**Automating the analysis of information generated in IIEP’s online forums**

Themes: training, e-learning, knowledge sharing

Associated skills: Data processing, analysis and visualization; learning management systems

1. **Presentation of the project team**

The IIEP-UNESCO Training Unit, through its Moodle platform, organizes capacity development online training and knowledge sharing activities for ministries of education, and national as well as international institutions around the world on different themes related to educational planning. During the training, participants share insights on policy and legislation, experiences and good practices related to their national contexts. This information is extremely helpful for their learning, but also for IIEP. In many courses, this information, used with the participants’ consent, contributes to IIEP’s research and technical cooperation work.

1. **Issue**

Learning Management Systems (LMS) contain an important volume of resources and information generated by instructors and learners. Many of the interactions and exchanges between learners and instructors take place in discussion forums available in the courses hosted in the LMS. The information generated through this knowledge sharing process can be extremely rich and useful for research and technical cooperation IIEP projects. However, it is difficult to retrieve and generally remains unexploited.

In different online courses and knowledge sharing activities, IIEP Program Specialists have tried to retrieve the data by coping and pasting information from the discussion forums. In the past years, Moodle added a function to export data generated in the discussion forums. However, this data is downloaded in a disorganized manner, which hinders its visualization, categorisation, and analysis.

1. **The proposed challenge**

The challenge consists of designing an integrated solution that helps to easily extract, process, analyse and display data from Moodle discussion forums. The solution should allow users to extract data in a structured manner, identify and visualize key terms and concepts and elaborate a summary of the information.

1. **Expected Deliverables**

A prototype of a solution that helps to extract, categorize, analyse and summarize information from discussion forums in Moodle.

The plugin should identify key words in forums and then display the data in a way that we can easily identify the posts which contain the key word, the user who mentioned the word and his/her country. This latest information is part of the user’s profile. If possible, the plugin should also count the number of posts that mention the selected key word and identify users by countries (and if possible regions), displaying the data in a simple dashboard.

This solution should be able to produce:

* A file which contains all the information generated in a forum or selection of forums displayed in a structured manner.
* A visualization of key concepts and terms that come out from the discussions through visualization techniques such as a word cloud.
* A file which contains a summary of discussions

The table below describes the types of functions and filters that the solution could offer:

|  |  |
| --- | --- |
| **Function 1** | User Interface – Control panel – Dashboard:https://powerbi.microsoft.com/en-us/data-dashboards/#:~:text=A%20data%20dashboard%20is%20a,or%20even%20a%20specific%20process. |
| **Function 2** | Connect forum post with participants' profile (name, country, job role) |
| **Function 3** | Present data in a visualized manner based on content |
| **Filter 1: Source of information**   * **Multiple forums from different courses based on a keyword(s) or** * **Multiple forums within one course based on keyword(s) or** * **Multiple forums from different courses based on single user or** * **Multiple forums within a course based on single user** | * Connect data from different forums within different courses based on key words and phrases from all users * Connect data from different forums within one course based on keywords and phrases from all users in the course * Connect data on a specific user from different courses or in a set time spam * Connect data from a single user within a course |
| **Filter 2: User name** | Participants username is normally their email account used to enrol in the platform. This info is in participants' profile. That could be potentially a problem. |
| **Filter 3: Country** | Participants' can state their country in their profile. We just need to be consistent and clarify that we mean the country/workstation. |
| **Filter 4: Date - Time** | Period of time. This is available in the discussion forum (date of post) |
| **Filter 5: Course(s) (title)** | This is available in the MOODLE platform (courses are named and have codes) |
| **Filter 6: Role (ex. Student, teacher, …)** | This is set up in MOODLE. Every user is assigned to one or multiple roles. |

1. **Resources available to address the challenge**

The Hackathon planners will provide access to the IIEP Moodle platform. In addition to the Moodle function to extract information from the discussion forums, some “forum scraping” Python code was developed last year by the IIEP Development team, in collaboration with an intern coder, in order to, process the HTML output of a Moodle forum, and re-output it in an Excel column format, so authors, dates, subjects, etc. could be more easily identified. This work did not however include analysis of the actual text of the forum messages. This code could be further improved to better organize the extracted information.

**6. Points of contact during the hackathon**

Jimena Pereyra: [j.pereyra@iiep.unesco.org](mailto:j.pereyra@iiep.unesco.org)

Jennifer Pye : [j.pye@iiep.unesco.org](mailto:j.pye@iiep.unesco.org)

Jacques Lecavalier : [j.lecavalier@iiep.unesco.org](mailto:j.lecavalier@iiep.unesco.org)

Ioanna Kosteridou : [i.kosteridou@iiep.unesco.org](mailto:i.kosteridou@iiep.unesco.org)

**7. Possible types of analysis that could be done of discussion content**

* Displaying individual participant contributions (with country affiliation) organized by theme reflective of the headings of the framework

(multi-forum)

* Displaying contributions from multiple participants from the same country organized by theme reflective of the headings of the framework

(multi-forum)

* Displaying forum synthesis (single forum) (text or bullets)
* Dashboard/ visual representation of sub-themes from framework ( as part of **demand, quality and supply**, and **sub-themes** of the enabling environment) that have been covered most prominently (focus on module 2 debate on framework) (message classification and possible graphical representation of frequency/prominence of themes)
* Sentiment analysis: overall attitude toward children with disability and inclusive education by participant/ by forum (with focus on Srei Story and Fahma story fora and Module 1 Debate on key concepts related to inclusive education- versions 2 and 3)
* Automatic correction of posts/messages from non native EN or FR speakers.