A	B	A	B
Stimuli		No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No
Emotion	Anger	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Contempt	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Disgust	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Fear	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Joy	0.0000	0.0545	0.1455	0.0000	0.0000	0.0909	0.0364	0.0909	0.0000	0.1182	0.0000	0.0000
	Sadness	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Surprise	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Behaviour	Off_Task	0.0000	0.1727	0.1818	0.0000	0.0000	0.3909	0.2727	0.0000	0.0000	0.2636	0.3818	0.0000
	Engaged	1.0000	0.7818	0.7818	1.0000	1.0000	0.5909	0.7636	1.0000	1.0000	0.7727	0.6909	1.0000

The analysis to the Emotion properties, detected with the sensors, revealed more significant changes in the average values for Contempt, Joy and Surprise, on most participants, and is a good indicator of these emotions potential value.

The change in the average value, between each different reading, was most significant with Joy, ranging between 3.2% (participant 3) and 17.1% (participant 6). Surprise average values, between each different reading, changed between 1.5% (participant 2) and 6.9% (participant 1). Contempt average values, between each different reading, changed between 0.8% (participants 1 and 6) and 5.9% (participant 2).

Comparing the emotions detected through the sensors with the observations made by the expert, it is possible to conclude that the average values for Joy were very similar, while the other emotions were not. This lead to the conclusion that the detection of Joy is a valid measurement, and is suitable to be used for setting the conditions of the Attention_Disorder Profile, contrasting with some assumptions made in Table 3-1.

5.2 Validation

The developed platform, based in the Framework for Student's Profile Detection, proved capable of handling different concurrent sensors, used for the automatic detection of student's emotions and behaviours, effectively creating a knowledge based student monitoring solution. With a Graphical User Interface suited to connect the teachers with the students' affective needs, through a real-time alert system, capable of warning for potential problematic profile situations, when set conditions are detected by the sensors.

Therefore, the detection of changes in student's behaviours is greatly facilitated, which is a step forward in the prevention of drop-out intentions and academic disengagement, thus validating the hypothesis, proposed in section 1.3.2.

The proposed application scenario was validated in the publication of a co-authored paper, called "Affective Computing to Enhance Emotional Sustainability of Students in Dropout Prevention", presented in "DSAI 2016 – International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion" [77]. This paper is attached to this dissertation document as Annex A.

The platform Ontology was reviewed and approved by Educational Psychologist, Dr. Maria de Lourdes Mata, from *Instituto Superior de Psicologia Aplicada* (ISPA).

The presented framework, its architecture implementation and the experimental results will also be submitted in a paper, called “A framework to bridge teachers, students affective state, and improve academic performance”, for the “IMECE 2017 – International Mechanical Engineering Congress and Exposition”, with the abstract already accepted.

The work developed has been discussed in project ACACIA, and these results are already being used in the CADEP centres, in Colombia, Nicaragua, and Peru.

6 Conclusions

The research work, presented in this dissertation, was developed as a response to a growing problem in HEI, concerning students' engagement and learning problems.

As society shifts to make use of more technological aids, the need to adapt the academic context, to include these benefits, is demanded. The proposed framework is one of some possible answers to mitigate these needs.

Designed to be interoperable with different types of sensorial systems, the framework proved capable of detecting the student's behaviour, emotional and affective states, and to manage all the collected information into a knowledge-based system, able to pre-emptively and pro-actively detect situations consistent with profiles of students' problems.

Following the framework guidelines, the implementation and deployment of a system, using an eye tracker for behaviour identification, the recognition of emotions through facial expression analysis software and recorded observations from experts in educational psychology, was successfully achieved and validated.

This system capability for real-time profile warnings, is a valuable asset to assist teachers identifying problems, during the students learning process, and to help in the prevention of school drop-outs.

6.1 Future Work

Although the primary goal of this dissertation was successfully achieved, through the validation of the hypothesis set in the beginning of this work, it was partially incomplete, mainly due to timing constrains, that prevented the timely implementation of a gait and posture detection system, using the Kinect sensor, the integration of a machine learning algorithm for pattern outlier analysis, as a recommendation system for profile detection, and the implementation and experimental validation of the Sociological_Issue and Drop_Out Profiles' detection.

It is expected that future work can integrate the implemented architecture with the mentioned devices, as well as expand integration to other sensors, given the interoperability of the framework platform, by the usage of the EML standard. It is also expected that the correlation, between the Digital_Observation and Human_Observation values, can be used in a machine learning implementation, that attempts to predict the detection of the Profiles, further enabling the prevention of drop-out intentions and academic disengagement.

Due to the sensitive information handled in the platform, the database files and processes' communication will require end-to-end encryption, in order to provide student's privacy, required for the deployment in a real-world scenario.

Given the proof-of-concept environment, and the obvious hardware limitations, the developed platform was only tested in a single student session, even though it supports multiple sensors and students.

A future step in scientific validation would be the deployment in a larger environment, like a real classroom, where significant data could be collected, further improving the machine learning dataset and, subsequently, the algorithm itself, for the more accurately detection of academic problematic profiles, thus fulfilling the objectives set for the *Apoya* module, of the ACACIA project

References

[1] W. J. Pelgrum, "Obstacles to the integration of ICT in education: Results from a worldwide educational assessment," *Comput. Educ.*, vol. 37, no. 2, pp. 163–178, 2001.

[2] U.S. Department of Education, "Future Ready Learning: Reimagining the role of technology in education," p. 101, 2016.

[3] E. Commission, *Modernisation of Higher Education*, no. June. 2014.

[4] "Use of Technology in Teaching and Learning | U.S. Department of Education." [Online]. Available: http://www.ed.gov/oii-news/use-technology-teaching-and-learning. [Accessed: 17-Oct-2016].

[5] L. M. Camarinha-matos, "Unit 2 : SCIENTIFIC METHOD," *Sci. Res. Methodol. Tech.*, pp. 2009–2012, 2012.

[6] J. Broekens, D. DeGroot, and W. A. Kosters, "Formal models of appraisal: Theory, specification, and computational model," *Cogn. Syst. Res.*, vol. 9, no. 3, pp. 173–197, 2008.

[7] D. L. Schacter, "Psychology Second Edition, 41 Madison Avenue, New York, NY 10010." Worth Publishers, 2011.

[8] K. R. Scherer, "Psychological models of emotion.," in *The Neuropsychology of Emotion*, Joan C. Bo., New York: Oxford University Press, 2000, pp. 137–162.

[9] J. Lin, M. Spraragen, and M. Zyda, "Computational models of emotion and cognition," *Adv. Cogn. Syst.*, vol. 2, pp. 59–76, 2012.

[10] P. Ekman, "What Scientists Who Study Emotion Agree About," *Perspect. Psychol. Sci.*, vol. 11, no. 1, pp. 31–34, Jan. 2016.

[11] K. R. Scherer, "What are emotions? And how can they be measured?," *Soc. Sci. Inf.*, vol. 44, no. 4, pp. 695–729, Dec. 2005.

[12] J. A. Russell, "A circumplex model of affect.," *J. Pers. Soc. Psychol.*, vol. 39, no. 6, pp. 1161–1178, 1980.

[13] R. Plutchik, *The emotions*. University Press of America, 1991.

[14] J. R. J. Fontaine, K. R. Scherer, E. B. Roesch, and P. C. Ellsworth, "The World of Emotions is not Two-Dimensional," *Psychol. Sci.*, vol. 18, no. 12, pp. 1050–1057, Dec. 2007.

[15] "Facial Expressions - Eye Tracking Software and Solutions - iMotions." [Online]. Available: https://imotions.com/facial-expressions/. [Accessed: 28-Feb-2016].

[16] J. Ocumpaugh, R. S. J. D. Baker, and M. M. T. Rodrigo, "Baker Rodrigo Ocumpaugh Monitoring Protocol (BROMP) 2.0 Technical and Training Manual," 2015.

[17] P. Baggia, Catherine Pelachaud, Christian Peter, and E. Zovato, "Emotion Markup Language (EmotionML) 1.0." [Online]. Available: https://www.w3.org/TR/emotionml/.

[Accessed: 13-Oct-2016].

[18] R. G. Lupu and F. Ungureanu, “A survey of eye tracking methods and applications,” *Math. Subj. Classif.*, vol. LXIII, no. 3, pp. 72–86, 2013.

[19] “myGaze Eye Tracker.” [Online]. Available: http://www.mygaze.com/products/mygaze-eye-tracker/. [Accessed: 13-Mar-2016].

[20] “GP3 Eye Tracker | Hardware Only – Gazepoint.” [Online]. Available: http://www.gazept.com/product/gazepoint-gp3-eye-tracker/. [Accessed: 13-Mar-2016].

[21] “The Eye Tribe.” [Online]. Available: https://theeyetribe.com/products/. [Accessed: 13-Mar-2016].

[22] “Tobii EyeX.” [Online]. Available: http://www.tobii.com/xperience/products/. [Accessed: 13-Mar-2016].

[23] E. S. Dalmaijer, “Is the low-cost EyeTribe eye tracker any good for research?,” *PeerJ Prepr.*, vol. 4, no. 606901, pp. 1–35, 2014.

[24] S. Zugal and J. Pinggera, “Low–Cost Eye–Trackers: Useful for Information Systems Research?,” *Adv. Inf. Syst. Eng. Work.*, pp. 159–170, 2014.

[25] “Cognitive Workload, the only completely unobtrusive way to measure workload.” [Online]. Available: http://www.eyetracking.com/Software/Cognitive-Workload. [Accessed: 28-Feb-2016].

[26] D. Bernhardt, “Human Body Motion,” no. 787, 2010.

[27] E. Crane and M. Gross, “Motion Capture and Emotion : Affect Detection in Whole Body Movement,” *Affect. Comput. Intell. Interact.*, pp. 95–101, 2007.

[28] D. Glowinski, A. Camurri, G. Volpe, N. Dael, and K. Scherer, “Technique for automatic emotion recognition by body gesture analysis,” *2008 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. CVPR Work.*, no. July 2016, 2008.

[29] A. Barliya, L. Omlor, M. A. Giese, A. Berthoz, and T. Flash, “Expression of emotion in the kinematics of locomotion,” *Exp. Brain Res.*, vol. 225, no. 2, pp. 159–176, 2013.

[30] B. Bonnechère, B. Jansen, P. Salvia, H. Bouzahouene, L. Omelina, F. Moiseev, V. Sholukha, J. Cornelis, M. Rooze, and S. Van Sint Jan, “Validity and reliability of the Kinect within functional assessment activities: Comparison with standard stereophotogrammetry,” *Gait Posture*, vol. 39, no. 1, pp. 593–598, 2014.

[31] J. Preis, M. Kessel, M. Werner, and C. Linnhoff-Popien, “Gait Recognition with Kinect,” *Work. Kinect Pervasive Comput. Pervasive 2012*, no. July 2015, pp. 1–5, 2012.

[32] M. Gabel, R. Gilad-Bachrach, E. Renshaw, and A. Schuster, “Full body gait analysis with Kinect,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, no. June 2016, pp. 1964–1967, 2012.

[33] A. Sinha, K. Chakravarty, and B. Bhowmick, “Person identification using skeleton information from kinect,” *ACHI 2013, Sixth …*, no. c, pp. 101–108, 2013.

[34] R. A. Clark, Y. H. Pua, K. Fortin, C. Ritchie, K. E. Webster, L. Denehy, and A. L. Bryant, “Validity of the Microsoft Kinect for assessment of postural control,” *Gait Posture*, vol. 36, no. 3, pp. 372–377, 2012.

[35] E. Stone and M. Skubic, “Evaluation of an Inexpensive Depth Camera for Passive In-Home Fall Risk Assessment,” *Proc. 5th Int. ICST Conf. Pervasive Comput. Technol. Healthc.*, vol. 3, no. 4, pp. 349–361, 2011.

[36] P. Ekman and W. V Friesen, “Constants across cultures in the face and emotion,”

Journal of personality and social psychology, vol. 17, no. 2. pp. 124–129, 1971.

[37] P. Ekman and W. V. Friesen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. 1978.

[38] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, 2004.

[39] B. Fasel and J. Luettin, "Automatic facial expression analysis: A survey," *Pattern Recognit.*, vol. 36, no. 1, pp. 259–275, 2003.

[40] M. Pantic and M. Bartlett, "Machine analysis of facial expressions," 2007.

[41] M. S. Bartlett, G. Littlewort, I. Fasel, and J. R. Movellan, "Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction.," *2003 Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 5, no. July, pp. 53–53, 2003.

[42] Affectiva, "Emotion as a Service," 2016. [Online]. Available: http://www.affectiva.com/solutions/apis-sdks/. [Accessed: 07-Jun-2016].

[43] Noldus, "FaceReader," 2016. [Online]. Available: http://www.noldus.com/human-behavior-research/products/facereader. [Accessed: 07-Jun-2016].

[44] NViso, "EmotionAdvisor," 2016. [Online]. Available: http://www.nviso.ch/index.html. [Accessed: 07-Jun-2016].

[45] Microsoft, "Microsoft Cognitive Services," 2016. [Online]. Available: https://www.microsoft.com/cognitive-services/en-us/emotion-api. [Accessed: 07-Jun-2016].

[46] V. Cantoni, C. J. Perez, M. Porta, and S. Ricotti, "Exploiting eye tracking in advanced e-learning systems," *Proc. 13th Int. Conf. Comput. Syst. Technol. - CompSysTech '12*, p. 376, 2012.

[47] M. Porta, S. Ricotti, and C. J. Perez, "Emotional e-learning through eye tracking," *IEEE Glob. Eng. Educ. Conf. EDUCON*, 2012.

[48] M. Pivec, C. Trummer, and J. Pripfl, "Eye-Tracking Adaptable e-Learning and Content Authoring Support," vol. 30, pp. 83–86, 2006.

[49] H. Wei, C. Muntean, H. Wei, A. Moldovan, and C. H. Muntean, "Sensing Learner Interest Through Eye Tracking," 2009.

[50] H. Wang, M. Chignell, and M. Ishizuka, "Empathic tutoring software agents using real-time eye tracking," *Acm Etra*, 2006.

[51] C. Calvi, M. Porta, D. Sacchi, and U. Pavia, "e5Learning, an E-Learning Environment Based on Eye Tracking," vol. 2008, no. 1, 2008.

[52] "OWL Web Ontology Language Overview." [Online]. Available: https://www.w3.org/TR/owl-features/. [Accessed: 08-Jan-2017].

[53] "RDF 1.1 Concepts and Abstract Syntax." [Online]. Available: https://www.w3.org/TR/rdf11-concepts/. [Accessed: 08-Jan-2017].

[54] "RDF Schema 1.1." [Online]. Available: https://www.w3.org/TR/rdf-schema/. [Accessed: 08-Jan-2017].

[55] "Protégé." [Online]. Available: http://protege.stanford.edu/. [Accessed: 08-Jan-2017].

[56] M. A. Musen, "The protégé project: a look back and a look forward," *AI Matters*, vol. 1, no. 4, pp. 4–12, Jun. 2015.

[57] "OWL 2 Web Ontology Language Document Overview (Second Edition)." [Online]. Available: https://www.w3.org/TR/owl2-overview/. [Accessed: 08-Jan-2017].

[58] "A Visual Notation for OWL Ontologies (VisioOWL)." [Online]. Available: https://sites.google.com/site/semanticsimulations2/visioowl. [Accessed: 10-Dec-2016].

[59] A. P. Association, "Diagnostic and statistical manual of mental disorders 5th edition American Psychiatric Publishing," *Arlington, VA, USA doi*, vol. 10, 2013.

[60] "Apache Jena." [Online]. Available: https://jena.apache.org/. [Accessed: 08-Jan-2017].

[61] mdesalvo, "RDFSharp - Start playing with RDF! - Home." [Online]. Available: http://rdfsharp.codeplex.com/. [Accessed: 08-Jan-2017].

[62] B. Pellens, "owldotnetapi." [Online]. Available: https://github.com/bpellens/owldotnetapi. [Accessed: 08-Jan-2017].

[63] "linqtordf." [Online]. Available: https://code.google.com/archive/p/linqtordf/. [Accessed: 08-Jan-2017].

[64] cognitum, "OWL API for .NET - Home." [Online]. Available: https://owlapinet.codeplex.com/. [Accessed: 08-Jan-2017].

[65] "dotNetRDF." [Online]. Available: http://www.dotnetrdf.org/. [Accessed: 08-Jan-2017].

[66] "SKOS Simple Knowledge Organization System Reference." [Online]. Available: https://www.w3.org/TR/skos-reference/. [Accessed: 09-Jan-2017].

[67] "Notation3 (N3): A readable RDF syntax." [Online]. Available: https://www.w3.org/TeamSubmission/n3/. [Accessed: 09-Jan-2017].

[68] "SPARQL 1.1 Overview." [Online]. Available: https://www.w3.org/TR/sparql11-overview/. [Accessed: 08-Jan-2017].

[69] "SPARQL 1.1 Update." [Online]. Available: https://www.w3.org/TR/sparql11-update/. [Accessed: 08-Jan-2017].

[70] "SPARQL 1.1 Evaluation Test Results." [Online]. Available: https://www.w3.org/2009/sparql/implementations/. [Accessed: 08-Jan-2017].

[71] "2.4 FSRE categories." [Online]. Available: https://www.w3.org/TR/2011/WD-emotion-voc-20110407/#fsre-categories. [Accessed: 02-Feb-2017].

[72] "Affectiva Developer Portal." [Online]. Available: http://developer.affectiva.com/accuracy/. [Accessed: 13-Jan-2017].

[73] "Creative Commons Legal Code." [Online]. Available: https://creativecommons.org/publicdomain/zero/1.0/legalcode. [Accessed: 20-Mar-2017].

[74] "The MIT License | Open Source Initiative." [Online]. Available: https://opensource.org/licenses/MIT. [Accessed: 20-Mar-2017].

[75] "Affectiva Developer Portal." [Online]. Available: http://developer.affectiva.com/sdklicense/. [Accessed: 20-Mar-2017].

[76] "Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)." [Online]. Available: http://www.coe.int/t/dg4/linguistic/cadre1_en.asp. [Accessed: 20-Mar-2017].

[77] M. Kadar and J. Calado, "Affective Computing to Enhance Emotional Sustainability of Students in Dropout Prevention," *DSAI 2016 Proc. 7th Int. Conf. Softw. Dev. Technol. Enhancing Access. Fight. Info-exclusion*, pp. 85–91, 2016.

Annex A – Paper: Affective Computing to Enhance Emotional Sustainability of Students in Dropout Prevention

Affective Computing to Enhance Emotional Sustainability of Students in Dropout Prevention

Manuella Kadar[1]
mkadar@uab.ro

Jorge Calado
jmd.calado@campus.fct.unl.pt

[1] Decembrie 1918 University, 5, Gabriel Bethlen street, 510009 Alba Iulia, Romania

Emmanuelle Gutiérrez y Restrepo[2]
egutierrez@dia.uned.es

Andreia Artifice
afva@uninova.pt

[2] aDeNu Research Group, Artificial Intelligence Department, Computer Science School, UNED, C/Juan del Rosal, 16. 28040 Madrid, Spain

Fernando Ferreira[3]
flf@uninova.pt

Joao Sarraipa[3]
jfss@uninova.pt

Ricardo Jardim-Goncalves[3]
rg@uninova.pt

[3] CTS, UNINOVA. DEE/FCT - Universidade NOVA de Lisboa, Portugal

ABSTRACT

In ACACIA project, through its Apoya module seeks to provide means and methods to enhance emotional sustainability as innovative approach to student's dropout prevention. Emotional state of the students at risk of dropout have to be assessed and innovative methods for counselling and curricula adaptation should be applied for getting out the student from the risk zone.

The aim of this study is to propose an innovative solution to meliorate both emotional state and attention of students in risk of dropout. A scenario is presented in which eyetrackers and webcams are integrated in a platform in order to infer and manage students' emotional state in a smart classroom environment.

CCS Concepts

• Applied computing → Education → Computer-assisted instruction.

Keywords

Affective computing, sensors, emotional state detection, student counselling.

1. INTRODUCTION

Cultural and social backgrounds of students may strongly influence the educational process and often lead to marginalization and social exclusion of certain students from meaningful participation in learning activities and community life. Such exclusion further reduces students' perspectives to learn, grow, and develop. Adapted educational systems facilitating the modern level of knowledge and skills are crucial components of positive change and successful development of the society. The use of technologies is not the only requirement of the new century. Educational planning and policy-making are also of great importance. Any educational policy must be able to meet diverse challenges and enable everyone to find his/her place in the community, which they belong to, and at the same time be given the means to open up in other communities.

The roadmap for inclusive education was set forth in 1994 at the World Conference on Special Needs Education in Salamanca [1]. Inclusion is concerned with learning, participation, and equal opportunities for all children, youth, and adults with a specific focus on the groups that are vulnerable to marginalization and exclusion from society life. It could apply to any or all of the following categories:

- girls and boys who have gender issues;
- ethnic and faith minority groups, travelers;
- asylum seekers, and refugees;
- children who need support in learning the language of instruction (second language);
- children with special educational needs, including those considered to have emotional, behavioral, sensory, physical, or mental disabilities;
- gifted and talented pupils;
- children with social difficulties, such as street children, prison inmates;
- people in disadvantaged, remote areas, poorly served by educational services;
- people who missed the opportunity to study in childhood;
- children in need, including those in public care, orphan children;
- other children, such as the ones with specific health needs, young careers, the children whose families are under stress, pregnant school girls, and teenage mothers;
- any pupils at risk of disaffection and exclusion.

These groups are usually excluded from the mainstream education. Therefore, education for them requires special approaches and techniques.

This article is organized as follows: Section 2 presents an overview of the definition of the emotional states, section 3 presents approaches for detection emotional states, section 4 presents the Acacia project and especially the Apoya module dealing with methods for detection emotional states, section 5 presents the architecture proposed for Apoya module while section 6 shows a working scenario within the centers for educational and professional development (CADEP) which will be setup in the Acacia project. Finally, section 6 summarizes the principal conclusions and future work.

2. EMOTIONAL STATES

As humans, our ability to perform every task in our daily life, from work to amusement or any other conscientious task depends on our emotions. Emotions play a central role as they ensure our survival and all activities from the most basic to the most elaborated tasks. Damasio proposed that the relationship between learning, emotion and body state runs much deeper than many educators realize and that the original purpose for which our brains evolved was to manage our physiology, to optimize our survival, and to allow us to flourish [2]. Emotions are as vast as the diversity of persons and their relationship with the environment including objects people and other living beings. Two aspects are important to be taken into consideration: first aspect is connected to the situation when exposed to external stimuli a person has a physiological flow of activity and that will trigger thoughts emotions and acts in response, and second, how a person reacts to threats or to friendly situations in different ways according to his feelings or thoughts towards those persons or environments[3].

There are several concepts such as emotion, feeling and mood that, sometimes are confused and used interchangeably without regarding to their differences in meaning. According to [4] emotions are cognitive data arising from events (internal and external) used to inform responses, and attributed to concepts and states while feelings are subjective experience of an emotion or set of emotions and mood is an overall state of emotion, which is sustained over longer periods of time and is less changeable than emotions themselves.

Emotions are important in nonverbal communication, and emotions influence cognition in many ways; how we process information, our attention, and our biases towards information [5].

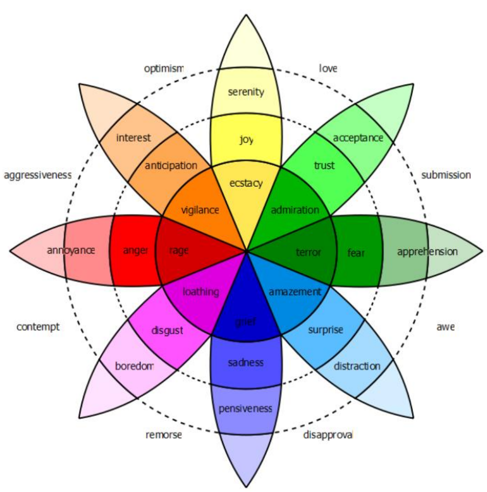

Figure 1 Plutchik's wheel model of emotions (after [6])

The Plutchik's wheel Figure 1, represents the families of emotions, which can be used as a reference list when one looks to find means of detection.

The Circumplex Model of Affect proposes that all affective states arise from two fundamental neurophysiological systems, one related to valence (a pleasure–displeasure continuum) and the other to arousal, or alertness. Each emotion can be understood as a linear combination of these two dimensions, or as varying degrees of both valence and arousal as depicted in Figure 2 [7].

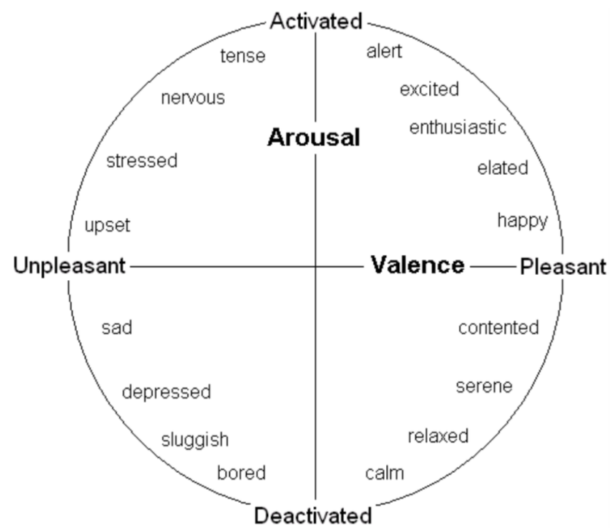

Figure 2 Arousal and Valence in Circumplex Model of Affect

3. EMOTION DETECTION APPROACHES

Current research in affective computing combined with learning assessment may control the affective dimension of the learner. In other words it may detect motivation [8] or learning effectiveness [9]. In this respect, the first issue is affective detection, and some works have been done, mainly centered in using just one data source aimed to get enough information to detect the affective state of learner. Two aspects are of interest and have been analyzed: i) emotions detection, and ii) data labeling. Facial gestures detection can also provide useful information about the users' emotions as in a research conducted by one of the psychologists of the aDeNu research group and project partners from the Universidad de Valencia. A review of the state of the art regarding emotions detection with facial gestures has been published in [10].

Psycho physiological sensors. One of the most common approaches on emotion detection relies on the use of Psycho physiological sensors, which consider different measures, including electrocardiography, skin conductance response or skin temperature. In [11] authors use a commercial pulsimeter and alternatively skin conductance to detect stress using continuous wavelet transforms and decision tree algorithms (J48). It has been noticed that the skin conductance presented broad differences in the relaxed and stressed stages. To perform the experiment, the subjects were asked to listen to annoying sounds and solve some Stroop tests (i.e., participants are asked to tell aloud the names of colors appeared in a different ink than the named color) [12].

Heart rates have been used in [13], but this time combined with speech to assess people's emotions while watching 30 pictures from the International Affective Picture System (IAPS) [14]. The evaluation was achieved with the Self-Assessment Manikin scale - SAM (i.e., a non-verbal pictorial assessment technique that directly measures several emotional dimensions) [15]. In this case some personality traits were also used. The study presented in

[16] shows the relationship between hyperventilation and affective states, measuring inspiratory and expiratory time, tidal volume, and pulse rate to study their connection with valence (pleasantness) and arousal (activation) measured using also the SAM scale. For elicit emotions in this experiment eight imagery scripts were used: three relaxation (positive valence, low arousal), two fear (negative valence, high arousal), one depressive (negative valence, low arousal), one action (neutral valence, high arousal) and one "desire" script (positive valence, high arousal). In [17] authors used Blood Volume Pulse, Galvanic Skin Response, Pupil Diameter and Skin Temperature to detect stress from 32 subjects while solving a series of Stroop tests [12]. To recognize stress patterns in the captured signals, three different learning techniques have been applied: decision tree learners, Naïve Bayes classifiers and Support Vector Machines (SVM). The last technique had the best prediction results (90.10%), followed by decision tree learner (88.02%). The lowest prediction rate was achieved with the Naïve Bayes (78.65%). While dealing with e-learning materials, [18] links e-learning material provided to the learner with his galvanic skin response as only data source. The results suggest that SVM (97.06%) and discriminant analysis (94.12%) shows higher accuracy levels than the k-nearest neighbour technique (79.42%). In [19] authors use both facial electromyography and electro dermal activity as a man-machine interface for empathic consumer products. Trying to detect four emotions (neutral, positive, negative and mixed), they first applied analysis of variance and principal component analysis to perform the selection of a subset of extracted features. After that, they performed classification using k-nearest neighbours (k-NN), SVM and artificial neural networks (ANN). The average prediction rates obtained (60.71% for SVM, 61.31% for k-NN and 56.19% for ANN) questioned the success of the predictors. Conclusions pointed out the need of the development of a generic, self-calibrating bio-signal-driven classification framework. In [20] electro-dermal activity was used to detect 12 proposed emotions during an 8 session experiment with 3 subjects while using ALEKS, a web-based ITS for math, statistics, science, and other domains. In [21] a framework is proposed to recognize learner's emotions. Using electro-encephalography, skin conductance and blood volume pressure during an experiment that included three different environments (trigonometry, backward digit span, and logic) where the participants were asked to rate their experienced levels of stress, confusion, frustration, and boredom. In [22] participants suffering from cerebral palsy were asked to evaluate some affective sounds from the International Affective Digitized Sounds database [23]. The experiment tried to predict their affective state evaluated as valence and arousal using SVM with galvanic skin response data and electro-encephalography data. The accuracy was of 51%. In [24] electrocardiography is used together with electroencephalography in order to detect emotions elicited by means of a set of 60 emotional images from the IAPS database [14]. [24], [25] present studies comparing different works using autonomic nervous system responses as data input for emotion detection, classifying more than 60 studies according to the approach followed and the classification techniques used.

There are many examples on the progress of affect computing in educational settings as well. For instance, a framework has been proposed to recognize learner's emotions using electroencephalography, skin conductance and blood volume pressure in [21]. In [18] authors used e-learning material provided to learners in relation to their galvanic skin response. Electro-dermal activity was used to detect the 12 proposed emotions in [20]. There are also instances on using non-verbal communication, such as body movements and facial expressions, which were used for evaluating learners' states [26].

The aDeNu Research Group at UNED has designed, implemented and evaluated the Ambient Intelligence Context-aware Affective Recommender Platform (AICARP) infrastructure to explore the potential of context-aware affective feedback beyond computer-based recommendation approaches taking advantage of the possibilities of ambient intelligence [27]. The corresponding personalized support is provided without interrupting the learning activity by delivering the recommended action to the learner meanwhile the learning activity is achieved (e.g. while the learner is talking, the system can tell her to slow down by switching on a light or playing a sound). Such feature requires to enrich the system with capabilities for detecting changes in the learners' affective state (e.g. from physiological sensors), as well as to interact with the user through the preferred sensorial channel (e.g. sight, hearing, touch, smell) [28]. [29] considers six main perspectives for modeling affect: emotions as expressions, emotions as embodiments, cognitive approaches to emotions, emotions as social constructs, affective neuroscience, core affect and psychological construct of emotion. All those trends are important for analyzing and characterizing emotional states, and currently influence AC.

AC systems can create different scenarios that help and improve educational conditions. A system for emotions identification may detect signals of frustration during the learning process, or lack of understanding during the study of concepts and definitions. Applications include tracking of emotional trends in groups, detection of emotional interactions, and detection of anxiety or depression patterns. With such identification at the beginning of processes, the educational staff can start individual psychological assistance for students, avoiding future problems that interfere in the learning process, and even more, in their lives.

Some tools are currently available for detecting emotions, such as Affectiva. Affectiva is based on machine learning techniques and provides a non-invasive way to detect emotions, since it is based purely on the images captured via webcam. It considers 42 action units of the face and, from them, 6 major emotional expressions are mapped and analysed. The database of videos (used to train the Affectiva machine learning tool), documentation, software Affdex SDK are freeware, and UFOPA considers Affectiva tools as starting point for investigations related to future AC developments in education in the context of the ACACIA project.

4. ACACIA PROJECT

ACACIA project is grounded on the principle of education for all. Although there have been many advances registered, higher education (HE) in Latin and Central America still faces several problems [30], including: student desertion caused by emotional factors, academic, economic or social marginalization; and the communication gaps between involved members that inhibit the management of collective actions to face transversal problems referred to the access and successful permanence in the university.

The main goal of ACACIA project is to contribute to the dissipation of that exclusion, discrimination and marginalization by disparity or inequality. In this context, the goals of ACACIA are:

- Recognize HE Institutions as social and political means to develop inter, multicultural and multi-linguals education programs matching the several real educational needs;

- Strengthen teaching staff qualification and capacitation;
- Use ICT as a tool to complement teaching processes and learning;
- Propose new forms of institutional organization to promote the integration of groups that combine efforts and resources to the solution of problems previously mentioned.

4.1 CADEP

In terms of organization, as a result from the analysis of multiple theoretical approaches that address the problem of student retention, the project setup supporting centers for educational and professional development, namely CADEP. The CADEP has an integrated systems of modules (Empodera, Innova, Cultiva, Apoya, Convoca) which act together in order to: (i) monitor students at risk; (ii) training and support the academic, technical and administrative staff; (iii) use new strategies for university teaching and innovative use of ITC in practical scenarios by stimulating the entrepreneurship between students and professors.

4.2 Apoya

This module has two components. First one is technological and the second one is human. The goal of the technological component is to implement an automated emotion detection system that allows to monitors and supports students by providing automatic recommendations.

The human component seeks to educate, inform and divulge throughout the academic community, guidelines for the recognition and handling of persons in situations of exclusion (attention disorder, including disability, cultural differences, extreme emotional situations, etc.).

This module directly relates to all other development modules, providing valuable information to PT2 (Empodera), PT3 (Innova), PT4 (Cultiva), and PT6 (Convoca). In the final design the module will be integrated in CADEP, and will communicate with PT8 (Disemina). PT8 Disemina provides guidelines for the detection and execution of situations that may generate social exclusion attending the broadcasting actions to raise awareness, for apprehension and for the action.

APOYA tasks descriptions:

T.1 Automatic detection of affective states. The work focuses on the implementation of a system for detecting emotions, also defining a methodology for data collection and methodology for emotional labeling and sys-tem management.

T.2 Tracking and recommendations automation. This work deals with the automation of monitoring the student and the system that generates recommendations to help improve their academic level.

T.3 Promotion of multiculturalism and diversity. The objective of this task is to generate guidelines for the detection and recognition of people at risk of social exclusion, a guide for action (treatment, activities, visibility) and a system of courses to train university and technical personnel (teachers, administrative staff and technical supporters included).

T.4 Integration strategies. The objective of this task is to define the protocol operation and the internal articulation of Apoya module, define the operation of the laboratory module, define the diffusion and disclosure system and the system of internal evaluation of the module itself.

5. ARCHITECTURE

Physical data comes from the assessed subject and/or ambience, as the environment contains various sensors that capture data. Captured data are stored in a warehouse. The combination and fusion of multiple sensor output can be used. A rule-based expert system transforms this data into relevant contextual data. The problem of sensor fusion is particularly important in the extraction of contextual data: a sensor might not produce sufficient information due to uncertainty and unreliability of the sensor itself. Two types of sensor fusion may be applied: the competitive and the complementary. The competitive sensor fusion is based on sensors which detect equivalent physical data, trying to diminish the errors in the measurements of every sensor. The complementary sensor fusion uses different typologies of sensors to extract high level data. Various sensors and actuators connect the environment to the middleware level. Data and knowledge are transferred to applications and services that are accessed by end users on mobile devices or client-server-type information systems.

The main goal of this environment is to create an enhanced learning and teaching environment. Our process is designed to identify the necessary improvements that will have an immediate positive impact, manage efficient implementation of those improvements, and regularly communicate progress and results to all stakeholders.

5.1 Apoya Recommendations

The fundamental issue in a ubiquitous learning environment is how to provide learners with the right material at the right time in the right way. Context aware adaptation is therefore indispensable to all kinds of learning activities in ubiquitous learning environments.

The ubiquitous environment should be personalized according to the learner's situation. Personalization is defined in [31] as the way in which information and services can be tailored in a specific way to match the unique and specific needs of an individual user. While a learner is doing learning task or activity, it usually looks for some knowledge. In a ubiquitous learning environment, it is very difficult for a learner to know who has this knowledge even though they are at the same place. In this case, the learner needs to be aware of the other learners' interests that match his request [31].

There are only a few studies that have attempted to induce the educational affordances of context - aware ubiquitous learning environment. [32] devoted attention to the problem of what educational affordances can be provided by a context - aware ubiquitous learning environment. They proposed a system named EULER that can provide eight educational affordances:

1. knowledge construction,
2. apply,
3. synthesis,
4. evaluation,
5. interactivity,
6. collaborative learning,
7. game based learning,
8. context aware learning.

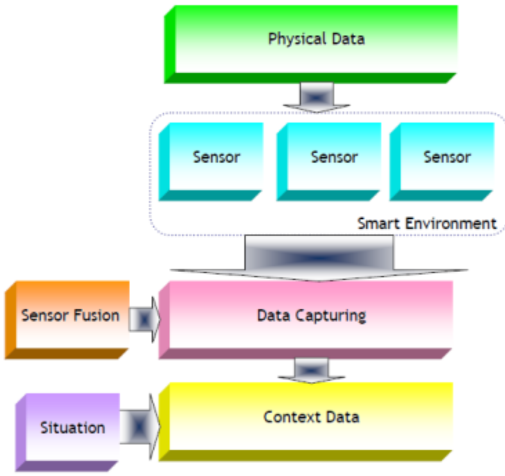

Figure 3 Emotional Gathering Architecture

Moreover, they stressed that ubiquitous learning provides context aware information and self-learning opportunities for learners. Therefore, it not only enables students to achieve learning goals anytime and anywhere, it is also cultivating their ability to explore new knowledge and solve problems. This should be considered to be one of most important characteristics of ubiquitous learning.

The process of ubiquitous learning is understood as a social process that happens at a time and place of the learner's choosing, continuing throughout one's life. It is collaborative, evolving and informed by a process of self-paced development.

6. SCENARIO

Student dropout is a problem affecting Higher Education Institutions (HEI) in Latin America and Caribbean. The dropout problematic can be originated by several factors such as emotional factors, academic, economic or social cultural marginalization.

A possible technological solution to avoid student's dropout and increase their performance is to develop a frameworks using the internet of things (IoT) paradigm integrating several devices such as biomedical sensors and eye trackers to collect information from the students. The goal is to use that information collected from different sources to support professors to identify and manage students' emotional state during classroom lessons.

The proposed scenario consists of the evaluation of the student's emotional state, for affective management and prevention of HEI dropouts (Figure 4).

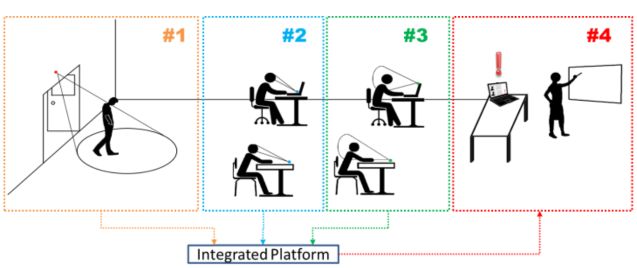

Figure 4 Scenario for affective management in student dropout prevention including four case studies: 1) gait and posture analysis; 2) eye tracking detection; 3) automatic facial emotion detection; 4) emotional record based on data integration.

In that context four case studies were identified as illustrated in Figure 4: gait and posture analysis; eye tracking detection; automatic facial emotion detection; emotional record based on data integration;

The gait and posture analysis case study is performed when the student enters or leaves the classroom. To implement this case study a 3D motion capture device will be used, composed by a RGB camera and an infrared depth sensor (i.e. a Kinect device) in order to track the student skeleton key points. It is expected that a change in the student regular gait also indicates a change in his/hers emotional state.

The eye tracking detection case study is performed when the students are sitting at a desk. The eye tracker will be used to track both head and eye movements. This analysis allows both to recognize the student affect states and to measure the engagement level of the student during learning tasks.

The automatic facial emotion detection case study is performed when the student is sitting at a desk. A RGB camera will be used for this case study to record facial expressions. An algorithm will extract and analyze the facial features in order to determine the correspondent emotion or emotions.

Both in the eye tracking detection and in the automatic facial emotion detection case studies there is the possibility of not using a computer, just the emotion tracking device.

In the emotional map based on data integration case study, data from the previous case studies will be collected and processed in real time in order to create an emotional map to detect potential problematic situations, such as disengagement, attention disorder, learning difficulties, emotional stress, and dropouts. This case study is performed in the classroom environment to support teacher manage the class. Technically, this integrative platform provides a real-time early alert system activated when deviations from regular patterns are detected. It is expected that the system, at the end of each class generates a report for the interested actors, i.e., teacher, students, parents, and the school.

For this scenario 3 different capturing devices will be used, as mentioned above.

The Kinect is comprised by an RGB camera and a infrared depth sensor, and its SDK allows to track a person's skeleton key points.

The eye tracker is a device usually composed by an infrared light source, an infrared or near-infrared camera and/or an RGB camera. It allows the tracking of head and eye movements as well as pupil variations, which has been proven to be an indicator of cognitive activity.

In the automatic facial emotion detection case study a simple RGB camera is used, and a computer algorithm first detects and isolates the face, then the facial features are extracted and the facial expression classified.

7. CONCLUSIONS AND FUTURE WORK

The goal of ACACIA project is to contribute to student dropout prevention in Latin America, caused by a variety of factors including emotional factors, academic, economic or social marginalization. The organization of the project promotes the implementation of centers for educational and professional development, namely CADEP. That are supported by a set of integrated modules, in which is included APOYA module which aims to use innovation as driving factor to avoid student dropout.

This paper addresses how affective computing may prevent student's dropout. Specifically, this paper presents a scenario in the context of a smart classroom or environment, in which an innovative technology will be used to detect and manage students emotional state based on Kinect devices, eyetracker and webcam data acquisition. The scenario is composed by four case studies: gait and posture analysis, eye tracking detection, automatic facial emotion detection and emotional record based on devices data integration.

Future work will be the technical design and implementation of the referred framework and the test of the described scenario in the CADEPs. Additionally, other innovative solutions and a methodology for prototype evaluation will be proposed and developed.

8. ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Union ERASMUS+ Program under grant agreement EAC/A04/2014 ACACIA.

9. REFERENCES

[1] UNESCO, "The Salamanca Statement and Framework For Action on Special Needs Education." Adopted by the World Conference on Special Needs Education: Access and Quality, Salamanca, Spain, 7-10 June 1994, 1994.

[2] M. H. Immordino-Yang and A. Damasio, "We Feel, Therefore We Learn: The Relevance of Affective and Social Neuroscience to Education," *Mind, Brain, Educ.*, vol. 1, no. 1, pp. 3–10, Mar. 2007.

[3] F. L. L. Ferreira, "Knowledge management framework based on brain models and human physiology," 2015.

[4] J. Lin, M. Spraragen, and M. Zyda, "Computational models of emotion and cognition," *Adv. Cogn. Syst.*, vol. 2, pp. 59–76, 2012.

[5] J. Broekens, D. DeGroot, and W. A. Kosters, "Formal models of appraisal: Theory, specification, and computational model.," in *Cognitive Systems Research*, 2008, pp. 173–197.

[6] R. Plutchik, *The Emotions*. University Press of America, 1991.

[7] J. A. Russell, "A circumplex model of affect.," *J. Pers. Soc. Psychol.*, vol. 39, no. 6, pp. 1161–1178, 1980.

[8] M. Ainley, "Connecting with learning: Motivation, affect and cognition in interest processes," *Educ. Psychol. Rev.*, vol. 18, no. 4, pp. 391–405, 2006.

[9] D. J. Lynch, "Motivational factors, learning strategies and resource management as predictors of course grades," *Coll. Stud. J.*, vol. 40, pp. 423–428, 2006.

[10] M. Saneiro, O. C. Santos, S. Salmeron-Majadas, and J. G. Boticario, "Towards emotion detection in educational scenarios from facial expressions and body movements through multimodal approaches," *Sci. World J.*, vol. 2014, 2014.

[11] M. V. Villarejo, B. G. Zapirain, and A. M. Zorrilla, "Algorithms based on CWT and classifiers to control cardiac alterations and stress using an ECG and a SCR.," *Sensors (Basel).*, vol. 13, no. 5, pp. 6141–6170, 2013.

[12] J. R. Stroop, "Studies of interference in serial verbal reactions.," *J. Exp. Psychol.*, vol. 18, no. 6, pp. 643–662, 1935.

[13] E. L. van den Broek, "Ubiquitous emotion-aware computing," *Pers. Ubiquitous Comput.*, vol. 17, no. 1, pp. 53–67, Oct. 2011.

[14] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, "International Affective Picture System (IAPS): Technical Manual and Affective Ratings," *NIMH Cent. Study Emot. Atten.*, pp. 39–58, 1997.

[15] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *J. Behav. Ther. Exp. Psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[16] I. Van Diest, W. Winters, S. Devriese, E. Vercamst, J. N. Han, K. P. Van de Woestijne, and O. Van den Bergh, "Hyperventilation beyond fight/flight: respiratory responses during emotional imagery.," *Psychophysiology*, vol. 38, no. 6, pp. 961–968, 2001.

[17] J. Zhai and A. Barreto, "Stress Recognition Using Non-invasive Technology," *Proc. 19th Int. Florida Artif. Intell. Res. Soc. Conf. (FLAIRS).*, pp. 395–400, 2006.

[18] S. Handri, K. Yajima, S. Nomura, N. Ogawa, Y. Kurosawa, and Y. Fukumura, "Evaluation of student's physiological response towards e-learning courses material by using GSR sensor," in *Proceedings - 9th IEEE/ACIS International Conference on Computer and Information Science, ICIS 2010*, 2010, pp. 805–810.

[19] E. L. Van Den Broek, V. Lisý, J. H. Janssen, J. H. D. M. Westerink, M. H. Schut, and K. Tuinenbreijer, "Affective man-machine interface: Unveiling human emotions through biosignals," *Commun. Comput. Inf. Sci.*, vol. 52, pp. 21–47, 2010.

[20] S. K. D'Mello, "Emotional Rollercoasters: Day Differences in Affect Incidence during Learning.," in *FLAIRS Conference*, 2014.

[21] I. Jraidi, M. Chaouachi, and C. Frasson, "A hierarchical probabilistic framework for recognizing learners' interaction experience trends and emotions," *Adv. Human-Computer Interact.*, vol. 2014, 2014.

[22] Y. Mohamad, D. T. Hettich, E. Bolinger, N. Birbaumer, W. Rosenstiel, M. Bogdan, and T. Matuz, "Detection and utilization of emotional state for disabled users," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2014, vol. 8547 LNCS, no. PART 1, pp. 248–255.

[23] M. M. Bradley and P. J. Lang, "The International Affective Digitized Sounds Affective Ratings of Sounds and Instruction Manual," *Tech. Rep. B-3. Univ. Florida, Gainesville, Fl.*, pp. 29–46, 2007.

[24] H. Hamdi, P. Richard, A. Suteau, and P. Allain, "Emotion assessment for affective computing based on physiological responses," in *IEEE International Conference on Fuzzy Systems*, 2012.

[25] M. Munih and D. Novak, "Interacting with Computers A

survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing q," *Interact. Comput.*, vol. 24, no. 3, pp. 154–172, 2012.

[26] S. Afzal and P. Robinson, "A study of affect in intelligent turoring," in *Supplementary Proceedings of the 13th International Conference of Arti fi cial Intelligence in Education (AIED 2007)*, 2007.

[27] O. C. Santos, M. Saneiro, J. G. Boticario, and M. C. Rodriguez-Sanchez, "Toward interactive context-aware affective educational recommendations in computer-assisted language learning," *New Rev. Hypermedia Multimed.*, vol. 22, no. 1–2, pp. 27–57, Jan. 2016.

[28] E. Gutiérrez y Restrepo, F. Ferreira, J. G. Boticario, E. Marcelino-Jesus, J. Sarraipa, and R. Jardim-Goncalves, "Enhanced Affective Factors Management for HEI Students Dropout Prevention," in *Learning and Collaboration Technologies: Third International Conference, LCT 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings*, P. Zaphiris and A. Ioannou, Eds. Cham: Springer International Publishing, 2016, pp. 675–684.

[29] R. A. Calvo and S. D'Mello, "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications," *IEEE Trans. Affect. Comput.*, vol. 1, no. 1, pp. 18–37, Jan. 2010.

[30] UNESCO, "The Education for All Global Monitoring Report - Teaching And Learning: Achieving quality for all, facilitated and supported by UNESCO (ED - 2014/GMR/GS/1)," *EFA Global Monitoring Report 201*, 2014. .

[31] M. M. El-Bishouty, H. Ogata, S. Rahman, and Y. Yano, "Social knowledge awareness map for computer supported ubiquitous learning environment," *Educ. Technol. Soc.*, vol. 13, no. 4, pp. 27–37, 2010.

[32] T.-Y. Liu and Y.-L. Chu, "Educational Affordances of Ubiquitous Learning Environment for Natural Science Learning," in *ICALT: 2009 IEEE International Conference On Advanced Learning Technologies*, 2009, pp. 290–292.