

## D3.1

# LA modules analysis and configuration





## **The colMOOC: Integrating Conversational Agents and Learning Analytics in MOOCs**

### **D3.1 - LA modules analysis and configuration**

<b>Project number:</b>	588438-EPP-1-2017-1-EL-EPPKA2-KA
<b>Grant Agreement No:</b>	2017-2841/001-001 (Amendments 002 & 003)
<b>Project acronym:</b>	colMOOC
<b>Project title:</b>	The colMOOC: Integrating Conversational Agents and Learning Analytics in MOOCs
<b>Programme, Key action, Action type, Topic:</b>	E+ KA2: Cooperation for innovation and the exchange of good practices, Knowledge Alliances
<b>Start date of the project:</b>	01/01/2018
<b>Duration:</b>	39 months (extended: 03/31/2021)
<b>Project web site:</b>	<a href="https://colmooc.eu">https://colmooc.eu</a>

<b>Deliverable type:</b>	Report, Software
<b>Deliverable reference number:</b>	D3.1
<b>Deliverable title:</b>	LA modules analysis and configuration
<b>WP contributing to the deliverable:</b>	WP3
<b>Delivery date:</b>	06/30/2019 (M18)

<b>WP Leader:</b>	AUTH
<b>Responsible organization:</b>	AUTH
<b>Abstract:</b>	This report documents the deployment, configuration, modification of the Learning Analytics modules and tools that will analyze the collected data to produce valuable information. One of the main tasks of the report is the identification of existing modules/tools that could be reused and the selection of the most appropriate to be deployed.
<b>Keywords:</b>	Learning analytics

<b>Dissemination level:</b>	Public
-----------------------------	--------

**Disclaimer:** “This project has been co-funded by the Erasmus+ Programme of the European Commission. This document reflects the views only of the authors, and the Education, Audiovisual and Culture Executive Agency and the European Commission cannot be held responsible for any use which may be made of the information contained therein”

## Authors and Reviewers<sup>1</sup>

---

Thrasyvoulos Tsiatsos, Aristotle University of Thessaloniki (AUTH), Greece, tsiatsos@csd.auth.gr

Stergios Tegos, Aristotle University of Thessaloniki (AUTH), Greece, stegos@csd.auth.gr

Stathis Nikolaidis, Centre for Research and Technology Hellas (CERTH), Greece, stathis.nikolaidis@iti.gr

## REVIEWERS

Stavros Demetriadis, Aristotle University of Thessaloniki (AUTH), Greece, sdemetri@csd.auth.gr

Anastasios Karakostas, Centre for Research and Technology Hellas (CERTH), Greece, akarakos@iti.gr

---

<sup>1</sup> Name Surname, Organization Name (Organization Name Acronym), Country, e-mail address

## Document Change Log

---

Version	Date (mm/dd/yyyy)	Author (s)	Sections Changed
0.1	29/11/2018	See author list	All
0.2	30/05/2019	See author list	All
1.0	30/06/2019	See author list	All
1.5	01/03/2021	See author list	All
2.0	31/03/2021	See author list	3

## Executive Summary

---

Massive Open Online Courses (MOOCs) have introduced a way of transcending formal higher education by realizing technology-enhanced formats of learning and instruction and by granting access to an audience way beyond students enrolled in any Higher Education Institution (HEI). MOOCs have been repeatedly praised for their potential to democratize education and help individuals gain access to quality education, regardless of their geographic location, financial means, schedule, language or background (Siemens, 2013).

Still, the potential for European HEIs to scale up and reach an international audience of diverse backgrounds has not been realized yet. Although MOOCs have been reported as an important educational tool, there are a number of issues and challenges relating to their educational aspect. The high number of dropouts, the low-level participation, and the lack of students' motivation and engagement are a few of the most critical concerns being reported (Floratos, Guasch, & Espasa, 2015).

A major issue often preventing massive-scale learning environments from reaching their transformative potential is that they fail to provide the kind of interactive environment that is conducive to sustained engagement and learning. Crafting compelling interactive activities, providing dynamic support, and assessing students' performance and engagement are often left out of the picture while creating an online course (Hone & El Said, 2016).

Under this prism, this deliverable presents the ongoing design and development efforts of a Learning Analytics (LA) dashboard.

The aim of this deliverable is to describe

- what type of data can and should be measured?
- for who are the metric visualizations intended?
- what is the scope of each metric?

Therefore, this deliverable presents the colMOOC LA Module design adopting the Conceptual Model introduced by Xiong, Li, Kornhaber, Suen, Pursel, & Goins (2015). Also, it presents 35 LA variables which are grouped in the following categories:

- Demographics
- Intrinsic motivation
- Extrinsic motivation
- Social motivation
- Engagement (Behavioral)
- Retention

In addition, this deliverable presents (a) the LA User Interface along with the metrics and the proposed representation of LA using graphs, and (b) use cases for Learners and Instructors.

## Table of Contents

---

Authors and Reviewers .....	4
Document Change Log.....	5
Executive Summary .....	6
Table of Contents .....	7
List of Tables.....	8
List of Figures .....	9
List of Acronyms.....	10
1 Introduction .....	11
1.1 Purpose of this document .....	11
1.2 Document structure .....	11
1.3 Audience.....	11
2 LA Module design.....	12
2.1 Conceptual model.....	12
2.2 LA Variables .....	12
2.3 LA User Interface .....	14
2.3.1 General user interface.....	14
2.3.2 Metrics.....	19
2.3.3 Use Cases.....	37
3 LA Module development.....	45
3.1 Database Model.....	45
3.2 Implementation.....	47
3.3 Evaluating the Consistency and Quality.....	49
3.4 Integration with external platforms .....	50
4 References .....	52

## List of Tables

---

Table 1: Measured Variables and Items .....	13
Table 2: colMOOC LA module metrics .....	20
Table 3. Learner Use Cases - My Performance in the Activity .....	37
Table 4. Learner Use Cases - My Team's Performance in the Activity .....	38
Table 5. Learner Use Cases - My Contribution in the Course .....	39
Table 6. Learner Use Cases - Peer Interaction in the Course .....	40
Table 7. Learner Use Cases - User-Agent Interaction in the Course.....	41
Table 8. Instructor Use Cases - View Learner Analytics .....	42
Table 9. Instructor Use Cases - Advanced Activity Analytics .....	43
Table 10. Instructor Use Cases - Advanced Course Analytics.....	44



## List of Figures

---

<b>Figure 1. Conceptual Model .....</b>	<b>12</b>
Figure 2: ColMOOC LA Module high-level IA diagram.....	15
Figure 3: User login.....	15
Figure 4: Student after login.....	16
Figure 5: Course overview .....	16
Figure 6: Activity Statistics .....	17
Figure 7: Teacher after login .....	17
Figure 8: Teacher selecting a student to ‘view as’ .....	18
Figure 9: Teacher viewing course as a student.....	18
Figure 10 : Schema of the colMOOC MySQL tables for LA module.....	45
Figure 11 : Landing page of the Learning Analytics Dashboard.....	47
Figure 12 : Course Stats page (Teacher view).....	48
Figure 13 : Activity Stats page (Teacher view).....	48
Figure 14 : Activity Stats page (Student view).....	49
Figure 15 : Public Communication API documentation at SwaggerHub (version 1.1) .....	51

## List of Acronyms

---

Acronym	Description
CA	Conversational Agent
HEI	Higher Education Institution
LA	Learning Analytics
MOOCs	Massive Open Online Courses
PC	Project Coordinator
ToC	Table of Contents
UG	User Group
UI	User Interface
WPL	Work Package Leader
OSI	Open Source Initiative
EUPL	European Union Public Licence
RDBMS	Relational Database Management System
SQL	Structured Query Language
PHP	Hypertext Preprocessor

## **1 Introduction**

---

### **1.1 Purpose of this document**

Learning Analytics is often regarded as a digital-from-birth technology/discipline, at the intersection of technical analysis of data and the learning sciences, that provides the opportunity for reclaiming decades of educational research into valuable daily practice for students (Akhtar, Warburton, & Xu, 2017).

Following the state of the art of the learning analytics systems, presented in D.1.2, this report describes the design and deployment of the colMOOC project Learning Analytics module. The module is responsible for analyzing the data collected in the colMOOC chat environment in order to produce valuable information for the students and the teachers participating in project activities.

### **1.2 Document structure**

The present deliverable is split into the following two major chapters:

- LA Module design
- LA Module development

### **1.3 Audience**

This document is public.

## 2 LA Module design

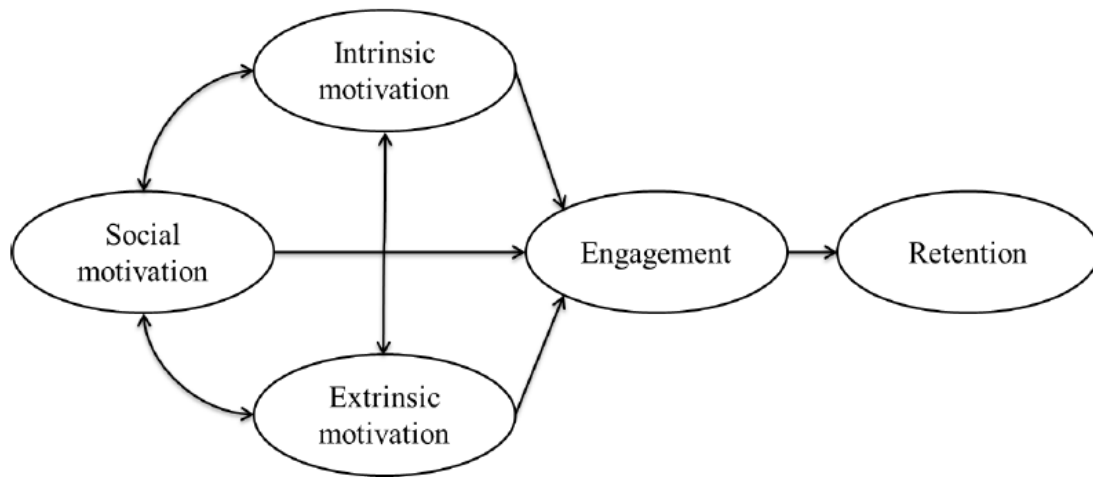
---

### 2.1 Conceptual model

Higher Education Massive Open Online Courses (MOOCs) introduce a way of transcending formal higher education by realizing technology-enhanced formats of learning and instruction and by granting access to an audience way beyond students enrolled in any one Higher Education Institution. The potential for European HEIs to scale up and reach an international audience of diverse backgrounds has not been realized yet. Although MOOCs have been reported as an efficient and important educational tool, there is a number of issues and problems related to the educational aspect. More specifically, there is an important number of drop outs during a course, little participation, and lack of students' motivation and engagement overall. This may be due to one-size-fits-all instructional approaches and very limited commitment to student-student and teacher-student collaboration.

The main GOAL of the LA module is to support the students' motivation and thus engagement and retention.

We decided to adopt the Conceptual Model introduced by Xiong, Li, Kornhaber, Suen, Pursel, & Goins (2015) and depicted in Figure 1.



**Figure 1. Conceptual Model**

### 2.2 LA Variables

This section presents the LA metrics, namely the data to be collected by the LA module in order to measure engagement and retention.

Henrie, Halverson, & Graham. (2015) presented existing approaches to measure engagement in technology-mediated learning. They have identified strengths and limitations of existing measures and outline potential approaches to improve the measurement of student engagement. Based on the existing approaches, and the nature of CA module design we propose a rather short list of metrics and methods presented at Table 1.

The metrics and methods are grouped in the following categories:

- Demographics
- Intrinsic motivation

- Extrinsic motivation
- Social motivation
- Engagement (Behavioral)
- Retention

This list of Variables and Items (Table 1) will operationalize the engagement and retention as presented in our Conceptual depicted in Figure 2.

**Table 1: Measured Variables and Items**

Variables	Item
Demographics	DE1. Gender DE2. Language DE3. Nationality DE4. Time zone
Intrinsic motivation	IM1. Interest: I am participating in this chat activity out of general <b>Interest</b> IM2. Interest: I am participating in this chat activity out of general <b>Curiosity</b> IM3. Interest: I am participating in this chat activity out of general <b>Enjoyment</b>
Extrinsic motivation	EM1. Certificate: I intend to earn a Statement of Accomplishment (or Verified Certificate) for this <b>course</b> . EM2. Credential: I am interested in earning a credential of the specific <b>activity</b> . EM3. Academic: The <b>course</b> relates to my current academic program. EM4. Academic: The <b>activity</b> relates to my current academic program. EM5. Job: The <b>course</b> relates to my current job responsibilities or company's line-of-business. EM6. Job: The <b>activity</b> relates to my current job responsibilities or company's line-of-business.
Social motivation	SM1. Connect: Request to start chat activity SM2. Partnership: Rate your collaboration with your peer
Engagement (Behavioral)	EN1. Posts: Number of chat posts EN2. Posts: Number of words in each post EN3. Agent: Number of answers to agent interventions directed to me EN4. Agent: Number of answers to agent interventions directed to my team EN5. Agent: Number of words in each answer to agent intervention directed to me EN6. Agent: Number of words in each answer to agent intervention directed to my team EN7. Agent: Number of unhandled agent interventions directed to me EN8. Agent: Number of unhandled agent interventions directed to my team EN9. Activity: Number of completed chat activities EN10. Activity: Number of chat activities participated EN11. Activity: Number of unique domain concepts discussed (timestamps)* EN12. Activity: Number of words in the final task answer EN13. Time: Response time to agent interventions directed to me EN14. Time: Response time to agent interventions directed to my team EN15. Time: Timestamp of (each) agent intervention EN16. Time: Timestamp of (each) chat post EN17. Time: Duration of activity EN18. Technical: Number of disconnects

Retention	R1. Retention in all colMOOC activities (Activity 1 – 75%, Activity 2 – 5%, ...) R2. Growth rate of other variables during the course (throughout all activities) ... { 1-13}
-----------	---

The following items

- DE1. Gender
- DE2. Language
- DE3. Nationality
- DE4. Time zone
- IM1. Interest: I am participating in this chat activity out of general Interest
- IM2. Interest: I am participating in this chat activity out of general Curiosity
- IM3. Interest: I am participating in this chat activity out of general Enjoyment
- EM1. Certificate: I intend to earn a Statement of Accomplishment (or Verified Certificate) for this course.
- EM2. Credential: I am interested in earning a credential of the specific activity.
- EM3. Academic: The course relates to my current academic program.
- EM4. Academic: The activity relates to my current academic program.
- EM5. Job: The course relates to my current job responsibilities or company's line-of-business.
- EM6. Job: The activity relates to my current job responsibilities or company's line-of-business.
- SM2. Partnership: Rate your collaboration with your peer

is proposed to be collected by questionnaires, whereas the rest of the items will be collected by monitoring the user's actions in the colMOOC platform.

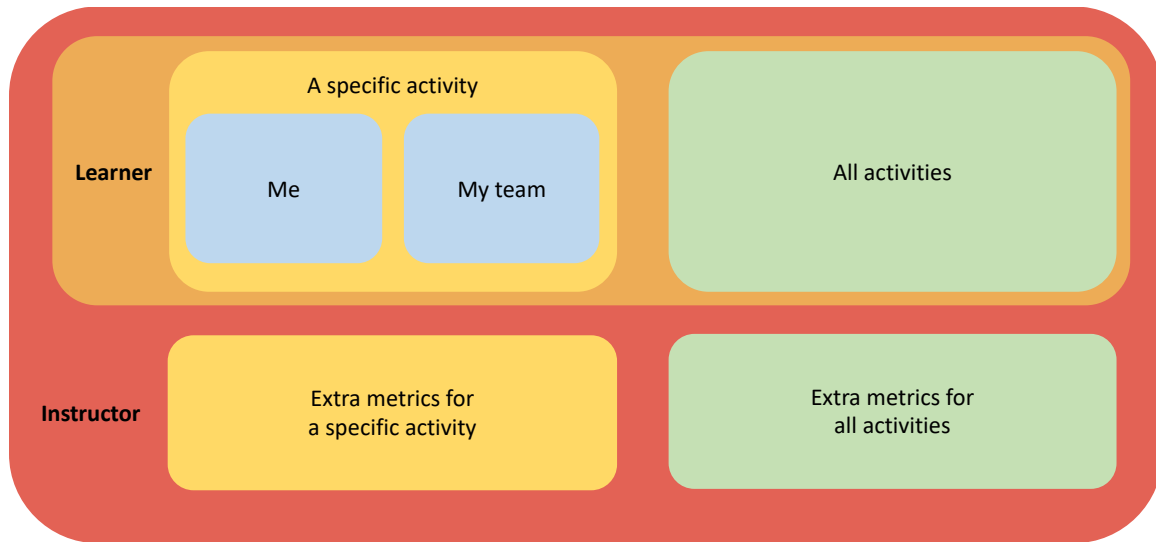
## 2.3 LA User Interface

The LA module will be used by Learners and Instructors visualizing various metrics according to their role and use case. This section will present in detail:

- The general UI
- The metrics as well as their visualization
- The specific use cases for the Learners and Instructors

### 2.3.1 General user interface

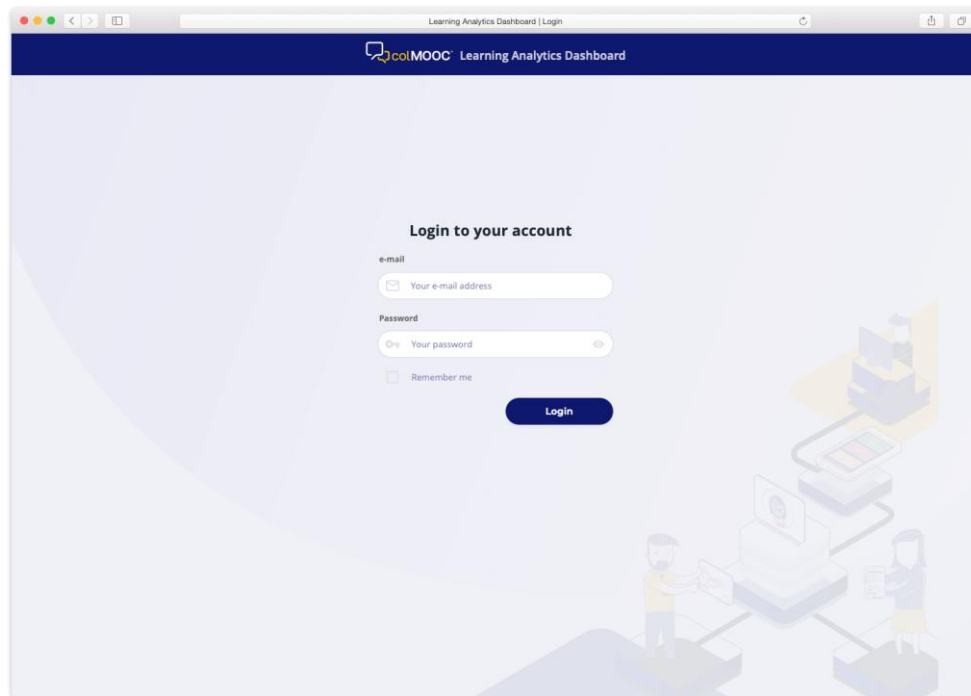
Considering the information architecture issues arising often while designing user interface with lots of information, a high-level information architecture diagram was created (Figure 2). In general, Information Architecture (IA) focuses on organizing, structuring, and labeling content in an effective and sustainable way. The goal is to help users find information and complete tasks. To accomplish this, it is critical to be able to understand how the pieces fit together to create the larger picture, how items relate to each other within the system.



**Figure 2: ColMOOC LA Module high-level IA diagram**

Since an agile software development methodology was followed throughout the project, a series of mockups were crafted by the design team before proceeding with the implementation of the LA dashboard interface. A collaboration tool, named Zeplin, was used during the prototyping phase of the LA user interface. During this phase, the UI designers and the front-end developers worked closely together. This iterative process resulted in crafting a highly modern user interface that followed the guidelines of the project Brand Toolkit Guide.

The final mockups that present the graphical user interface of the Learning Analytics Dashboard are presented below.



**Figure 3: User login**

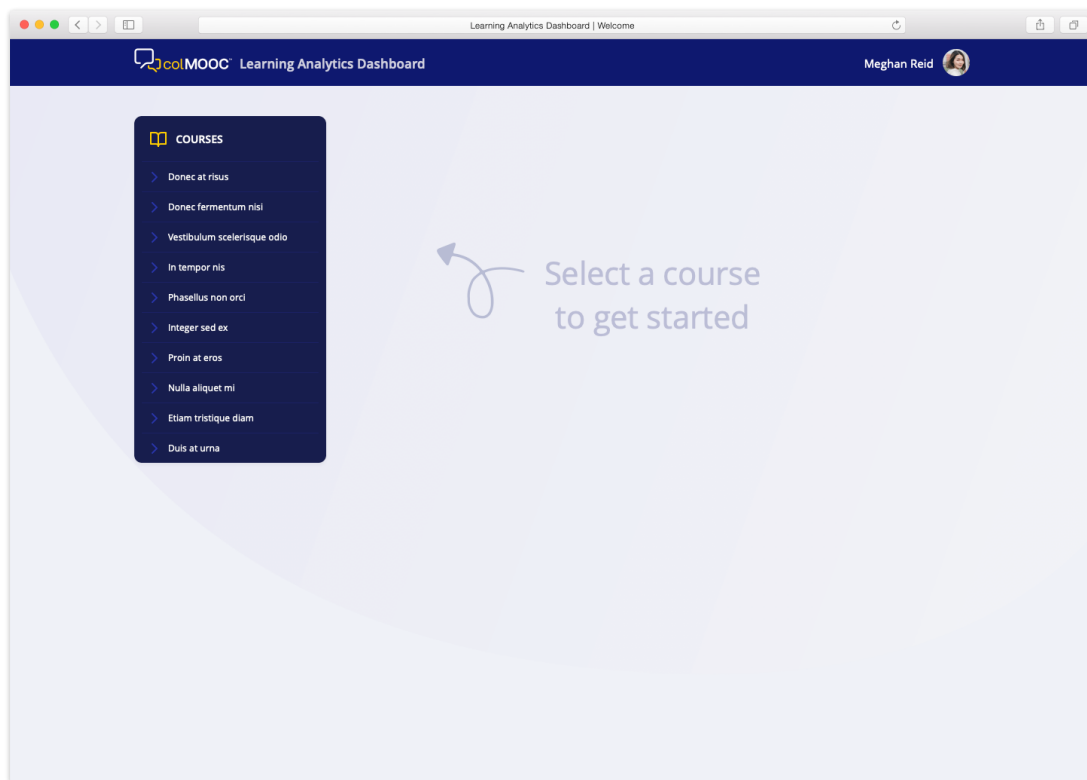


Figure 4: Student after login

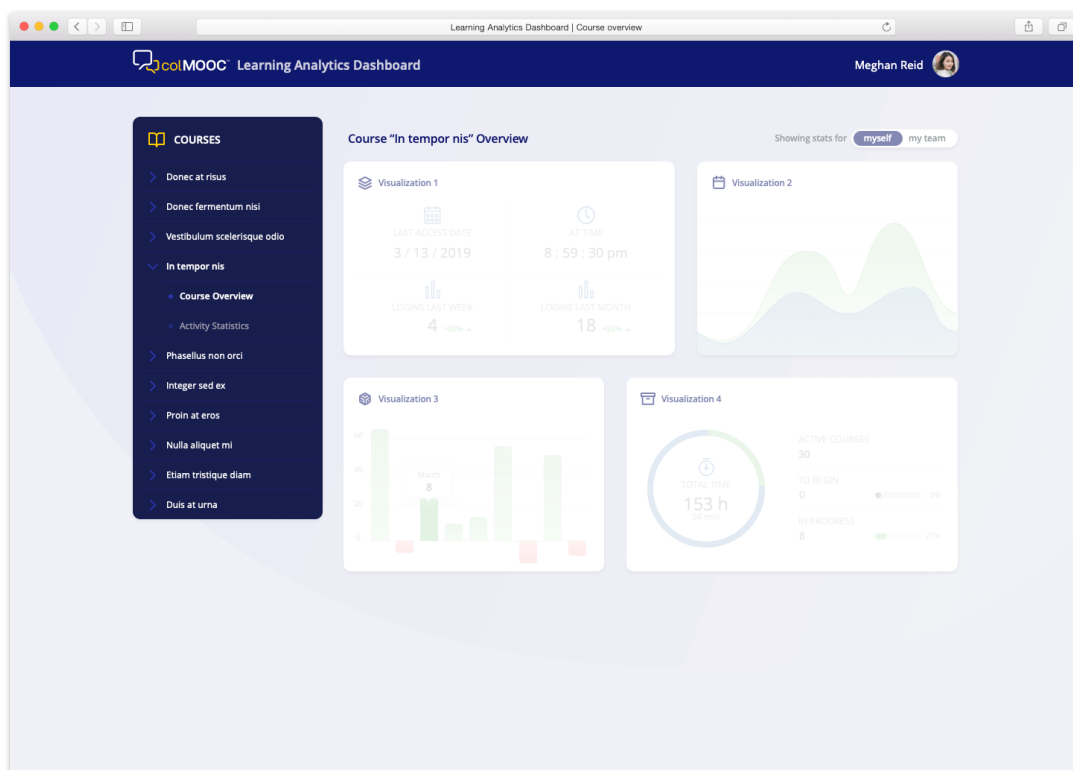


Figure 5: Course overview



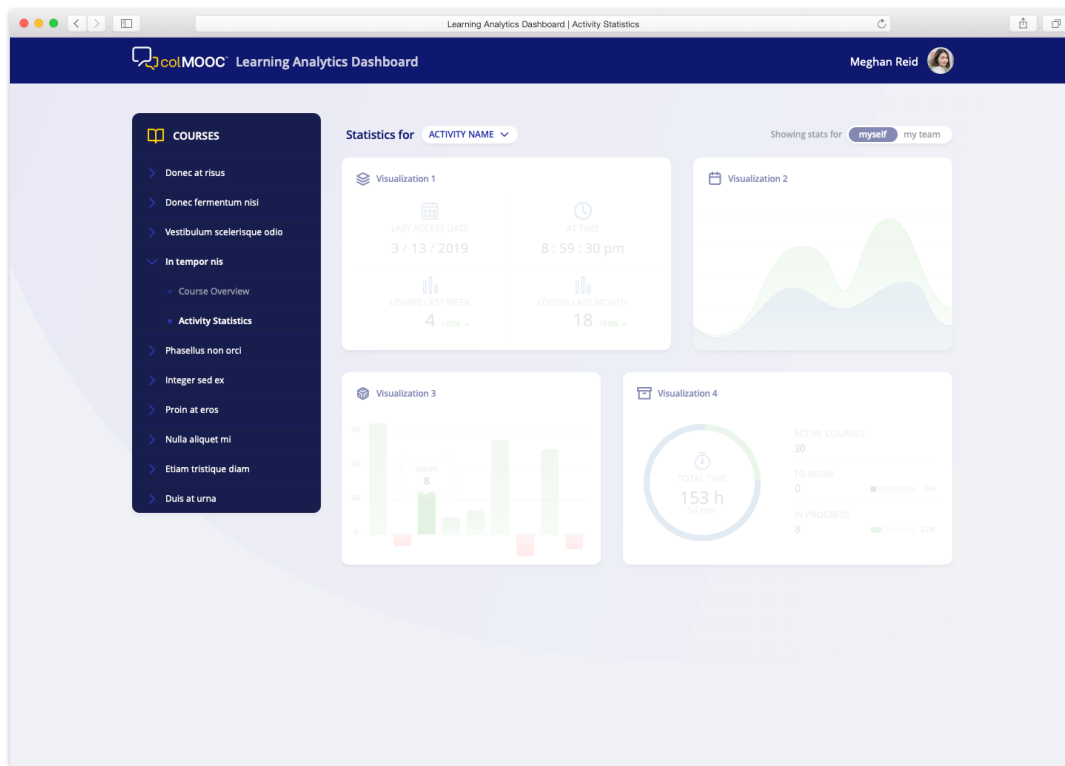


Figure 6: Activity Statistics

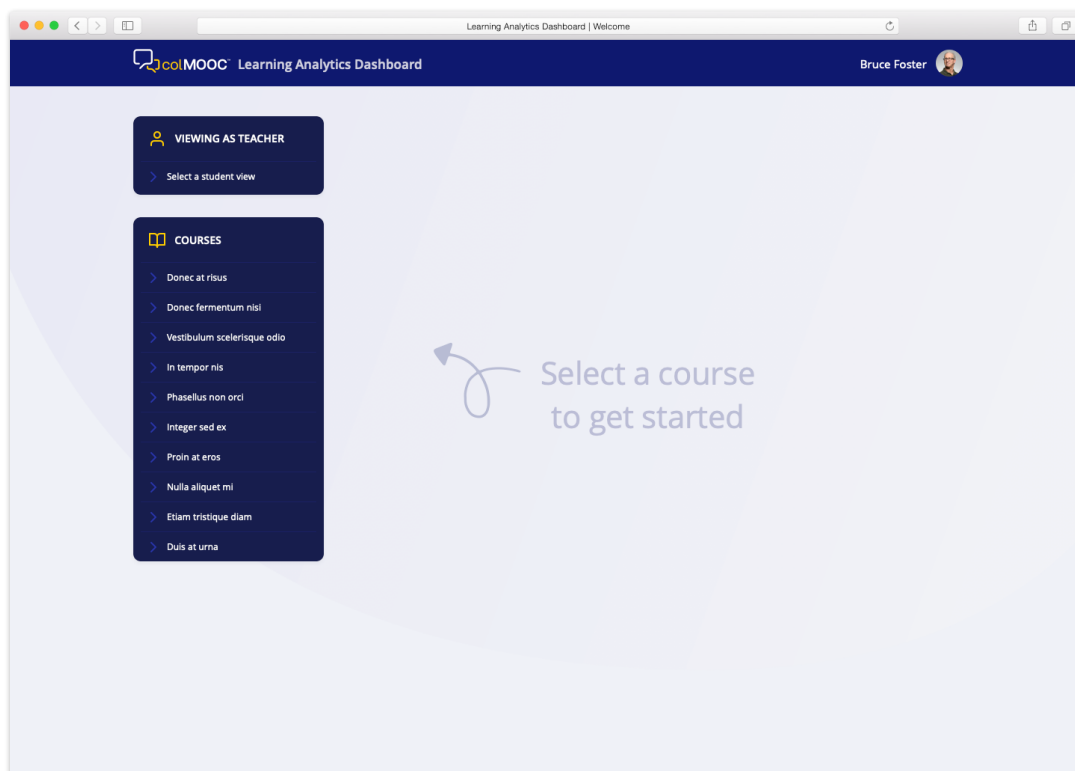


Figure 7: Teacher after login

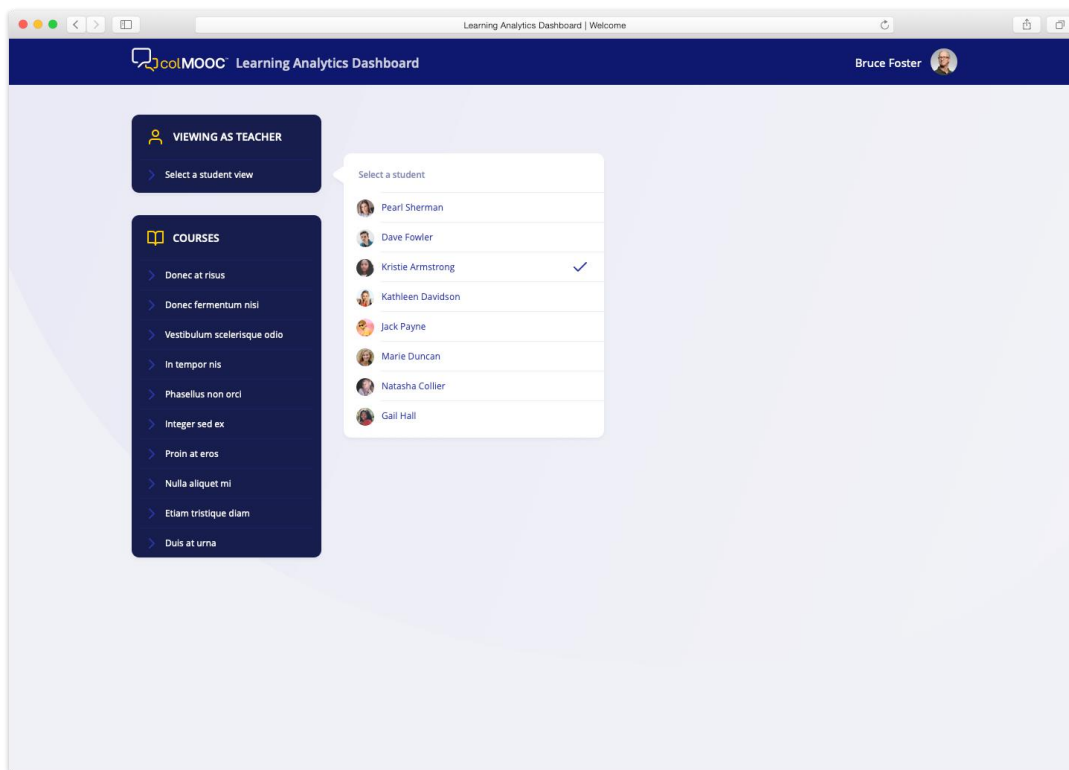


Figure 8: Teacher selecting a student to ‘view as’

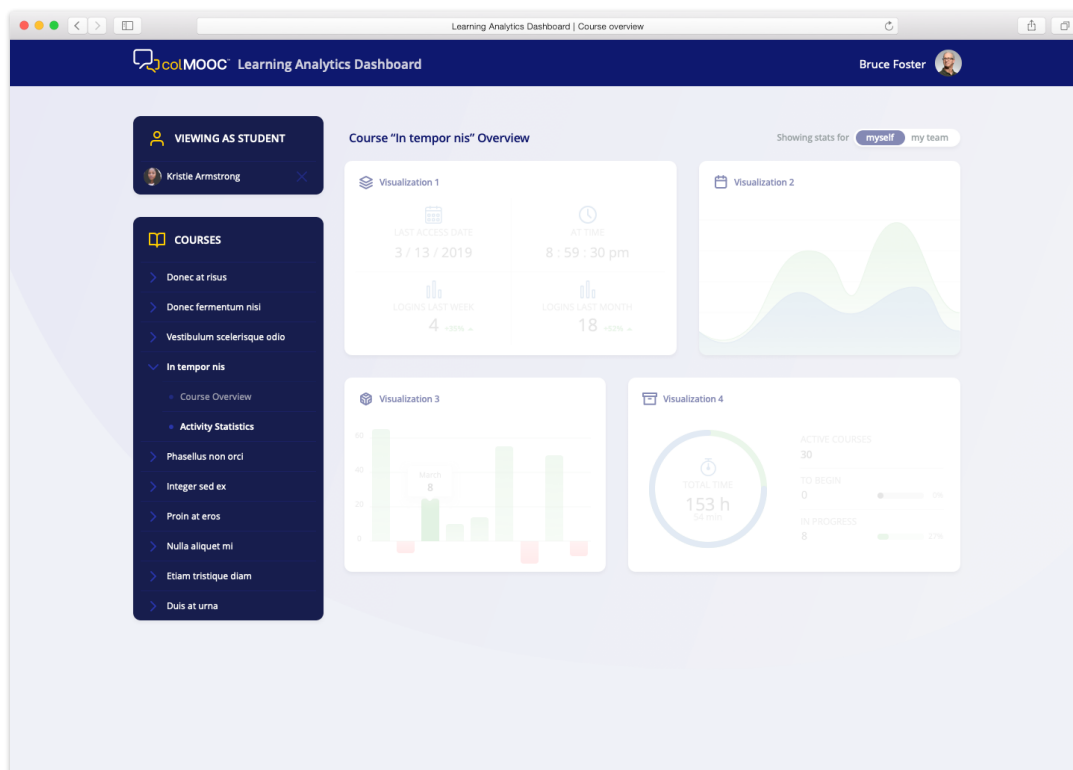


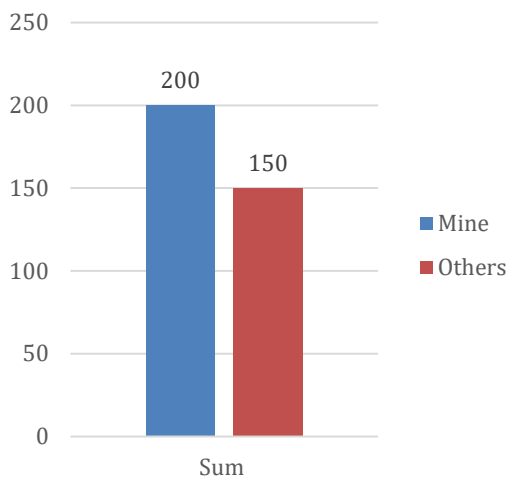
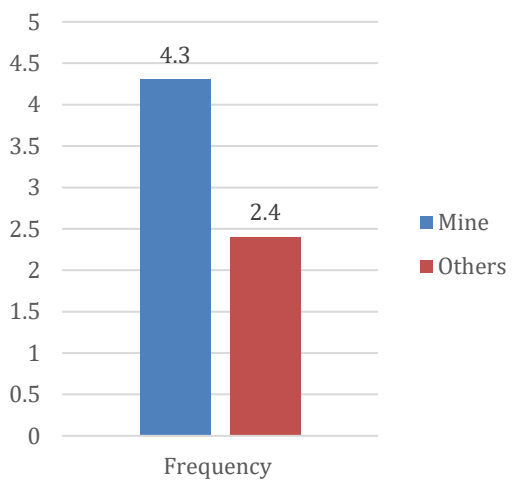
Figure 9: Teacher viewing course as a student

### 2.3.2 Metrics

The following table presents in detail the metrics of the LA module as well as their visualization. More specifically, the table presents the following info in the respective columns:

- **ID:** The metric ID to be used of easy reference to the metric
- **Metric:** The short description of the metric
- **User:** Learners or Instructor. Note that all analytics addressing the Learner can also be accessed by the Teacher, who is essentially the Activity Administrator.
- **Level:** (I)ndividual/(T)eam. Individual means that the metric depicts my performance/activity. Team means that the metric depicts my team's performance/activity.
- **Variable(s):** The ID of the variables presented at Table 1. The variable(s) is (are) the input(s) to calculate the metric.
- **Visualization:** The graph that depicts the metric.

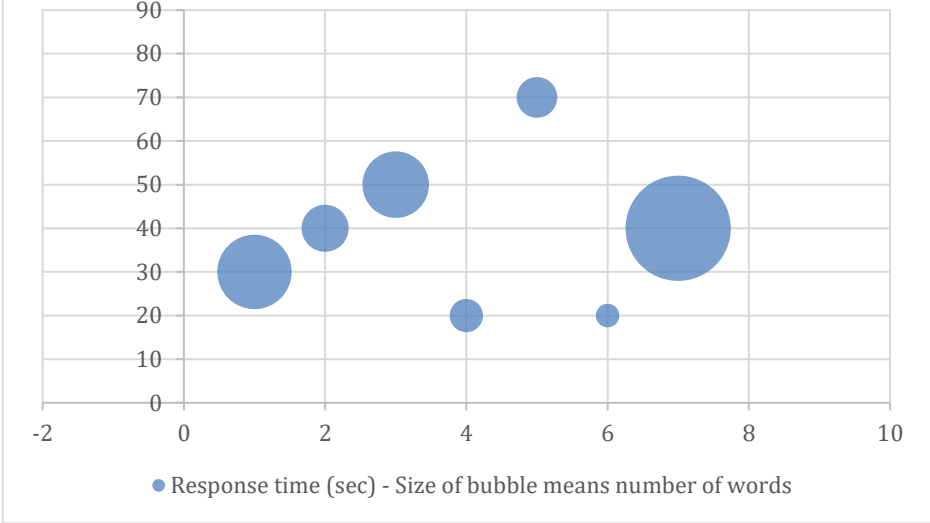
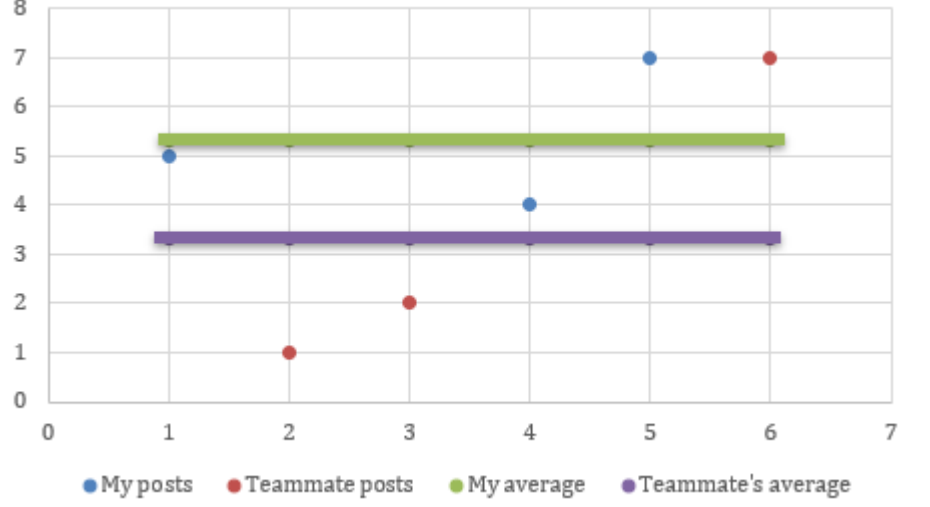
Table 2: colMOOC LA module metrics

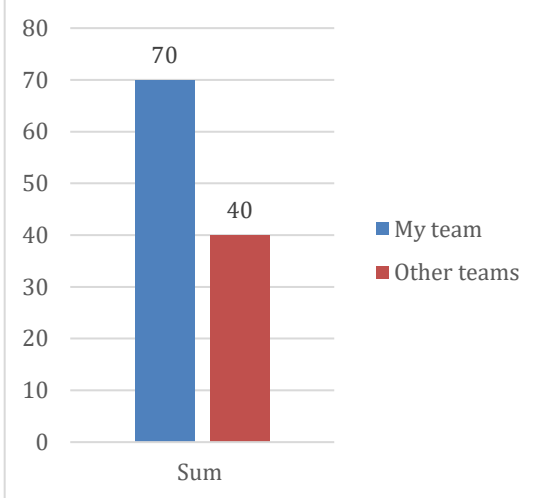
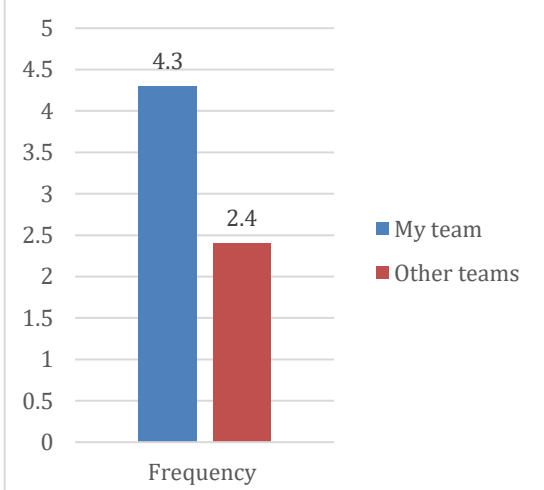
ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization	
M01	Sum of my posts vs average(sum) of all others	Learner	A	I	EN1		
M02	Frequency of my posts vs average(frequency) of all others	Learner	A	I	EN1, EN17		

<sup>2</sup> (A)ctivity/(C)ourse<sup>3</sup> (I)ndividual/(T)eam

\*Note: all analytics addressing the Learner can also be accessed by the Instructor, who is essentially the Activity Administrator

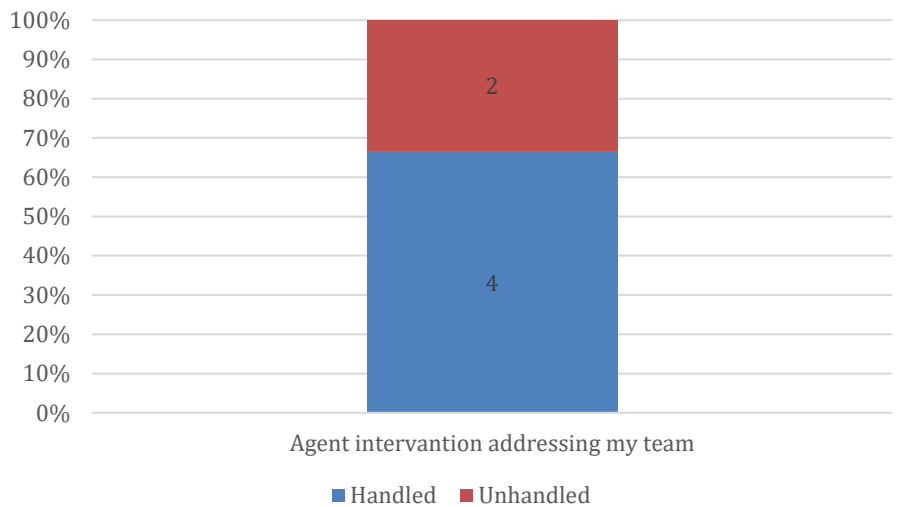
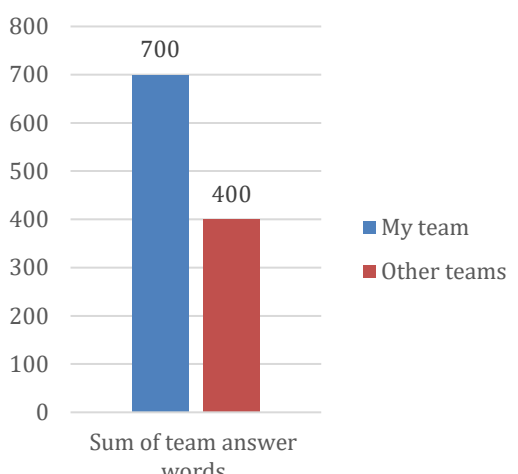
ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization
M03	Timeline of sum of words used in my posts during one activity. & All others average vs my average	Learner	A	I	EN2, EN16, EN17	<p>My posts    My average    Others average</p>
M04	Percentage & number of my replies to agent questions addressing me & Percentage & number of my replies to agent questions addressing my team (handled vs unhandled bot messages)	Learner	A	I	EN3, EN7, EN8	<p>Agent intervantion addressing me    Agent intervantion addressing my team</p> <p>Handled    Unhandled</p>

ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization
M05	Response time to agent interventions directed to me depicted along with the size of agent replies (in words)	Learner	A	I	EN3, EN5, EN16	 <p>● Response time (sec) - Size of bubble means number of words</p>
M06	Timeline of sum of words used in my posts during one activity vs my teammate.  & My teammate average vs my average	Learner	A	T	EN2, EN16, EN17	 <p>● My posts ● Teammate posts ■ My average ■ Teammate's average</p>

ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization						
M07	Sum of my team posts vs average(sum) of all other teams	Learner	A	T	EN1	 <table><thead><tr><th>Team</th><th>Sum</th></tr></thead><tbody><tr><td>My team</td><td>70</td></tr><tr><td>Other teams</td><td>40</td></tr></tbody></table>	Team	Sum	My team	70	Other teams	40
Team	Sum											
My team	70											
Other teams	40											
M08	Frequency of my team posts vs average(frequency) of all other teams	Learner	A	T	EN1, EN17	 <table><thead><tr><th>Team</th><th>Frequency</th></tr></thead><tbody><tr><td>My team</td><td>4.3</td></tr><tr><td>Other teams</td><td>2.4</td></tr></tbody></table>	Team	Frequency	My team	4.3	Other teams	2.4
Team	Frequency											
My team	4.3											
Other teams	2.4											

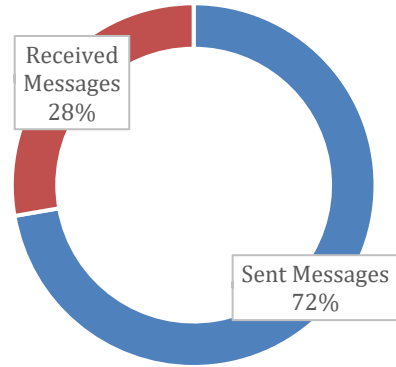
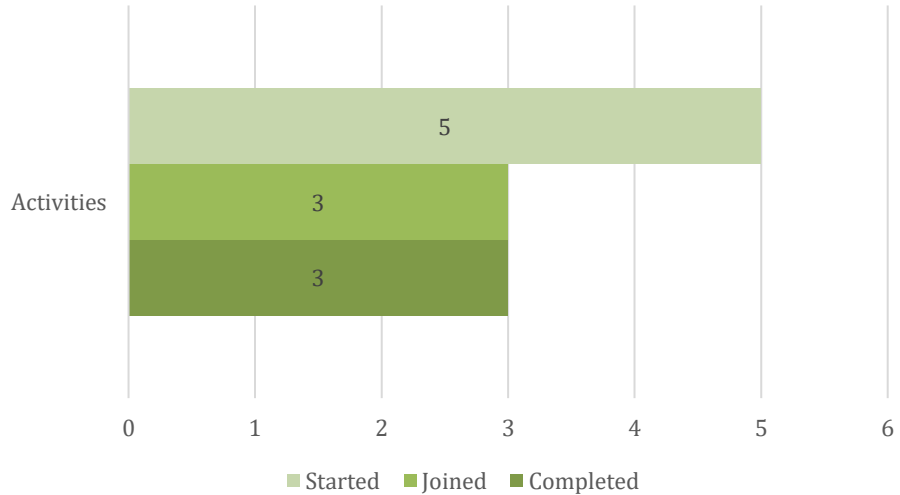
ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization
M09	Distribution of posts in my team. & Average distribution of other teams	Learner	A	T	EN1	<p>Stacked bar chart showing the distribution of posts in my team and the average distribution of other teams. The y-axis represents percentage from 0% to 100%. The x-axis is labeled 'Posts'. The bar for 'Mine' is blue and represents 30%. The bar for 'Teammate's' is red and represents 40%. A green horizontal line indicates the 'Average distribution of other teams' at approximately 55%.</p>
M10	Timeline of sum of words used in my teams' posts during one activity. & All other team's average vs my team's average	Learner	A	T	EN2, EN16, EN17	<p>Scatter plot showing the timeline of sum of words used in my teams' posts during one activity. The y-axis is 'Sum of words' (0 to 25) and the x-axis is 'Timeline of posts' (0 to 10). Blue dots represent 'My team' data points. A red horizontal line represents 'My team's average' at approximately 12. A green horizontal line represents 'Other teams' average' at approximately 6.</p>

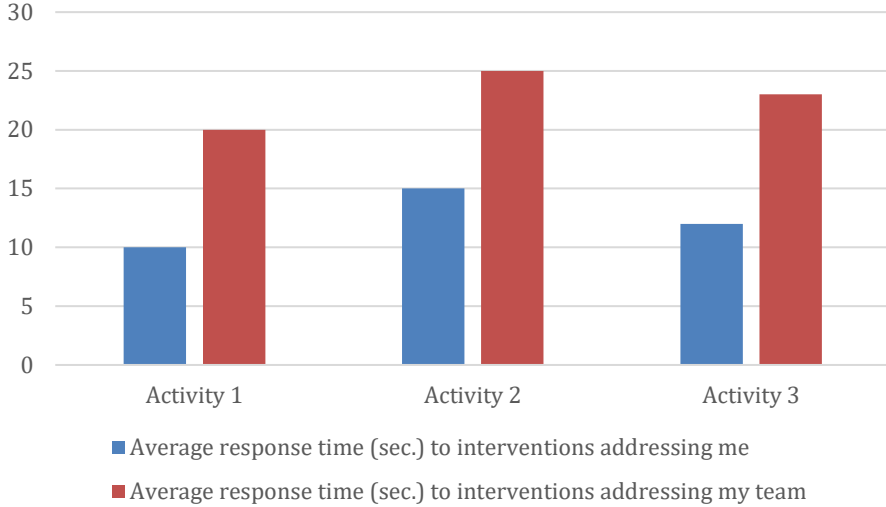
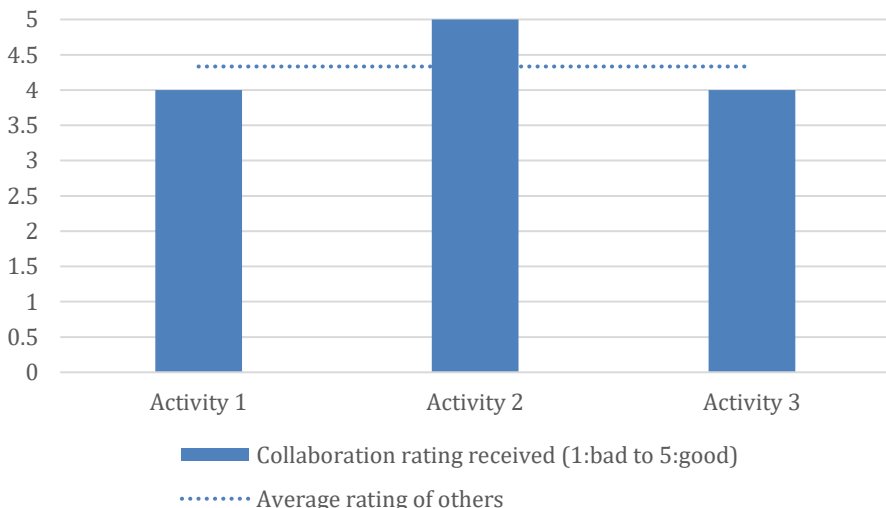


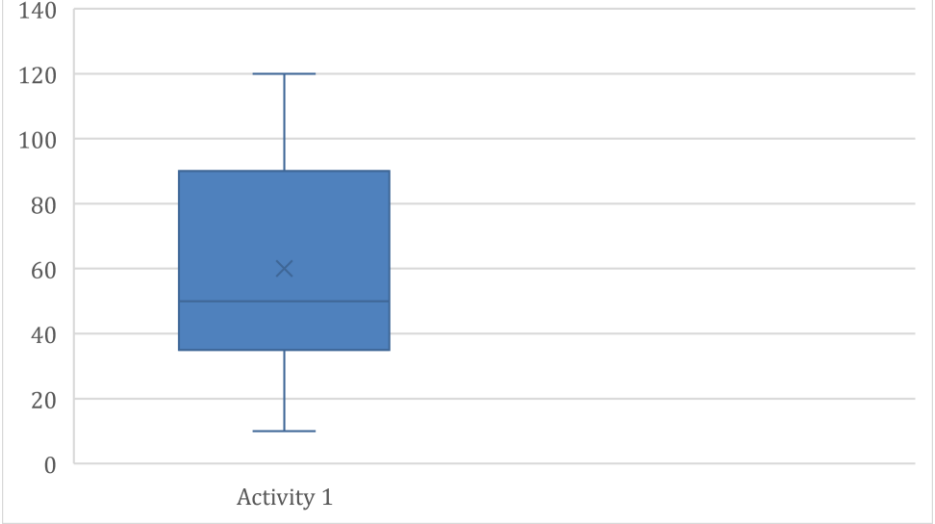
ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization									
M11	Percentage & number of replies to agent questions addressing my team (handled vs unhandled bot messages)	Learner	A	T	EN4, EN8	<div><p>Agent intervantage addressing my team</p><p>■ Handled ■ Unhandled</p><table><thead><tr><th>Category</th><th>Percentage</th><th>Count</th></tr></thead><tbody><tr><td>Handled</td><td>67%</td><td>4</td></tr><tr><td>Unhandled</td><td>33%</td><td>2</td></tr></tbody></table></div>	Category	Percentage	Count	Handled	67%	4	Unhandled	33%	2
Category	Percentage	Count													
Handled	67%	4													
Unhandled	33%	2													
M12	Sum of team answer words vs average of all other teams	Learner	A	T	EN12	<div><p>Sum of team answer words</p><p>■ My team ■ Other teams</p><table><thead><tr><th>Team</th><th>Sum of team answer words</th></tr></thead><tbody><tr><td>My team</td><td>700</td></tr><tr><td>Other teams</td><td>400</td></tr></tbody></table></div>	Team	Sum of team answer words	My team	700	Other teams	400			
Team	Sum of team answer words														
My team	700														
Other teams	400														

ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization
M13	Timeline of agent intervention and users' posts	Learner	A	T	EN15, EN16, EN17	<p>0 5 10 15 20</p> <p>● Agent intervention ● My posts ● Teammate's posts</p>
M14	Response time to agent interventions directed to my team depicted along with the size of agent replies (in words)	Learner	A	T	EN6, EN14, EN15	<p>-2 0 2 4 6 8 10</p> <p>90 80 70 60 50 40 30 20 10 0</p> <p>● Response time (sec) - Size of bubble means number of words</p>

ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization												
M15	Sum of my posts per activity vs average(sum) of all others per activity	Learner	C	I	EN1	<table><thead><tr><th>Activity</th><th>My posts</th><th>Avg. posts of others</th></tr></thead><tbody><tr><td>Activity 1</td><td>200</td><td>100</td></tr><tr><td>Activity 2</td><td>190</td><td>150</td></tr><tr><td>Activity 3</td><td>160</td><td>120</td></tr></tbody></table>	Activity	My posts	Avg. posts of others	Activity 1	200	100	Activity 2	190	150	Activity 3	160	120
Activity	My posts	Avg. posts of others																
Activity 1	200	100																
Activity 2	190	150																
Activity 3	160	120																
M16	Percentage & number of my replies to agent questions addressing me (handled vs unhandled bot messages)	Learner	C	I	EN3, EN7, EN9	<table><thead><tr><th>Activity</th><th>Handled</th><th>Unhandled</th></tr></thead><tbody><tr><td>Activity 1</td><td>80%</td><td>20%</td></tr><tr><td>Activity 2</td><td>35%</td><td>65%</td></tr><tr><td>Activity 3</td><td>25%</td><td>75%</td></tr></tbody></table>	Activity	Handled	Unhandled	Activity 1	80%	20%	Activity 2	35%	65%	Activity 3	25%	75%
Activity	Handled	Unhandled																
Activity 1	80%	20%																
Activity 2	35%	65%																
Activity 3	25%	75%																

ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization
M17	Overall ratio of sent messages to received messages across all activities	Learner	C	I	EN1, EN9	 <p>Received Messages 28%</p> <p>Sent Messages 72%</p> <p>■ Sent Messages ■ Received Messages</p>
M18	Statistics illustrating participation in activities (initiated vs joined vs completed)	Learner	C	I	EN9, EN10, SM3	 <p>Activities</p> <p>5</p> <p>3</p> <p>3</p> <p>0 1 2 3 4 5 6</p> <p>■ Started ■ Joined ■ Completed</p>

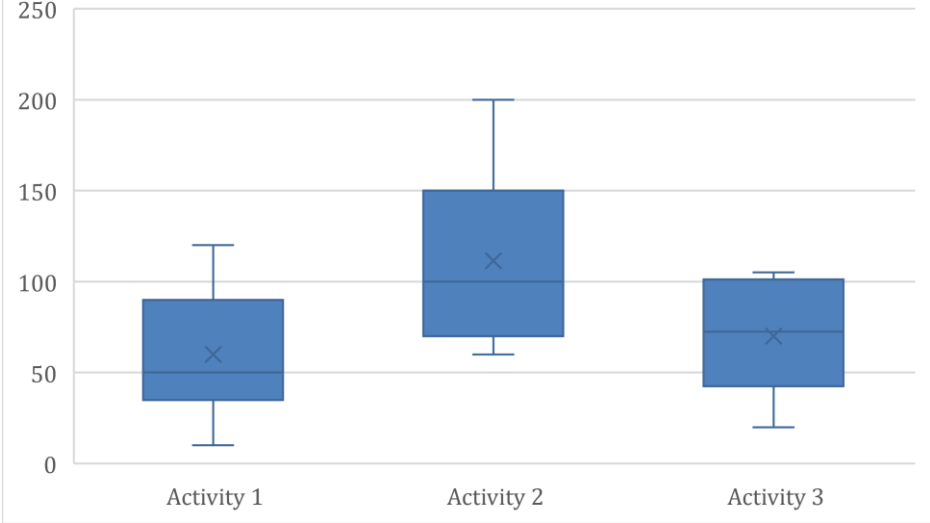
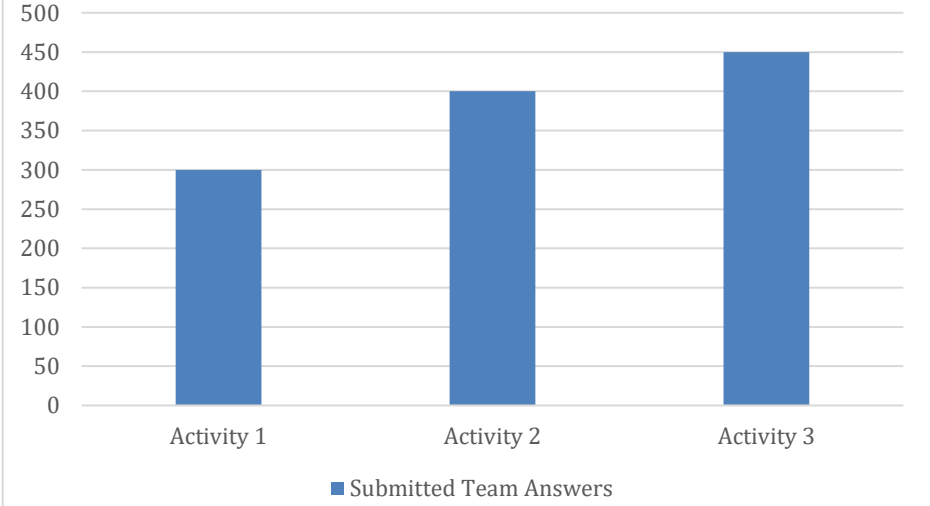
ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization												
M19	Average response time to agent interventions directed to me depicted along with the average response time to agent interventions directed to my team in all activities	Learner	C	I	EN13, EN14	 <table><caption>Data for M19 Visualization</caption><thead><tr><th>Activity</th><th>Average response time (sec.) to interventions addressing me</th><th>Average response time (sec.) to interventions addressing my team</th></tr></thead><tbody><tr><td>Activity 1</td><td>10</td><td>20</td></tr><tr><td>Activity 2</td><td>15</td><td>25</td></tr><tr><td>Activity 3</td><td>12</td><td>23</td></tr></tbody></table>	Activity	Average response time (sec.) to interventions addressing me	Average response time (sec.) to interventions addressing my team	Activity 1	10	20	Activity 2	15	25	Activity 3	12	23
Activity	Average response time (sec.) to interventions addressing me	Average response time (sec.) to interventions addressing my team																
Activity 1	10	20																
Activity 2	15	25																
Activity 3	12	23																
M20	Perceived collaboration rating per activity (based on partners' reviews) vs Average rating of others	Learner	C	I	SM5	 <table><caption>Data for M20 Visualization</caption><thead><tr><th>Activity</th><th>Collaboration rating received (1:bad to 5:good)</th><th>Average rating of others</th></tr></thead><tbody><tr><td>Activity 1</td><td>4</td><td>4.3</td></tr><tr><td>Activity 2</td><td>5</td><td>4.3</td></tr><tr><td>Activity 3</td><td>4</td><td>4.3</td></tr></tbody></table>	Activity	Collaboration rating received (1:bad to 5:good)	Average rating of others	Activity 1	4	4.3	Activity 2	5	4.3	Activity 3	4	4.3
Activity	Collaboration rating received (1:bad to 5:good)	Average rating of others																
Activity 1	4	4.3																
Activity 2	5	4.3																
Activity 3	4	4.3																

ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization
M21	Users' posts in the activity	Teacher	A	I	EN1	 <p>Activity 1</p> <p>Sum of all users' posts in the activity: 7942</p>
M22	List of topics covered by students' posts	Teacher	A	T	EN11	<ul style="list-style-type: none"> <li>✓ Educational software</li> <li>✓ Formative evaluation</li> <li>✓ Summative evaluation</li> <li>✗ Python</li> <li>✓ Software lifecycle</li> </ul>

ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization
M23	Popular topics detected in the answers submitted by the teams	Teacher	A	T	EN12	
M24	Topic flow diagram providing a visualization that maps the most common user journeys through the colMOOC chat activity	Teacher	A	T	LOG	<p>Note: in this type of flow diagram, the link thickness is proportional to the flow quantity.</p>
M25	Activity overview illustrating agent interventions (x-axis: teams, y-axis: minutes)	Teacher	A	T	EN12, EN15, EN16, EN17	

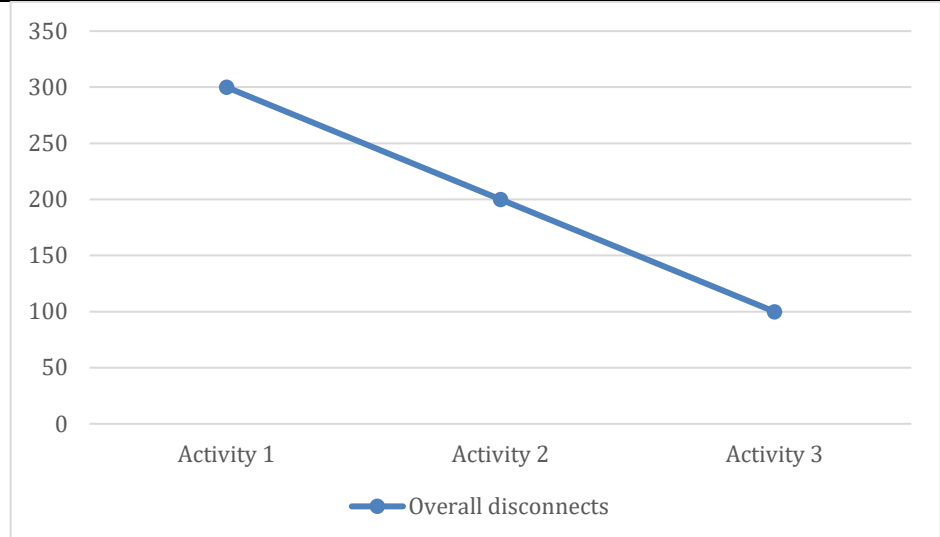
ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization
						<p>Line graph showing four interventions (Intervention 1, 2, 3, 4) over time points T1 to T7. The Y-axis ranges from 0 to 40. Intervention 4 (purple) shows the highest values, starting at ~14 and ending at ~36. Intervention 3 (green) starts at ~8 and ends at ~24. Intervention 2 (red) starts at ~4 and ends at ~13. Intervention 1 (blue) starts at ~1 and ends at ~5.</p>
M26	Growth of active colMOOC users over time	Teacher	C	I	R2	<p>Stacked area chart showing the growth of active colMOOC users (blue) and MOOC users (red) over three activities. The Y-axis ranges from 0 to 1400. At Activity 1, colMOOC users are ~100 and MOOC users are ~600. At Activity 2, colMOOC users are ~400 and MOOC users are ~700. At Activity 3, colMOOC users are ~600 and MOOC users are ~700.</p>



ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization
M27	Statistics concerning users' posts per activity	Teacher	C	T	EN1, EN9	 <p>Sum of users' posts in activity 1: 7942 Sum of users' posts in activity 2: 9906</p>
M28	Sum of answers submitted per activity	Teacher	C	T	EN9, EN12	

ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization															
M29	Average distribution of posts within the teams in each activity	Teacher	C	T	EN1	<table><thead><tr><th>Activity</th><th>Peer A (%)</th><th>Peer B (%)</th></tr></thead><tbody><tr><td>Activity 1</td><td>68</td><td>32</td></tr><tr><td>Activity 2</td><td>60</td><td>40</td></tr><tr><td>Activity 3</td><td>55</td><td>45</td></tr><tr><td>Activity 4</td><td>50</td><td>50</td></tr></tbody></table>	Activity	Peer A (%)	Peer B (%)	Activity 1	68	32	Activity 2	60	40	Activity 3	55	45	Activity 4	50	50
Activity	Peer A (%)	Peer B (%)																			
Activity 1	68	32																			
Activity 2	60	40																			
Activity 3	55	45																			
Activity 4	50	50																			
M30	Number of unhandled agent interventions directed to individuals & groups in each activity	Teacher	C	I	EN7, EN9, EN8,	<table><thead><tr><th>Activity</th><th>Unhandled agent interventions addressing an individual user</th><th>Unhandled agent interventions addressing a group</th></tr></thead><tbody><tr><td>Activity 1</td><td>10</td><td>20</td></tr><tr><td>Activity 2</td><td>15</td><td>25</td></tr><tr><td>Activity 3</td><td>12</td><td>23</td></tr></tbody></table>	Activity	Unhandled agent interventions addressing an individual user	Unhandled agent interventions addressing a group	Activity 1	10	20	Activity 2	15	25	Activity 3	12	23			
Activity	Unhandled agent interventions addressing an individual user	Unhandled agent interventions addressing a group																			
Activity 1	10	20																			
Activity 2	15	25																			
Activity 3	12	23																			

ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization
M31	Average activities duration and size (in words)	Teacher	C	T	EN2, EN9, EN17	<p>Note: The size of the bubbles represent the average size (in words) of discussions</p>
M32	Retention rate of users (from the first to the last activity)	Teacher	C	I	R1	

ID	Metric	User*	Scope <sup>2</sup>	Level <sup>3</sup>	Variable(s)	Visualization								
M33	Number of disconnects per activity	Teacher	C	I	EN9, EN18	 <table><tr><th>Activity</th><th>Overall disconnects</th></tr><tr><td>Activity 1</td><td>300</td></tr><tr><td>Activity 2</td><td>200</td></tr><tr><td>Activity 3</td><td>100</td></tr></table>	Activity	Overall disconnects	Activity 1	300	Activity 2	200	Activity 3	100
Activity	Overall disconnects													
Activity 1	300													
Activity 2	200													
Activity 3	100													

### 2.3.3 Use Cases

This paragraph presents the various use cases for the users of the LA module (namely Learner and Instructor)

#### 2.3.3.1 Learner Use Cases

##### 2.3.3.1.1 My Performance in the Activity

**Table 3. Learner Use Cases - My Performance in the Activity**

Use Case Title	My Performance in the Activity
<b>Description</b>	Following the completion of an activity, the learner decides to check their performance in the specific activity. The learner visits the offline analytics dashboard, sets the scope to ‘activity level’, selects the particular activity from the available dropdown menu, and selects my performance in the activity. The latter panel displays a series of personal statistics as regards the performance of the user and how well it compares to other users’ performance in the particular colMOOC activity.
<b>Actor</b>	Learner / Instructor of this course
<b>Input</b>	EN1, EN2, EN3, EN4, EN5, EN7, EN8, EN16, EN17
<b>Output</b>	<ul style="list-style-type: none"> <li>• M1: Sum of my posts vs average(sum) of all others</li> <li>• M2: Frequency of my posts vs average(frequency) of all others</li> <li>• M3: Timeline of sum of words used in my posts during one activity &amp; All others average vs my average</li> <li>• M4: Percentage &amp; number of my replies to agent questions addressing me &amp; Percentage &amp; number of my replies to agent questions addressing my team (handled vs unhandled bot messages)</li> <li>• M5: Response time to agent interventions directed to me depicted along with the size of agent replies (in words)</li> </ul>
<b>Flowchart</b>	<pre> graph TD     Learner((Learner)) --&gt; OfflineAnalytics[Offline Analytics Module]     OfflineAnalytics --&gt; SelectScope{Select Scope}     SelectScope -- activity --&gt; SelectChatActivity{Select a Chat Activity}     SelectChatActivity --&gt; ActivityAnalytics[Activity Analytics]     ActivityAnalytics --&gt; SelectLevel{Select Level}     SelectLevel -- individual --&gt; MyPerformance[My Performance in the Activity]     MyPerformance --&gt; Metrics[M1 M2 M3 M4 M5]   </pre>

### 2.3.3.1.2 My Team's Performance in the Activity

**Table 4. Learner Use Cases - My Team's Performance in the Activity**

Use Case Title	<b>My Team's Performance in the Activity</b>
Description	Following the completion of an activity, the learner decides to check how their team has performed in the specific activity. The learner visits the offline analytics dashboard, sets the scope to 'activity level', selects the particular activity from the available dropdown menu, and then chooses to view their team's performance in the activity. This panel visualizes a series of metrics showcasing the performance of the user's team and how well it compares to the performance of other teams in the particular colMOOC activity.
Actor	Learner / Instructor of this course
Input	EN1, EN2, EN4, EN6, EN8, EN12, EN14, EN15, EN16, EN17
Output	<ul style="list-style-type: none"> <li>• M06: Timeline of sum of words used in my posts during one activity vs my teammate; plus, my teammate average vs my average</li> <li>• M07: Sum of my team posts vs average (sum) of all other teams</li> <li>• M08: Frequency of my team posts vs average (frequency) of all other teams</li> <li>• M09: Distribution of posts in my team &amp; average distribution of other teams</li> <li>• M10: Timeline of sum of words used in my teams' posts during one activity; plus, all other team's average vs my team's average</li> <li>• M11: Percentage &amp; number of replies to agent questions addressing my team (handled vs unhandled bot messages)</li> <li>• M12: Sum of team answer words vs average of all other teams</li> <li>• M13: Timeline of agent intervention and users' posts</li> <li>• M14: Response time to agent interventions directed to my team depicted along with the size of agent replies (in words)</li> </ul>
Flowchart	<pre> graph TD     Learner((Learner)) --&gt; OfflineAnalytics[Offline Analytics Module]     OfflineAnalytics --&gt; SelectScope{Select Scope}     SelectScope -- activity --&gt; SelectChatActivity{Select a Chat Activity}     SelectChatActivity -- Activity Analytics --&gt; SelectLevel{Select Level}     SelectLevel -- team --&gt; Performance[My Team's Performance in the Activity]     Performance --&gt; Metrics[M06 M07 M08 M09 M10 M11 M12 M13 M14]   </pre>

### 2.3.3.1.3 My Contribution in the Course

**Table 5. Learner Use Cases - My Contribution in the Course**

Use Case Title	<b>My Contribution in the Course</b>
Description	Following the completion of an activity, the learner decides to check the level of their contribution in the course. The learner visits the offline analytics dashboard, sets the scope to 'course level', selects their course from the available dropdown menu, and then selects to view contribution-related metrics. The specific panel includes a series of visualizations illustrating the user contribution in the particular course and comparing those statistics with the ones of other users/teams participating in the specific colMOOC-compatible course.
Actor	Learner / Instructor of this course
Input	EN1, EN9, EN10, SM1
Output	<ul style="list-style-type: none"> <li>M15: Sum of my posts per activity vs average(sum) of all others per act</li> <li>M18: Statistics illustrating participation in activities (initiated vs joined vs completed)</li> </ul>
Flowchart	<pre> graph TD     Learner((Learner)) --&gt; OAM[Offline Analytics Module]     OAM --&gt; SS{Select Scope}     SS --&gt; CA[Course Analytics]     CA --&gt; SV{Select View}     SV -- Contribution --&gt; CC[Course Contribution]     CC --&gt; M15[M15]     CC --&gt; M18[M18] </pre>

### 2.3.3.1.4 Peer Interaction in the Course

**Table 6. Learner Use Cases - Peer Interaction in the Course**

Use Case Title	Peer Interaction in the Course
Description	Following the completion of an activity, the learner decides to view statistics regarding their peer interaction in the course. The learner visits the offline analytics dashboard, sets the scope to 'course level', selects their course from the available dropdown menu, and then selects to examine metrics relating to user's peer interactions. The specific panel showcases data relating to users' peer interaction in the selected colMOOC-compatible course.
Actor	Learner / Instructor of this course
Input	EN1, EN9, SM5
Output	<ul style="list-style-type: none"> <li>M17: Overall ratio of sent messages to received messages across all acti</li> <li>M20: Perceived collaboration rating per activity (based on partners' reviews) vs average rating of others</li> </ul>
Flowchart	<pre> graph TD     Learner((Learner)) --&gt; OAM[Offline Analytics Module]     OAM --&gt; SS{Select Scope}     SS --&gt; CA[Course Analytics]     CA --&gt; SV{Select View}     SV -- "Peer Interaction" --&gt; PIC[Peer Interaction in Course]     PIC --&gt; M17M20[M17   M20]           </pre> <p>The flowchart illustrates the process for a learner to view peer interaction statistics. It begins with the 'Learner' actor, represented by a black circle, who interacts with the 'Offline Analytics Module' (blue rectangle). This leads to a decision point 'Select Scope' (purple diamond). From there, the flow proceeds to 'Course Analytics' (blue rectangle), then to another decision point 'Select View' (purple diamond). From 'Select View', a path labeled 'Peer Interaction' leads to the 'Peer Interaction in Course' (blue rectangle), which finally outputs to a box containing 'M17' and 'M20' (blue rectangle).</p>



### 2.3.3.1.5 User-Agent Interaction in the Course

**Table 7. Learner Use Cases - User-Agent Interaction in the Course**

Use Case Title	User-Agent Interaction in the Course
Description	Following the completion of an activity, the learner decides to view statistics regarding their interaction with the agent in the course. The learner visits the offline analytics dashboard, sets the scope to ‘course level’, selects their course from the available dropdown menu, and then selects to examine metrics relating to their interaction with the colMOOC conversational agent. The specific panel showcases data relating to user-agent interactions throughout the colMOOC activities completed in the context of the specific course.
Actor	Learner / Instructor of this course
Input	EN3, EN7, EN9, EN13, EN14
Output	<ul style="list-style-type: none"> <li>M16: Percentage &amp; number of my replies to agent questions addressing me (handled vs unhandled bot messages)</li> <li>M19: Average response time to agent interventions directed to me depicted along with the average response time to agent interventions directed to my team in all activities</li> </ul>
Flowchart	<pre> graph TD     Learner((Learner)) --&gt; Analytics[Offline Analytics Module]     Analytics --&gt; Scope{Select Scope}     Scope --&gt; Course[Course Analytics]     Course --&gt; View{Select View}     View -- "User-Agent Interaction" --&gt; Interaction[User-Agent Interaction in Course]     Interaction --&gt; Metrics[M16   M19]           </pre> <p>The flowchart illustrates the process of viewing user-agent interaction statistics. It begins with the 'Learner' actor, represented by a black circle, who interacts with the 'Offline Analytics Module' (blue rectangle). This leads to a decision point 'Select Scope' (purple diamond). From there, the flow proceeds to 'Course Analytics' (blue rectangle), then to another decision point 'Select View' (purple diamond). From 'Select View', the flow goes to 'User-Agent Interaction in Course' (blue rectangle) via the path labeled 'User-Agent Interaction'. Finally, this leads to the output metrics 'M16' and 'M19' (blue rectangles).</p>

### 2.3.3.2 Instructor Use Cases

#### 2.3.3.2.1 View Learner Analytics

**Table 8. Instructor Use Cases - View Learner Analytics**

Use Case Title	<b>Advanced Activity Analytics</b>
Description	The instructor/teacher enters the Learning Analytics Dashboard and decides that he/she would like to view the statistics of a specific learner. The instructor clicks on the button ‘select a student view’ and searches for a particular student by entering their registered name or ID. When the instructor finds and clicks the student of their interest, the dashboard refreshes and the instructor is able to view all the analytics that the specific student has access to. The new analytics dashboard panel includes all student-specific options and metrics described in the previous use cases.
Actor	Instructor
Input	Learner’s name or ID
Output	In contrast to the other use cases included in this document, this use case is regarded as a higher-level one and its output items are a series of other use cases instead of direct metrics. These are listed below: <ul style="list-style-type: none"> <li>• 4.3.3.1.1: My Performance in the Activity</li> <li>• 4.3.3.1.2: My Team’s Performance in the Activity</li> <li>• 4.3.3.1.3: My Contribution in the Course</li> <li>• 4.3.3.1.4: Peer Interaction in the Course</li> <li>• 4.3.3.1.5: User-Agent Interaction in the Course</li> </ul>
Flowchart	<pre> graph TD     Teacher((Teacher)) --&gt; OAM[Offline Analytics Module]     OAM --&gt; LV[Learner View]     LV --&gt; SL{Select Learner}     SL --&gt; Learner((Learner))     Learner --&gt; U1[4.3.3.1.1]     U1 --&gt; U2[4.3.3.1.2]     U2 --&gt; U3[4.3.3.1.3]     U3 --&gt; U4[4.3.3.1.4]     U4 --&gt; U5[4.3.3.1.5] </pre>

### 2.3.3.2.2 Advanced Activity Analytics

**Table 9. Instructor Use Cases - Advanced Activity Analytics**

Use Case Title	<b>Advanced Activity Analytics</b>
Description	Following the completion of an activity, the instructor/teacher decides to view statistics regarding the participants' overall performance in an activity. The teacher visits the offline analytics dashboard, sets the scope to 'activity level', and selects a specific activity from the available dropdown menu. This analytics dashboard panel displays a series of metrics depicting the performance of all participants and teams in the selected activity.
Actor	Instructor
Input	EN1, EN11, EN12, EN15, EN16, EN17, LOG
Output	<ul style="list-style-type: none"> <li>• M21: Users' posts in the activity</li> <li>• M22: List of topics covered by students' posts</li> <li>• M23: Popular topics detected in the answers submitted by the teams</li> <li>• M24: Topic flow diagram providing a visualization that maps the most common user journeys through the colMOOC chat activity</li> <li>• M25: Activity overview illustrating agent interventions (x-axis: teams, y-axis: minutes)</li> </ul>
Flowchart	<pre> graph TD     Teacher((Teacher)) --&gt; OfflineAnalytics[Offline Analytics Module]     OfflineAnalytics --&gt; SelectScope{Select Scope}     SelectScope -- Activity --&gt; ActivityAnalytics[Activity Analytics]     ActivityAnalytics --&gt; ActivityAdvancedAnalytics[Activity Advanced Analytics]     ActivityAdvancedAnalytics --&gt; Metrics[M21 M22 M23 M24 M25] </pre>

### 2.3.3.2.3 Advanced Course Analytics

**Table 10. Instructor Use Cases - Advanced Course Analytics**

Use Case Title	<b>Advanced Course Analytics</b>
Description	Following the completion of an activity, the instructor/teacher decides to view statistics regarding the participants' overall performance in a course. The teacher visits the offline analytics dashboard, sets the scope to 'course level', and selects a specific course. The dashboard panel that emerges depicts a series of metrics showcasing the performance of all participants and teams in the specific course.
Actor	Instructor
Input	EN1, EN2, EN7, EN8, EN9, EN12, EN17, EN18, R1
Output	<ul style="list-style-type: none"> <li>• M26: Growth of active colMOOC users over time</li> <li>• M27: Statistics concerning users' posts per activity</li> <li>• M28: Sum of answers submitted per activity</li> <li>• M29: Average distribution of posts within the teams in each activity</li> <li>• M30: Number of unhandled agent interventions directed to individuals &amp; groups in each activity</li> <li>• M31: Average activities duration and size (in words)</li> <li>• M32: Retention rate of users (from the first to the last activity)</li> <li>• M33: Number of disconnects per activity</li> </ul>
Flowchart	<pre> graph TD     Teacher((Teacher)) --&gt; OfflineAnalytics[Offline Analytics Module]     OfflineAnalytics --&gt; SelectScope{Select Scope}     SelectScope -- Course --&gt; CourseAnalytics[Course Analytics]     CourseAnalytics --&gt; CourseAdvancedAnalytics[Course Advanced Analytics]     CourseAdvancedAnalytics --&gt; Metrics[M26 M27 M28 M29 M30 M31 M32] </pre>

### 3 LA Module development

#### 3.1 Database Model

As mentioned in the Technological roadmap of the project (Deliverable 4.1) in order to keep the architecture coherent and manageable, and simplify hosting and maintenance, a controlled set of technologies have been used to implement the colMOOC Learning Analytics module.

We also followed our general principle, using open-source technologies to implement the colMOOC platform (“open-source” defined as software licensed under an OSI -approved license that is included in the list of EUPL-compatible licenses).

Taking care all of the above requirements we chose the MySQL, an open-source relational database management system (RDBMS) for our storage solution. The latest version of MySQL is a very fast, reliable, and easy to use relational database management system based on SQL – Structured Query Language. It is developed, distributed, and supported by Oracle Corporation as open source and it is ideal for storing our data for all colMOOC modules.

From the server perspective we chose the MariaDB Server which is one of the most popular open-source relational databases and it’s made by the original developers of MySQL. Finally in order to handle the administration of the colMOOC MySQL database (managing databases, tables, columns, relations, indexes, users, permissions, etc) we chose the free software tool phpMyAdmin that is written in PHP.

The Database model of the Learning Analytics module uses three MySQL tables in UTF-8 encoding in order to ensure correct and consistent handling of multilingual content. Their schema is presented in the Figure below:

<b>colmooc.mooc_sessions_pairs</b> id : int(11) player1_id : varchar(250) player2_id : varchar(250) activity_id : int(11) room_id : int(11) created : timestamp session_data : longtext session_answer : varchar(250)	<b>colmooc.mooc_analytics</b> id : int(11) session_id : int(11) activity_id : int(11) player_id : int(11) partner_id : int(11) moderator : tinyint(3) posts_total : int(11) words_total : int(11) activity_duration : int(11) my_posts : longtext my_posts_avg : double my_frequency : double team_answer_words : int(11) agent_posts : longtext disconnects_total : int(11) agent_quest_me : int(11) agent_answer_me : int(11) response_answer_me : longtext agent_quest_team : int(11) agent_answer_team : int(11) response_answer_team : longtext
<b>colmooc.mooc_lasessions</b> lasession_id : int(11) platform_id : varchar(250) course_id : varchar(250) teacher_id : varchar(250) colstudent_id : int(11) created : timestamp closed : timestamp lasession_token : varchar(250) lasession_status : tinyint(3)	

**Figure 10 : Schema of the colMOOC MySQL tables for LA module**

All chat activities are stored in the `mooc_sessions_pairs` table as described below:

- **ID**: a unique id for each Conversational Agent Player session
- **player1\_id**: the id of the first student
- **player2\_id**: the id of the second student
- **activity\_id**: the activity's unique id
- **created**: timestamp of the database entry
- **session\_data**: data from this session used for processing
- **session\_answer**: the submitted final answer to the activity topic

For keeping track of the Learning Analytics module's usage and also for authentication functionality, we store each session request in the `mooc_lasessions` table. The structure of the table described below:

- **ID**: a unique id for each LA module session
- **platform\_id**: the unique id of the external platform
- **course\_id**: the unique id of the course
- **teacher\_id**: the teacher's unique id (zero value if the session is about a student)
- **colstudent\_id**: the student's unique id (zero value if the session is about a teacher)
- **created**: timestamp of the database entry
- **closed**: timestamp of the session ending
- **lasession\_token**: 16-bit random token for authentication functionality
- **lasession\_status**: current status of LA session (open or closed)

Moreover, we save an extra database entry for each student performance in the `mooc_analytics` table. By using this extra table, we are able to pre-calculate all necessary statistical data that we are using in various charts and graphs in the Learning Analytics module and ultimately achieving the best possible response performance for user experience. The structure of the table described below:

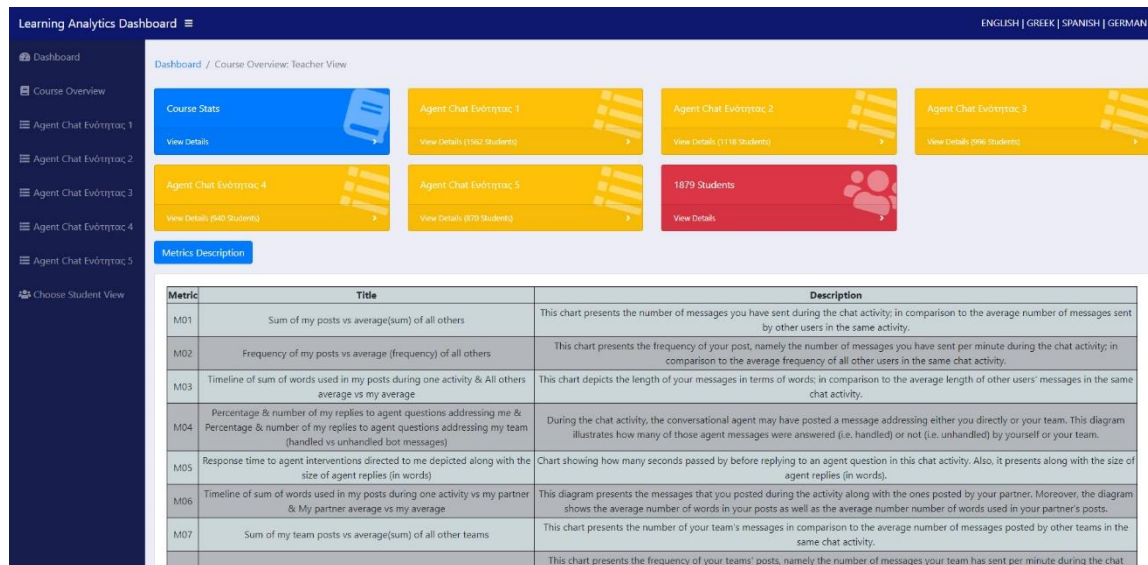
- **ID**: a unique id for each analytics entry
- **session\_id**: the unique id of the LA session
- **activity\_id**: the unique id of the colMOOC activity
- **player\_id**: the student's unique id
- **partner\_id**: the partner's unique id (student paired for a specific activity)
- **moderator**: indicates if the student was the moderator
- **posts\_total**: total posts written during the activity
- **words\_total**: total words written during the activity
- **activity\_duration**: duration of the activity (seconds)
- **my\_posts**: total posts written by this student during the activity
- **my\_posts\_avg**: average posts written by this student during the activity
- **my\_frequency**: frequency of posts written by this student during the activity
- **team\_answer\_words**: total words written in the final topic answer
- **agent\_posts**: total agent's interventions during the activity
- **disconnects\_total**: total disconnections occurred during the activity
- **agent\_quest\_me**: agent's questions directed at this student only
- **agent\_answer\_me**: answers from this student on agent's questions
- **response\_answer\_me**: response time (seconds) from this student's answers
- **agent\_quest\_team**: agent's questions directed for both participants
- **agent\_answer\_team**: answers from both students on agent's questions
- **response\_answer\_team**: response time (seconds) from both students' answers

### 3.2 Implementation

The implementation of the Learning Analytics module on the server side was accomplished with the popular scripting language PHP. The latest release of PHP (version 7) is an efficient and powerful tool for making dynamic and interactive web pages. The final output is built on the server and the result is sent as HTML5 to the client browser for rendering. For styling our web pages (colors, layout, fonts, etc.) and adapt the web page presentation for different types of devices, we are using CSS3 language. Furthermore, for database functionalities and interactive elements that engage the users, we chose to proceed with JavaScript language and the jQuery library.

For hosting service we are using an Apache HTTP Server, a free and open-source cross platform web server that delivers web content through the internet and it's by far the most popular HTTP client on the web. The colMOOC server runs with Linux operating system which is also an open-source Unix-like system.

Moreover, the Learning Analytics module works as a multilingual software system. We created a basic template in PHP format that includes all different UI/UX labels and with the collaboration of our colMOOC partners we managed to completely translate it in Greek, Spanish and German language (English is the initial basic language). Users can change the LA dashboard language by choosing their preference at the right top corner of their screen.



**Figure 11 : Landing page of the Learning Analytics Dashboard**

The landing page of the Learning Analytics module is a dashboard with a menu on the left for navigation purposes and the main content on the right of every web page. In this main content area the users can see all the activities that took place in their MOOC as yellow cards as well as one Course blue card that links to all the Course related statistics. For teachers view only there is another card with all the students who enrolled to the course and the capability to view the module as a specific student of their choice.

There is also a table with Metrics description that concludes all different metrics of the module accompanied with a brief description for quick reference.

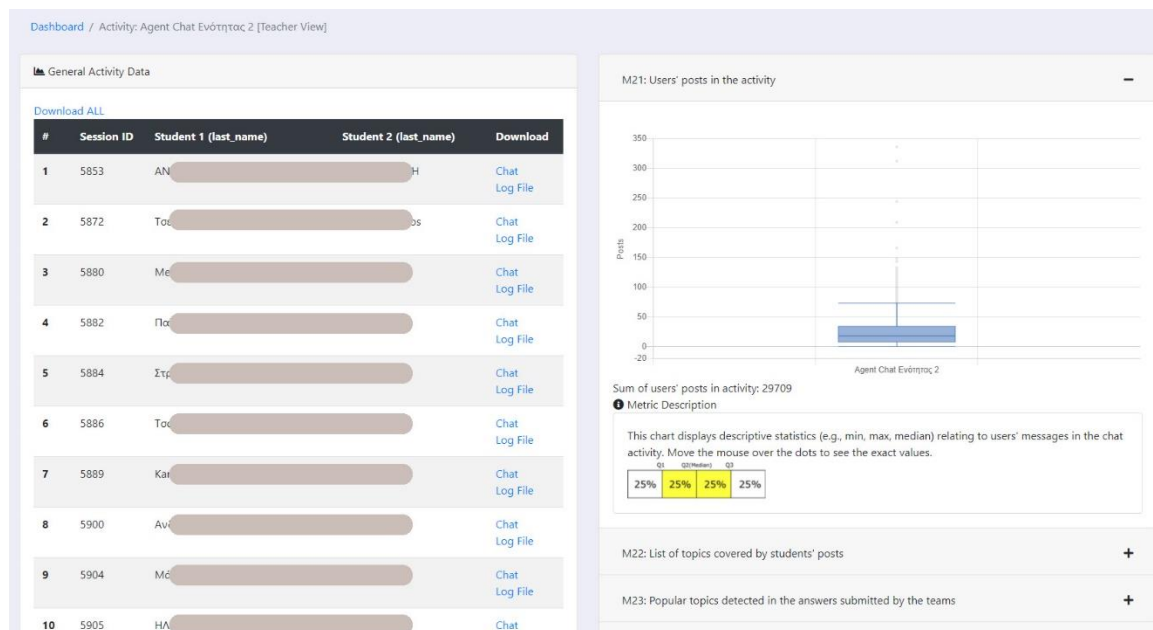
Most of the Learning Analytics module web elements were developed with the latest version of the Bootstrap framework, which in simple terms is a giant collection of handy, reusable bits of code written in HTML, CSS, and JavaScript.

The Course Stats page visualize all the needed metrics that described in Section 3 of this document. The users (teachers or students) can navigate through different metrics by clicking on their titles in order to load them. This accordion functionality is ideal because displays collapsible content panels (toggle between hiding and showing) for presenting information in a limited amount of space.



**Figure 12 : Course Stats page (Teacher view)**

The Activity Stats page works with the same logic. Metrics are hiding/showing with an accordion functionality that gives user freedom to navigate from one to another. Moreover, there is an extra section on the left of the screen with more information about the current activity.



**Figure 13 : Activity Stats page (Teacher view)**



The extra section for Teachers' view provides a full list of Student names who managed to finish this activity and also a unique session id for reference use. Teachers also have the ability to download the log file of each session in order to view the exact chat messages of a specific activity.

In a related way, the extra section for Students' view provides them with the Activity Topic answer that they have submitted during their participation in the Conversational Agent Player.



**Figure 14 : Activity Stats page (Student view)**

All the graphs of the Learning Analytics module developed and designed with the Chart.js open source Javascript library. Chart.js generates easy to use, interactive, animated and multilingual graphs with great rendering performance across all modern browsers.

Finally, all the tables developed with the DataTables plug-in for the jQuery Javascript library. Datatables is a highly flexible tool, built upon the foundations of progressive enhancement, that adds a bunch of advanced features to any HTML table like Pagination, Instant search and Ordering functionalities.

### 3.3 Evaluating the Consistency and Quality

For the evaluation of the consistency and quality of the colMOOC Learning Analytics module we selected to present some quality metrics of the module which are considered as the key factors of the implementation.

There are many scalability dimensions in the integration and mechanisms used which affect favorably the overall performance and throughput of the system. The final version of the module can comfortably accommodate the anticipated load of the colMOOC pilots, and has the capacity to support a higher load, given the current installation and deployment. In addition, there are various scalability factors, affecting performance that can be applied when the system load gets considerably larger.

API metrics helped us understand the operational performance of the colMOOC APIs, and gave us indications of the performance of the Learning Analytics module and colMOOC Communication API.

- CPU Usage (Infrastructure metric)

CPU usage is one of the most classic performance metrics that can be a proxy to application responsiveness. High Server CPU usage can mean the server is oversubscribed and overloaded or it can mean a performance bug in our application such as too many spinlocks. We used CPU usage (along with its affiliate metric, memory usage) for resource planning and measuring overall health. We monitored the colMOOC server during each colMOOC pilot via Activity Monitor on Linux OS and we never recorded a CPU usage higher than 50%.

- Memory Usage (Infrastructure API metric)

Like CPU usage, memory usage is also a good proxy for measuring resource utilization as CPU and memory capacity are physical resources unlike a metric which may be more configuration dependent. A server with extremely low memory usage can either be downsized or have additional services allocated to that server to consume additional memory. On the flip side, high memory usage can be an indicator of servers overloaded. We did the same monitoring process as CPU usage during each colMOOC pilot and the Memory Usage never recorded values higher than 60% of a total 16GB capacity.

- Number of API calls (Application API metric)

For scalability and fault tolerance the API can run with a number of servers acting as cooperating message brokers, cooperating for providing continuous service. Running multiple brokers means that for each partition there shall be a single broker acting as the designated leader, and a list of brokers acting as replicas. Currently the colMOOC API is deployed over 5 brokers. The number of brokers can be scaled up based on need, but for the foreseeable future there is no expectation that the platform would require more brokers to be deployed.

The number of API calls varied depending on day of the week or even hour per day especially when the colMOOC pilots were running but generally never exceed our usage prediction and threshold.

- Number of Errors (Application API metric)

Similar to number of API calls, error rate is the number of API calls with non 200 family of status codes per minute and is critical for measuring how buggy and error-prone our API is. In order to track total number of errors, it is important to understand what type of errors are happening. “500 errors” imply bad things are happening with the code whereas many “400 errors” imply user errors from a poorly designed or documented API. Thus, we developed a public API documentation that we continuously improved with the collaboration of our colMOOC partners. Lots of “400 errors” prevented by using the try-out functionality inside the documentation, while all of the “500 errors” were eliminated with the feedback after each evaluation cycle.

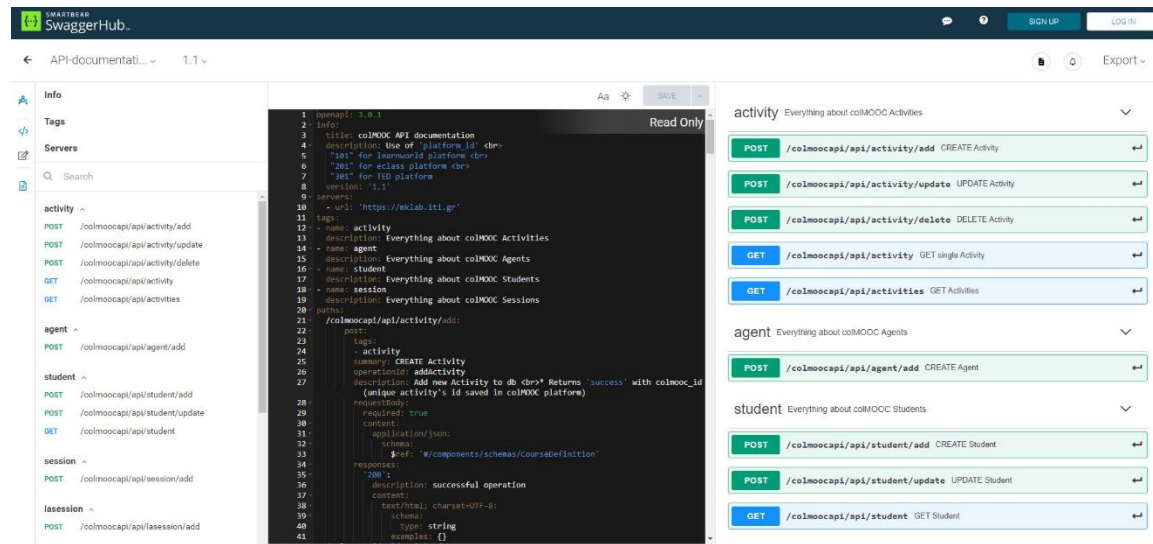
### **3.4 Integration with external platforms**

To accomplish connectivity between an external platform and the Learning module, we created a specific step with instructions for a basic usage flow (Step 3 of the colMOOC Communication API) that all technical partners had to follow carefully.

The colMOOC Communication API was developed with Slim, a PHP micro framework that helps developers quickly write simple yet powerful web applications and APIs. At its core, Slim is a dispatcher that receives an HTTP request, invokes an appropriate callback routine, and returns an HTTP response. Each API call is using 'https' protocol (SSL certificate installed by request) so the connection is reliable and stable.

A public API documentation was developed for helping all technical partners while give them the ability to test each API call before using it. The documentation was written at SwaggerHub platform and can be found below (latest version: 1.1):

<https://app.swaggerhub.com/apis/colMOOC/API-documentation/1.1>



**Figure 15 : Public Communication API documentation at SwaggerHub (version 1.1)**

The Step 3 of the colMOOC Communication API is basically a creation of two different types of buttons (one for the teacher and one for each student) that must be developed in an external platform UI, while accompanied with links to the colMOOC Learning Analytics platform. It's a process of requesting new unique tokens for each visit to the LA module (a similar flow used for the first two steps), that are used for authentication purposes and they need to be added as a URL parameter to its redirection link.

The exact flow of this procedure is being described below, accompanied with comments regarding the appropriate API calls that needs to be called.

- **Event 1: User press Learning Analytics button** on external platform

Actions needed:

A) POST CREATE lasession call (with *platform\_id*, *course\_id*, *teacher\_id* & *colstudent\_id* parameters) and you'll get the *lasession\_id* and *lasession\_token* as response. Note: You need to set *colstudent\_id*=0 (zero) if the button is for teacher use or *teacher\_id*=0 (zero) if the button is for student use. The *course\_id* is the ID inserted at API Step 1 when the activity had initially created.

B) Redirect Teacher/Student to colMOOC Learning Analytics with *lasession\_id* and *lasession\_token* parameters that received previously:

[https://mklab.iti.gr/colmoocapi/analytics/?lasession\\_id={id}&lasession\\_token={token}](https://mklab.iti.gr/colmoocapi/analytics/?lasession_id={id}&lasession_token={token})

Each platform had to implement this button wherever it was more appropriate based on their judgment (each platform has its own design) and by keeping in mind though that the Learning Analytics module is "Course-specific", so it was more appropriate to create those buttons at Course Main Page or somewhere near each Activity link.

## **4 References**

---

- Akhtar, S., Warburton, S., & Xu, W. (2017). The use of an online learning and teaching system for monitoring computer aided design student participation and predicting student success. *International Journal of Technology and Design Education*, 27(2), 251–270.
- Floratos, N., Guasch, T., & Espasa, A. (2015). Recommendations on Formative Assessment and Feedback Practices for stronger engagement in MOOCs. *Open Praxis*, 7(2), 141-152.
- Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, 90, 36-53.
- Hone, K. S., & El Said, G. R. (2016). Exploring the factors affecting MOOC retention: A survey study. *Computers & Education*, 98, 157-168.
- Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10), 1380-1400.
- Xiong, Y., Li, H., Kornhaber, M. L., Suen, H. K., Pursel, B., & Goins, D. D. (2015). Examining the relations among student motivation, engagement, and retention in a MOOC: A structural equation modeling approach. *Global Education Review*, 2(3).