An online ecosystem for the collection and analysis of learner data

"Improving the quality and sustainability of learning using early intervention methods based on learning analytics"

Project No. 2023-1-FI01-KA220-HED-000159757

**Co-funded by
the European Union**

Project ref. number	2023-1-FI01-KA220-HED-000159757
Project title	**ISILA - Improving the quality and sustainability of learning using early intervention methods based on learning analytics**
Document title	An online ecosystem for the collection and analysis of learner data
Document Type	Deliverable
Document version	1.0.0
Planned date of delivery	2024-06-30
Language	English
Dissemination level	Public
Number of pages	9
Partner responsible	UEF
Author(s)	Sonsoles López Pernas (UEF), Ramy Elmoazen (UEF)
Revised by:	Kamila Misiejuk (UiB), Tanya Yordanova (SU)
Abstract	This report presents the tools developed within the ISILA project. The developed tools support data collection and integration across various Learning Management Systems (LMSs) and external educational resources to enhance institutional capabilities in analyzing student learning behaviors. The core tools include the Enhanced Learning Locker, an open-source Learning Record Store (LRS) that centralizes learning data, and csv2xAPI, which converts tabular data into xAPI-compatible logs for interoperability. Additionally, the data collection tools extend analytics beyond LMS environments: xapi-youtube captures fine-grained interaction data from educational videos, while ChatGPTscraper enables the collection and analysis of students' interactions with conversational AI tools.
Keywords	Learning Analytics, xAPI, Learning Record Store, Data Collection, ChatGPT, Artificial Intelligence, Video Analytics, Higher Education

Table of Content

1 Introduction

The goal of Activity 2.2 is to develop an infrastructure that will be used by all participating HEIs, and that will collect data from a Learning Management System (LMS) and other sources used in teaching. Based on the requirements of each institution highlighted in Report 2.1, we designed the following infrastructure. The main source of data collected by educational institutions takes place in the LMS. The participating HEIs use three types of LMSs: Moodle (UEF, ULe, SU), Canvas (UiB), and LAMS (BMU). The infrastructure developed needs to be compatible with the HEIs LMSs. It needs to be available to easily integrate with institutions other than ISILA HEIs and other data sources beyond the LMS.
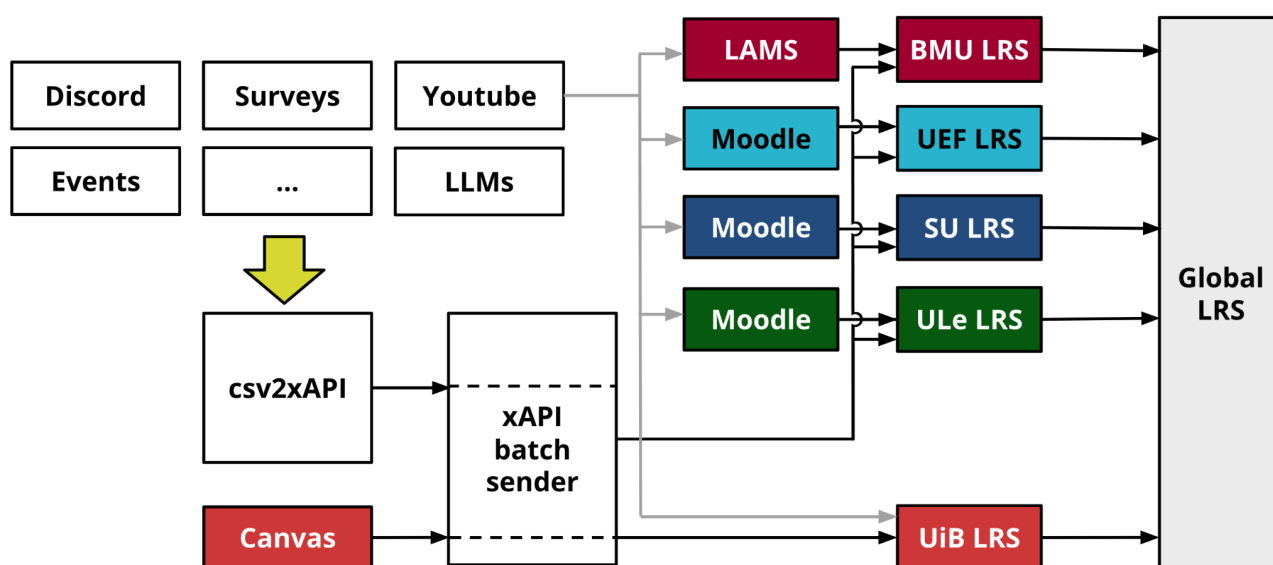

Figure 1 : ISILA infrastructure

In this report, we describe the set of tools developed or adapted to compose the designed infrastructure. We classify the tools into two categories: core tools, and data collection tools.

2 Core Tools

The core tools are designed to facilitate the integration and management of learning data across various systems.

2.1 Enhanced Learning Locker

In order to store the data collected by the different LMSs in a common format that enables a shared analysis, we chose to employ an open-source Learning Record Store (LRS). An open-source LRS, serves as the backbone for storing and retrieving learning records. It is simple to install on the cloud, ensuring that institutions can

quickly set up and start using the platform without extensive technical overhead. Specifically, we chose the Learning Locker LRS[1], one of the most powerful open-source options available, compatible with the xAPI standard. We used existing and newly developed plugins[2] to forward data from the LMSs to the LRS. We modified the existing version of Learning Locker to be able to develop extremely customized multimodal dashboards[3]

2.2 csv2xAPI

To combine data from multiple sources in the scope of the ISILA project, csv2xAPI tool was developed to convert from CSV file (or other tabular format) to xAPI JSON logs that can be sent to a Learning Record Store (Figure 2). The tool is publicly available online[4]. The source code is available on GitHub with an open-source license[5]. A tutorial on how to use the tool is available on YouTube as well[6]. Using the tool requires the following 3 steps:

1. Uploading the data: Users can upload their CSV (or Excel, tsv, etc.) data into the app. If the data is in *wide format*, for example, a survey where each row represents one student with multiple columns representing different responses to different questions, users can choose which columns to pivot so that each row corresponds to an xAPI statement. If the data is already in the *long format*, for example, an event log, where each row represents an interaction and there are —potentially— multiple rows per user, then no additional preprocessing is required.

2. Mapping the data columns to xAPI: Users then map the columns from their data to the xAPI components (actor, verb, object, etc.). If the column names of the dataset match the xAPI parts (timestamp, actor.email, verb.id, object.name, etc.) the mapping is performed automatically.

3. Sending the records to the LRS: Users must provide the details for their LRS (Learning Record Store): endpoint, client and secret. After hitting "Send," the statements are sent to the LRS.

---

[1] Learning Locker, Open Source Documentation: https://learninglocker.atlassian.net/wiki/spaces/DOCS/overview

[2] Logstore xAPI plugin https://github.com/ulerobroticsgroup/logstore-xapi

[3] ISILA version of Learning Locker: https://github.com/aggeorgiev/learninglocker/tree/isila

[4] csv2xAPI: https://sonsoleslp.shinyapps.io/csv2xapi/

[5] csv2xAPI repository: https://github.com/sonsoleslp/csv2xAPI

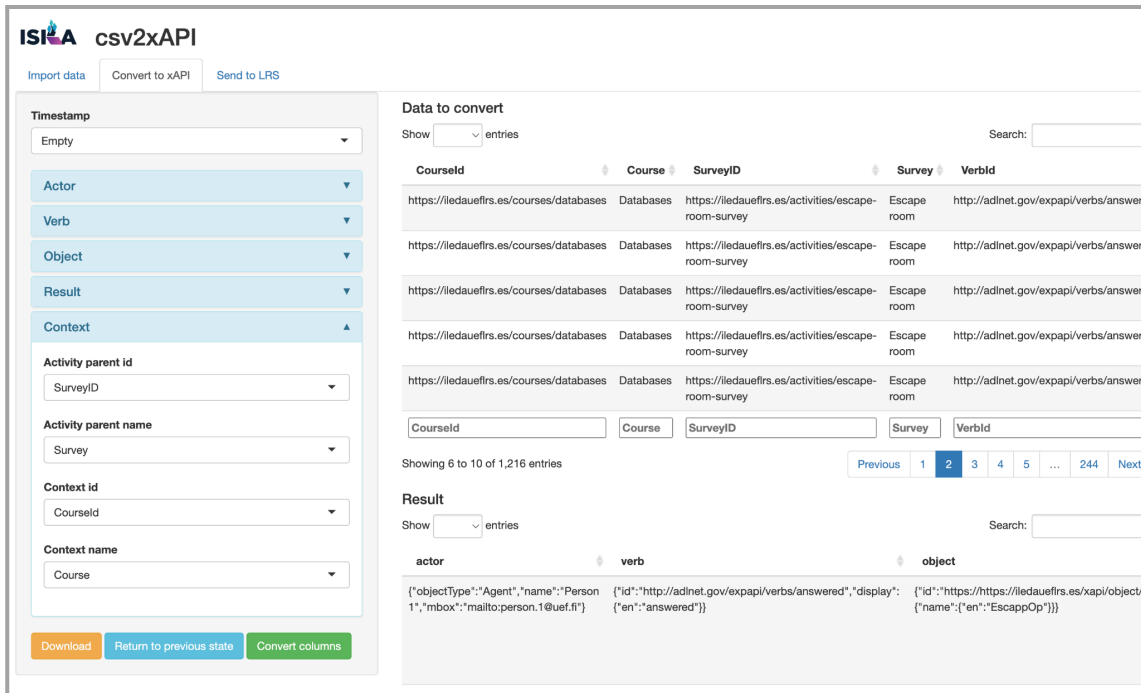[6] csv2xAPI tutorial: https://www.youtube.com/watch?v=OoPYZcQ5YEc

Figure 2 : csv2xAPI Tool

3 Data Collection Tools

The data collection tools are tools used to capture meaningful learning interactions beyond the LMS.

3.1 xapi-youtube

A common problem with LMS data is the granularity of the collected data. For example, if a teacher uploads a video lecture for students to watch, the only data collected is whether students opened the video link or not, but not whether they followed the video for its entire duration, the moments they paused or rewinded, etc.

ADL —the organization in charge of developing the xAPI standard— developed a web application several years ago that allows the capturing of fine-grained data about students when watching YouTube videos and sending them to a LRS using the xAPI format. However, this application was discontinued due to —among other things— some changes in browser restrictions that prevented data from being sent when the window was being closed, and therefore, information on when students stopped watching the video was not collected.

As part of the ISILA project, these technical issues have been fixed by developing a new version of the application that allows to easily embed it into the learning management system (Figure 3).
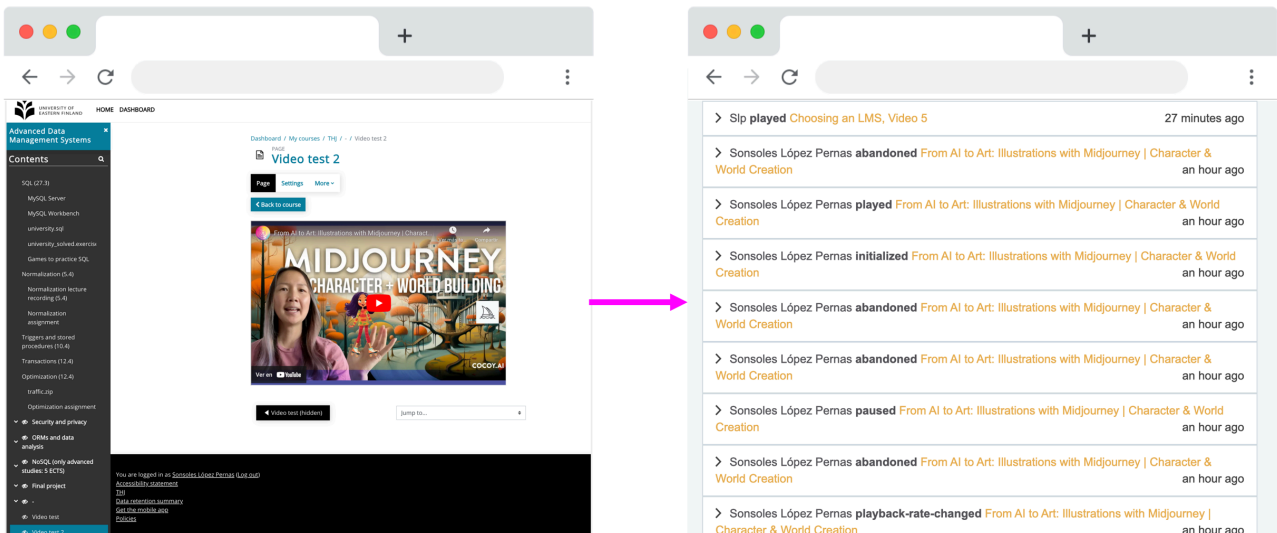
Figure 3 : xapi-youtube Tool

The code of the application can be found on Github[7] and an instance is publicly available online[8], which can be used by passing the user and LRS information on the query parameters. Below is an example of how it can be used from the LMS.

```html
<iframe                                         style="width:700px;height:450px;"id="video-xapi"
src="https://sonsoles.me/xapi-youtube/"></iframe>
<script>
    document.body.onload = function() {
        /* Configuration */
        var endpoint = "YOUR_DOMAIN_WITHOUT_HTTPS/data/xAPI/"; // Your LRS endpoint
        var user = "YOUR_USER"; // Your LRS client user
        var password = "YOUR_PASSWORD"; // Your LRS client secret
        var video = "fK29-99vW48"; // Youtube video code
         var course = "Databases"; // The name of your course
        /* */

        /* Customize how to get student id */
        var student = document.querySelector("a[title='View profile']").innerText;
        /* */

        var iframeSrc = document.getElementById("video-xapi").src +
            "?endpoint=https" + "://" + endpoint +
            "&user=" + user +
            "&password=" + password +
            "&course=" + course +
            "&video=" + video +
            "&student=" + (student);
        document.getElementById("video-xapi").setAttribute("src", iframeSrc);
    };
</script>
```

---

[7] xAPI Youtube ISILA version repository: https://github.com/sonsoleslp/xapi-youtube.
[8] xAPI Youtube ISILA version https://sonsoles.me/xapi-youtube/

3.2 ChatGPTscraper

With the rise of conversational AI tools like ChatGPT, educators now have access to powerful resources that can enrich the learning experience. However, effectively utilizing these tools requires efficient methods for collecting and analyzing interaction data. Traditional approaches —such as relying on APIs or manually retrieving data— can be cumbersome and inefficient. To address these challenges, we are excited to introduce the ChatGPTscraper app —a tool designed to streamline the integration of ChatGPT interactions into learning analytics interventions. The tool is publicly available online[9].

The app allows teachers to download students' interactions with ChatGPT by entering the URLs of the conversations provided by the students. This way teachers can easily analyze the content of the conversations both for research and educational purposes. Below is a brief guide to getting started:

1. **Upload Data**: Input a csv/Excel file containing the URLs of the conversations (Figure 3).
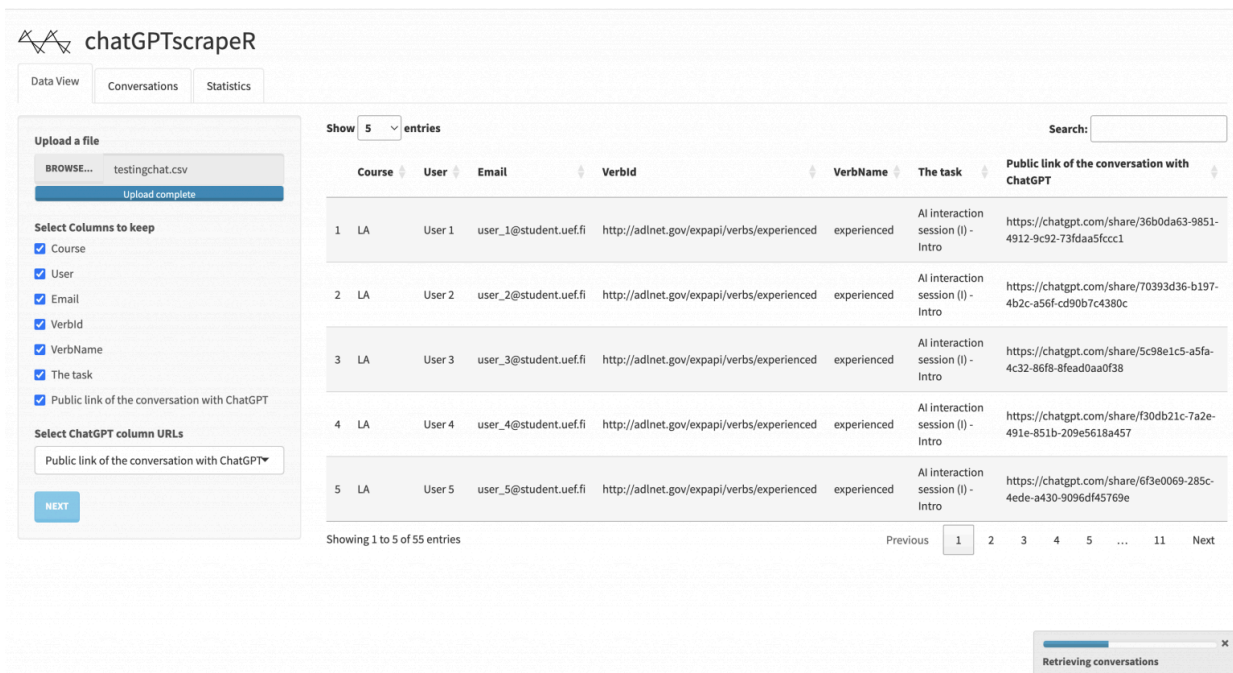

Figure 4 : Upload data to chatGPTscrapeR

2. **Analyze Interactions**: Use the app's interactive tools to view and analyze conversation logs and statistics (Figure 4). You can also download the conversations to analyze them offline or send them to an LRS using our csv2xapi tool.

---

[9] ChatGPTscraper: https://sonsoleslp.shinyapps.io/chatgptscraper

Figure 5 : ChatGPT conversation logs and statistics

The ChatGPTscraper app represents a significant advancement in integrating AI tools into learning analytics. It offers a user-friendly interface that empowers educators to make informed decisions and enhance their teaching practices. This innovative tool aligns perfectly with the ISILA project's mission to improve student engagement and success, making it a valuable addition to any educational toolkit.