



Template for AI characterization

Public in August 2022(updated December 2022)

Author: Jiajun PAN, Azim ROUSSANALY, Anne BOYER

Contents

Abstract	2
1. Background: why we need to propose the template	3
2. Definition of the template: what kind of template do we propose	6
3. Methodology for the questionnaire:how should we design the questionnaire .	13
4. Example of the template: how should we use the template	17
5. Reference	24

Abstract

This document has been prepared in the framework of the European AI4T project. It is one of the deliverable works of WP2.

The purpose of this work is to provide a template that demonstrates the attributes associated with the use of artificial intelligence (AI) technology in educational resources. The template therefore aims to link each resource to a form generated from an instance of the proposed template and to help users understand the characteristics of the resource. In our case, the template is mainly intended for teachers who are selected by WP1 in AI4T and are already aware of the use of AI in the educational resources proposed to them.

For the purpose of this study, an AIER (AI Educational Resource) has a broad meaning and can be an educational platform or a standalone or online application. To investigate these AIERs and represent them in our proposed template, we designed a questionnaire for the creators and companies of AIERs. This questionnaire and its design of it can not only be used in the AI4T project but also be suited to other projects which investigate or evaluated AI educational resources.

This document will be organized as follows:

- In the background section, we present the current status of AI education resources and our reasons for proposing templates. And the focus and limitations of this template are clarified.
- In the definition section, we define a template with 7 layers. For each layer of the template, we introduce its meaning and necessity and give examples.
- In the questionnaire section, we propose a methodology for designing the questionnaire. And we give an example of the questionnaire.

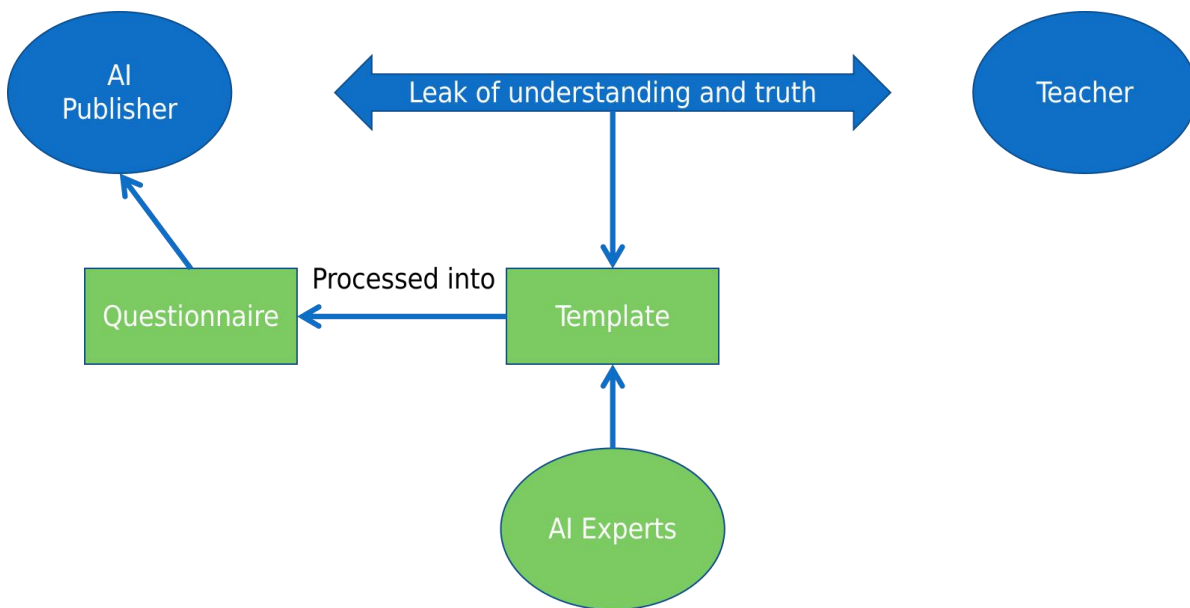
1. Background: why we need to propose the template

Artificial Intelligence in Education (AIEd) is one of the emerging areas in education today. Although it has been around for about 30 years, it is still unclear how it can be fully leveraged to influence teaching and learning and how it can be used to its pedagogical advantage on a larger scale. A survey of K12 education in Europe in 2021 reported [1] that 72% of countries are interested in AIEd, yet most are only beginning to experiment with AI involvement in education or have not even begun to do so. over 47% of learning management tools will be AI-enhanced within the next three years, according to the survey by eLearning Industry [2].

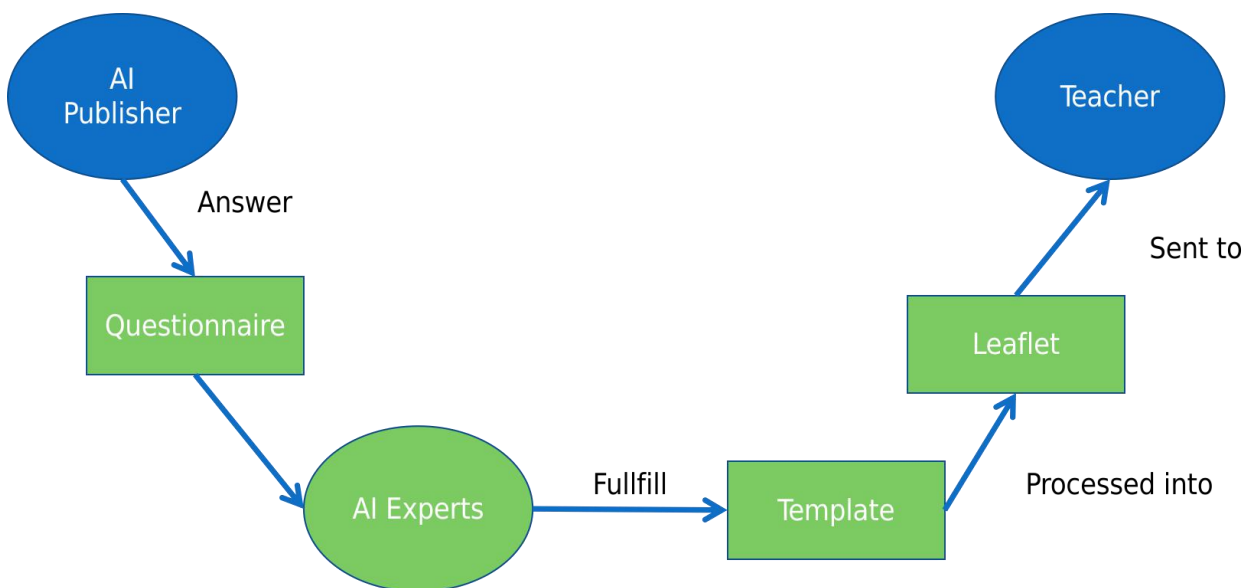
Please note that when we discuss AI in education, we do not include AI that can be used for general tasks in both educational and non-educational scenarios, for example, AI translators that can be used to translate mathematics textbooks and translate non-educational novels. In this report, we focus on AI that is used only for educational tasks. However, the AI technology used in AIER can be not only the AI methods particularly designed for educational tasks but also the common AI methods, we will not limit the definition of AIER by the technology.

AIER is the core of AIEd. Educational resource means the resources used in education, including not only the course, the text teacher designed or used but also the online platform and so on. AIERs are educational resources working with AI technology.

However, in this project AI4T, we have noticed a lack of knowledge and trust in AIER among our main service group, teachers, related groups, students and educational institutions. The existing companies that develop AIERs are also more focused on the technical domain and commercial promotion of AIERs. This results in the lack of an objective and credible standard template that can comprehensively demonstrate the characteristics of AIERs and help users who do not have the relevant knowledge to understand AIERs.



Therefore, we propose a template that focuses on the characteristics of AI integration into educational resources at all levels: scientific, technical, regulatory, ethical, etc. In addition, the information we have compiled is intended to be as objective as possible. It is primarily intended to give teachers a better understanding of the resources they use or that their students use. It is not intended to categorize or compare the solutions proposed by the various resource publishers.



Based on this template, a questionnaire has been designed to gather information about AIER from its publisher. The details of this questionnaire will be shown in the questionnaire section. After getting the answers to the questionnaire, the AI specialist (in the AI4T project, which will be our work in WP2) can fill in the template as a meaningful example. This example can be provided to teachers in the form of a leaflet to make it easier to understand.

Although the primary use of this template is for use in AI4T projects, it is mainly designed for teachers selected in AI4T. However, in the next section, we have also proposed the sources and methods of the design of the layers of the template, so that, following the method of designing this template, it can easily be extended for use in other contexts.

2. Definition of the template: what kind of template do we propose

According to the four-part classification framework of AI technology[3], the four dimensions to classify AI are:

- Context: The environment where the system is being deployed and who is deploying it.
- Data and Input: The data the system uses and the kinds of input it receives.
- AI Model: The underlying particularities that make up the AI system
- Task and Output: The tasks the system performs and the outputs that make up the results of its work.

In the case of AIERs, the basic characteristics to describe it should be the educational environment, the input data from the user (teacher, student or related group), the AI technology (including algorithms) and the educational task (usage scenario).

A survey of AIEd [4] shows a Technology-Organisation-Environment Framework, which denoted AIERs with 3 parts and 1 output. The 3 parts are technology, organization (users and usage), environment (support and input data). The output is the decisions. And the level of decision making that the user can decide in the final output is noteworthy. The usage scenario and decision level of the users have been pointed out again in [2].

In Explainable Artificial Intelligence in education[5], besides the characteristics mentioned before, another dimension is proposed for AIERs which is transparency. To increase trust in AI systems, it is necessary to promote the use of methods that generate transparent explanations and justifications for the decisions made by AI systems.

Another dimension to increase the trust in AI is data security, in [1][6], as for the AI4T project, it can see that concerns about data security are an important source of distrust in AI technology.

Not only to satisfy the requirement of teachers and related users to understand AIERs, but also based on the basic characteristics of AIERs, summarized by the dimensions mentioned before, we design a 3-class 7-layer template. The layers of template are the following:

Usage
Decision level
Technology
Algorithm
Personal data
Data security
Transparency

The first group is the Usage Scenario and the Decision level. This class is to clarify the purpose and users of an AIER and to help teachers understanding what their role is in using this AIER. The core questions to ask in these layers are as follows:

Usage	Who is the resource for?
	What are the purposes of the resource?

Usage scenarios layer describes the users and applications of the AIER, which is often the first information teachers want to know. AIER could be classified into 3 types: Learner-oriented AI, Instructor-oriented AI and Institutional system-oriented AI[1]. There are many different usage scenarios based on the type of AIER. For example, a learner-oriented AIER could be an intelligent computer system designed to help students learn a course, while a teacher-oriented AIER could be one that automatically generates questions and tests for teacher-prepared courseware for a particular course.

Decision level	What kind of the output from the AI? Is it a decision, recommendation or propose?
	Who bears the responsibility for the decision? Could he change or modify the decision? Could teacher change or modify the decision?
	Is teacher take part of this decision? How could he work on that?

Decision level layer describes the proportion of the final outcome that is attributable to the faculty when teacher using AIER. Normally, the decision level depends on the output of the AIER. For example, for AI that uses scoring strategies to automatically assess student learning outcomes, the level of decision making depends on whether the scores are sent directly to the student or to the teacher and who makes the subsequent assessment results. Decision level help teacher know their right and role in using an AIER.

The second group is the Technology and Algorithm one. This class is to clarify the approaches the AIER used to reach its purpose and to help teachers understand how it works. The core questions to ask in these layers are as follows:

Technology	What is the issue addressed by the AI? Or what is the educational task the AI support?
	The result of AI technology?

Technology layer describes the AI technologies used in the AIER. Considering that an AIER may be an educational platform, a standalone or online application, etc., several different technologies may be used in a single resource. The purpose of this layer is to clarify whether AI-based technologies are actually being used and to introduce teachers to what kind of AI technologies are being used. The technology could be one of the following list:

- ◆ Automatic generation of educational content(Courses, texts and so on).

- ◆ Improved educational content
- ◆ AI assistance for teachers to provide personalized instruction for each student.
- ◆ Fast feedback to students.
- ◆ Assistance in monitoring students.
- ◆ Automatic assesses students' learning behaviour and learning routes.
- ◆ Automatic records the student's learning process.
- ◆ Targeted improvement based on learning analysis.
- ◆ Chatbot between teachers, students, parents and relevant groups.
- ◆ Other technology for educational tasks.

Algorithm	Which family of approaches do the algorithms related?
	What are the algorithms?
	Is open source AI software or resources used in this resource? If so, what is it?

Algorithm layer describes the algorithms used in the AIER. Compared to the previous layer, this layer focuses more on the scientific level than on the technical level. We will declare the family of the algorithms from three common group: Knowledge-based Systems, Machine Learning and Deep Learning. Note that we are aware of the debate about whether deep learning is machine learning or not. However, we have separated them here because it is easier for teachers to understand. For each group, the algorithm could be selected from the following list:

- ◆ Knowledge-based Systems: Rule-based systems (or expert systems), Ontology, Semantic networks and so on.

- ◆ Machine Learning: Clustering, Approximate possibility, Regression Analysis, Representation and Dimensionality reduction, Active learning, Decision Trees, and so on.
- ◆ Deep Learning: Convolutional Networks (CNNs), Long Short Term Memory Networks (LSTMs), Recurrent Neural Network (RNN), Generative Adversarial Networks (GANs).

Please note that we will only identify the algorithms used in this layer and briefly introduce the algorithms, without comparing and evaluating the algorithms used.

The third group is Personal data, Security and Transparency. This class is to clarify the input data used by AIER and the secure handling of that data, to make it clear that AIER demonstrates the transparency of the internal mechanisms and to help teachers trust in the security of AIER. The core questions to ask in these layers are as follows:

Personal data	What kind of personal data is used?
----------------------	-------------------------------------

Personal data layer list the personal data used in the AIER. Nowadays the information security is increasingly valued. In addition to the data explicitly proposed for input in the AIER, external sources, such as shared social accounts, may collect additional personal information. The creation of this layer is necessary to enable teachers to clarify this information.

Data security	Is the personal data anonymous? If yes, how it works?
	How does the resource under the consent under local regular?
	What are the possibilities for outsiders to audit the resource?

The data security layer specifies whether non-user visitors and external visitors to AIER can access the data and how it can be handled securely. This layer is related to

the previous one. Different personal data should be handled in different ways, such as anonymity, encryption, and denial of access.

Anonymization means that the information recipient is completely unable to identify the data subject directly or briefly. There are five common data anonymization operations: generalization, suppression, anatomization, permutation, and perturbation[7]. Whereas encryption or also known as pseudonymization means that the information receiver cannot identify the data subject directly, but the information can be de-anonymized by the data manager after cross-identification with other information. Denial of access is a simple and straightforward way to keep the information only in the information manager, without giving access to others.

In this layer, it also clarifies whether the data will be used by external visitors. Many software share data with their own data partners nowadays, and if this is the case, it can also lead to different security of the data.

Transparency	What is the outcome or end service provided by this resource?
	In what way is this result presented to the user?
	Is there an explanation of the AI technology and algorithms used? If so, are there any relevant published papers?
	Is there a website or other public way to explain how this resource works? If so, what is it?

The transparency layer clarifies the level of understanding of how the internals of AIER work. This level is more dependent on the technical layer and the educational tasks that the AIER is intended to accomplish. For example, if the educational task is to assess students through their learning behaviors, the internal mechanism of this AIER is about which learning behaviors were selected and which parameters were used to generate the final report. If the final assessment report given to the teacher that shows these parameters or if the AIER's introductory website explains the internal mechanism of this AIER, it will help teachers understand the possibilities, limitations, and risks of this AIER in education.

In addition to this, the transparency to teachers of the intermediate results generated by AIERs in the educational tasks is worth declaring at this layer. For example, an AIER delivering personalized instruction may have to first generate a profile of each student that it uses to recommend different educational content in the process of accomplishing this task. If the profiles is available for teachers, it will help teachers understand and trust this AIER.

Overall, this 3-class 7-layer template can describe AIERs by introducing and charity the basic features of AIERs and the dimensions of interest to those who use AIERs in education (e.g. teachers, etc.), more importantly, this template can be targeted to increase user trust.

It is worth noting that the primary use of this template is for AI4T projects, designed to be teacher-facing. As such, we have discarded some elements that are less important or redundant for an educational environment, such as the operating environment, operation and maintenance staff, support for requirements, etc. However, following the methodology used to propose this template, it could easily be extended to a template for describing AIERs in other situations or other AI resources for use in a wider range of situations.

3. Methodology for the questionnaire: how should we design the questionnaire

In order to fill in the template mentioned above, a questionnaire was designed for the providers of AIERs. Note that, though the questionnaire’s purpose is to create examples of template for teachers and pedagogical experts, the questionnaire is fulfilled by the designers to investigate the characteristics of the AIER. Therefore, the design and order of the questions in the questionnaire are from the view of the AIER designer, rather than following the order of our proposed template.

In the questionnaire, we used colors to express the correspondence between the questionnaire questions and the various layers of the templates. After obtaining the completed questionnaire, we will fill in the examples of the template according to the correspondence.

Note that, the options and full details of the questionnaire is not shown in this report, you can test it by the link:

<https://sondages.inria.fr/index.php/282721?lang=en>

The questions in the questionnaire were designed into five groups, ranging from easy to answer to more involved details:

Questionnaire of the template for AI characterization	
1. What is this resource?	Name of the resource
	Short description of the resource
	Information of the publisher of this resource
	Is there a public information website about the resource? If so, what is the url?

	Is there any related public papers about the resource?
	If so, what are the related publics?
2.What is this resource used for?	Who is this resource for?
	What is the purpose of the resource for the user?
3.How does this resource achieve its purpose?	What functions are artificial intelligence technologies used in this resource?
	What are the algorithms used in this resource?(There is a list we used for options in the end of this section)
	Are there any scientific publications on the algorithms implemented?If so, what are the related publics?
	Is open source AI software or resources used in this resource? If so, what is it?
4.What does the resource require as input?	What type of data does the resource need to collect?
	Is additional personal data collected (for creating login accounts or other functions) in addition to the data supporting the resource? If so, what is it?
	What is the general level of data security processing for the used personal data?
	What is the method of secure data processing used?
	Is data from external sources (cookies, social accounts) used? If so, what is it?
	Is any additional data (the environment in which the resource is running, additional parameters that need to be entered, etc.) collected



	besides personal data? If so, what is it?
	How does the resource obtain consent under local regulations?
	Is the collected data shared with partners?
5.What is this resource output?	What is the final output of this resource?
	What is the final result of the AI technology or algorithms of this resource?
	What type of end result does this belong to?
	For this result, who has direct access? Who can make adoption or modification?
	What is the presentation of this output?
	If the result is quantitative outputs or predictive models, for this output, what parameters (student behavior, test scores, etc.) are used to calculate the results?
	Are these parameters presented or accessible to the user?
	How does the work of the teacher interact with this output?
	Is there a web page or other way to show more explanation of how this resource gets such results? If yes, what is the url of this website?

List of algorithms	
Knowledge-based	Rule-based systems (or expert systems)



Systems	Ontology
	Semantic networks
Machine Learning	Clustering(K-Means, Spectral Clustering, Hierarchical Clustering, etc.)
	Approximate possibility(Markov Process, Markov Random Fields, Expectation Maximisation, etc.)
	Regression Analysis(Linear Regression, Logistic Regression, etc.)
	Representation and Dimensionality reduction(Metric Learning, Principal Component Analysis (PCA), Principal Component Analysis (PCA), etc.)
	Active learning(Neural Network (ANN), Support Vector Machine (SVM), etc.)
	Decision Trees, Random Forests, etc.
	Reinforcement Learning (Q-Learning, SARSA, Armed Bandit, etc.)
Deep Learning	Convolutional Networks (CNNs)
	Long Short Term Memory Networks (LSTMs)
	Recurrent Neural Network (RNN)
	Generative Adversarial Networks (GANs)



4. Example of the template: how should we use the template

Here are some examples we get:

Basic Information	Name of the resource	Duolingo for Schools
	Short description	Duolingo for Schools is a dashboard right inside of a teacher’s Duolingo account that allows them to create classrooms and assignments, and to follow students’ activity. We’re excited to work with educators to bring the world’s number one language learning app into classrooms, with features designed to maximize teacher effectiveness and student learning.
Usage	Who is the resource for?	Doulingo for Schools is a learner-oriented AI
	What are the purposes of the resource?	With Duolingo for Schools, teachers can create one or more classrooms and either add students, or allow their students to join themselves. Teachers can manage student access to course content and track student progress throughout their learning journey by tracking student activity, including XP earned and time spent working on a course, as well as individual student level activities, such as lessons completed and scores earned
Decision level	What kind of the output from the AI? Is it a decision, recommendation, propose or service?	Adapt learning pathway of students or of group of student when teachers assign them tasks via a dashboard It is a servise.
	Who bears the responsibility for	Decision to give a task or assessment to a learner is taken by the teacher.Elements from the dashboard



	<p>the decision? Could he change or modify the decision? Could teacher change or modify the decision?</p>	<p>(skills /levels / lessons) are not directly transformed into grades, teacher could modify part of them.</p>
	<p>Is teacher take part of this decision? How could he work on that?</p>	<p>Teachers can propose two types of assignments to learners, the first is to collect experience points (or XP) and the second one is to level up a skill.</p>
Technology	<p>What is the issue addressed by the AI? Or what is the educational task the AI support?</p>	<p>A dashboard right inside of a teacher's Duolingo account that allows them to create classrooms and assignments, and to follow students' activities.</p>
	<p>What is the result of this AI technology?</p>	<p>Analytics report</p>
Algorithm	<p>Which family of approaches do the algorithms related?</p>	<p>Machine Learning</p>
	<p>What are the algorithms?</p>	<p>Sleeping recovering Bandit = for notifications on duolingo https://research.duolingo.com/papers/yancey.kdd20.pdf Bayesian Knowledge training =for knowledge tracing (estimating a student's current knowledge state and predicting his/her performance in future interactions, students can receive personalised learning materials https://sharedtask.duolingo.com/papers/chen.slam18.pdf</p>





	Is open source AI software or resources used in this resource? If so, what is it?	Unknown
Personal data	What kind of personal data is used?	Name, Contact information: email address for creating account.
Data security	Is the personal data anonymous? If yes, how it works?	Anonymized Allows the creation of profiles without name, surname or email addresses; Teachers can also create accounts for their students without using name or email address.
	How does the resource under the consent under local regular?	No consent procedure(UNknown)
	What are the possibilities for outsiders to audit the resource?	Unknown
Transparency	What is the outcome or end service provided by this resource? In what way is this result presented to the user?	A dashboard right inside of a teacher's Duolingo account that allows them to create classrooms and assignments, and to follow students' activities. It is presented as a dashboard with report of students' behavior.
	Is there an explanation of the	No



	AI technology and algorithms used? If so, are there any relevant published papers?	
	Is there a website or other public way to explain how this resource works? If so, what is it?	Incidentally many scientific articles are published by the Duolingo teams in scientific publication to explain how some elements of the model work : https://research.duolingo.com/

Basic Information	Name of the resource	BacAnalytics
	Short description	Tools for organizing baccaulaureat exams
Usage	Who is the resource for?	Educational Institutions
	What are the purposes of the resource?	Define and assess students' learning behavior, motivation, attention and engagement. This tool is used by the academic staff for the second group of exams of the french baccalauréat when learners choose 2 subject for oral sessions
Decision level	What kind of the output from the AI? Is it a decision, recommendation, propose or service?	Score for each possible choice (subject) Recommendation of teachers assignement for each exam center and for each subject

	<p>Who bears the responsibility for the decision?</p> <p>Could he change or modify the decision?</p> <p>Could teacher change or modify the decision?</p>	Educational Institutions bear the responsibility.
	<p>Is teacher take part of this decision?</p> <p>How could he work on that?</p>	No
Technology	<p>What is the issue addressed by the AI?</p> <p>Or what is the educational task the AI support?</p>	Ranking of estimated best choices of subjects for each learner
	<p>What is the result of this AI technology?</p>	Classification / Ranking
Algorithm	<p>Which family of approaches do the algorithms related?</p>	Machine Learning
	<p>What are the algorithms?</p>	Neural Network, Decision Trees Base on Mahartta's paper
	<p>Is open source AI software or resources used in this resource?</p> <p>If so, what is it?</p>	Use Java WEKA API
Personal data	<p>What kind of personal data is used?</p>	<p>Address information: place where they studied;</p> <p>Educational background information: marks, options;</p> <p>Contextual data : organization information + teachers assignment rules</p>
Data security	<p>Is the personal data anonymous? If</p>	<p>Anonymized</p> <p>Identities informations are replaced</p>

	yes, how it works?	by an anonymized ID
	How does the resource under the consent under local regular?	No consent procedure
	What are the possibilities for outsiders to audit the resource?	No
Transparency	What is the outcome or end service provided by this resource? In what way is this result presented to the user?	Recommendation of teachers assignment for each exam center and for each subject. It is peresented as score for each possible choice (subject).
	Is there an explanation of the AI technology and algorithms used? If so, are there any relevant published papers?	Mahartta's paper
	Is there a website or other public way to explain how this resource works? If so, what is it?	No

Basic Information	Name of the resource	CNED Project
	Short description	Early warning system for at risk students detection
Usage	Who is the resource for?	Teacher
	What are the purposes of the resource?	Identify and address topic with struggling students or inadequate feedback It is used by teachers for getting alerts concerning at risk student. The tool could be used every week.

Decision level	<p>What kind of the output from the AI? Is it a decision, recommendation, propose or service?</p>	<p>Score for each possible choice (subject) Recommendation of teachers assignment for each exam center and for each subject</p>
	<p>Who bears the responsibility for the decision? Could he change or modify the decision? Could teacher change or modify the decision?</p>	<p>Teacher bears the responsibility. The classification could not be changed but only teacher know the score.</p>
	<p>Is teacher take part of this decision? How could he work on that?</p>	<p>No</p>
Technology	<p>What is the issue addressed by the AI? Or what is the educational task the AI support?</p>	<p>Classification : 1) at risk, 2) medium risk, 3)no risk possibly o) : drop out</p>
	<p>What is the result of this AI technology?</p>	<p>Classification / Ranking</p>
Algorithm	<p>Which family of approaches do the algorithms related?</p>	<p>Machine Learning</p>
	<p>What are the algorithms?</p>	<p>Clustering,Active learning,Decision Trees 5 articles of Amal</p>
	<p>Is open source AI software or resources used in this resource? If so, what is it?</p>	<p>Open source Python AI libraries</p>
Personal data	<p>What kind of personal data is used?</p>	<p>Personal physical information:gender, age; Comment:marks, options;</p>

		Contextual data : organization information + teachers assignement rules
Data security	Is the personal data anonymous? If yes, how it works?	Anonymized Identities informations are replaced by an anonymized ID
	How does the resource under the consent under local regular?	No consent procedure
	What are the possibilities for outsiders to audit the resource?	No
Transparency	What is the outcome or end service provided by this resource? In what way is this result presented to the user?	Recommendation of teachers assignement for each exam center and for each subject. It is peresented as score for each possible choice (subject).
	Is there an explanation of the AI technology and algorithms used? If so, are there any relevant published papers?	5 articles of Amal
	Is there a website or other public way to explain how this resource works? If so, what is it?	No

5. Reference

[1] European Schoolnet . “Artificial Intelligence Role in K12 Education: Agile Collection of Information”, 2021, Brussels, Belgium.

[2] Amit Gautam “Role Of Artificial Intelligence In Shaping The Future Workforce”, July 7, 2019,eLearning Industry, <https://elearningindustry.com>

- [3] Karine Perset, Dewey Murdick, Jack Clark, Marko Grobelink “A first look at the OECD’s Framework for the Classification of AI Systems, designed to give policymakers clarity”, November 24, 2020, <https://oecd.ai/en/>
- [4] Tahiru, Fati. "AI in education: A systematic literature review." *Journal of Cases on Information Technology (JCIT)* 23.1 (2021): 1-20.
- [5] Khosravi, Hassan, et al. "Explainable artificial intelligence in education." *Computers and Education: Artificial Intelligence* 3 (2022): 100074.
- [6] Sarah Hunt. “How AI is Being Used in Education”, October 10, 2021, <https://www.datamation.com/>
- [7] Eyupoglu C, Aydin M A, Zaim A H, et al. An efficient big data anonymization algorithm based on chaos and perturbation techniques[J]. *Entropy*, 2018, 20(5): 373.