Fostering the Creation of Personalized Content for Cultural Visits

Riccardo De Benedictis, Carlo De Medio, Augusto Palombini, Gabriella Cortellessa, Carla Limongelli and Amedeo Cesta

Abstract: Among more dramatic effects, the COVID-19 scenario also raised the need for new online information and communication services, promoting the spread of software solutions whose usefulness will last well beyond the pandemic situation. Particularly in the cultural heritage domain, it has been unveiled the relevance of new AI-based approaches, able to dynamically aggregate information and making them available for a customized fruition aimed to the individual cultural growth. Here, we integrate machine learning techniques for the automatic generation of contents for an intelligent tutoring system grounded on automated planning techniques. We present a solution for semantic, intelligent creation of personalized cultural contents, born as a lesson-making assistant, but developed as to become a multi-function “cultural crossover”, useful in the frame of a wide range of planning, dissemination, and managing activities for cultural heritage contents.

Keywords: artificial intelligence; educational technology; personalization; AI modules integration; cultural heritage

1. Introduction

Cultural heritage, cities of art and, in general, cultural experiences can contribute to the development of creativity also in learning processes. For this reason, cultural heritage represents a favorable context for stimulating and enriching one’s knowledge, being a stimulating starting point for learning.

This opportunity appears particularly critical in the current period, when we are living with the crisis caused by the COVID-19 health emergency, which has changed our habits also by limiting physical access to cultural heritage.

This change highlights, on the one hand, the need to develop alternative forms of learning and fruition, and, on the other, the great opportunities provided by technology.

One of the useful tools in this context is represented by technology enhanced learning (TEL), which are systems based on the use of technology to support and improve learning processes [1–3]. TEL environments offer the possibility to access a wide range of materials, contents and information, and to promote an active participation of students in building their own knowledge [4,5]. TEL systems offer valid opportunities and alternatives to classic learning methods, although still an open challenge exists there that is linked to the authoring problem to create content of lessons with estimates of 200–300 h of development per hour of fruition [6].

This work is situated in the above context, describing the latest efforts to integrate and strengthen the past activities of the authors aimed at providing, on the one hand,
dynamic adaptation services and personalized fruition of learning experiences and, on the other, support services to the editors of the multimedia contents that enrich and improve the learning offer. Specifically, we started with a system, called EXPLORAA [7], aiming at supporting experiential learning through the use of automated planning. Specifically EXPLORAA allows generating a collection of "stimuli", called learning objects, aiming to some didactic objectives that are, simultaneously, personalized to the profile of the user and adapted according to their current and dynamically evolving state. The reason to rely on automated planning [8], in this context, stems from the opportunity it offers to model the user’s profile as the initial state of a planning problem and the didactic objectives as goals. A specific form of automated planning technology, called timeline-based [9] is used that, thanks to its flexibility, enables the dynamic adaptation of the learning objects to the changes in the user’s state that may occur during the fruition of a specific session (plan adaptation).

Although effective in addressing the problem of planning personalized lessons and dynamically adapting them at execution time, EXPLORAA suffers from the problem of creating and organizing content, adopting a manual form of editing and, as such, extremely slow and prone to errors [10]. Thanks to the initial integration with an authoring tool, called Wiki course builder (WCB) [11,12], we began, at a later stage, to tackle the editing problem introducing easier methods, relying on forms of semantic analysis of Wikipedia pages, for building the content, from which to generate personalized lessons [13]. This work aims to strengthen the content generation and lesson planning phases, rethinking and enriching the modeling of the users, so as to further support the creation of personalized content, and further integrating the EXPLORAA and WCB systems addressing, more homogeneously, the use of personalized contents and their creation and organization. The joint use of different AI technologies, indeed, ranging from semantic reasoning and automated planning to machine learning techniques, allows the implementation of the aforementioned services, resulting in the definition of a flexible tool for managing learning materials, in this specific case focused on the cultural heritage domain, and making them available, in “customized packages”, to different types of users, with their own interests and preferences.

This work is organized as follows: Section 2 introduces some related works with the aim of better contextualizing our framework; Section 3 describes some basic concepts underlying the EXPLORAA system; Section 4 describes our most recent efforts to support the editing process through intelligent techniques, thus leading to the creation of a complete and integrated system, called WIKITEL, for the definition and use of personalized lessons; Section 5 introduces our choices for user modeling and how such choices relate to the capabilities of supporting the definition of the content for personalized lessons; although the proposed system is general and, therefore, domain independent, Section 6 describes how we contextualized the framework to the case of the city of Matera, showing some benchmarking results; finally, in Section 8 we discuss some conclusions and possible future system extensions.

2. Related Works

Many works in the cultural heritage domain mainly focus on the faithful reproduction of the contents [14–17], allowing extremely realistic virtual tours while tackling, among other things, accessibility aspects, ensuring access for people with mobility issues. The present work focuses on the problem of generating and planning personalized content and building lessons for students that are tailored to specific interests. Personalization can occur at different levels, having different influence on the effectiveness of use. The work presented in [18], for example, identifies three types of customization: adaptability (also called customization), that offers the user the possibility of setting options so that the system adopts a desired behavior; context-awareness that is the system’s ability to perceive dynamic changes in the environment and adapt accordingly to them; adaptivity that is the system’s ability to adapt its behavior by inserting, for example, additional content. The
adaptivity property, in particular, is the most complex to implement, requiring more effort from the author. Therefore, the greater the ability of the system to customize, the greater the difficulty for the content author to specify the customization rules.

WCB proposes a methodology which helps the content author to quickly configure the course material, based upon Wikipedia pages. Many studies have been performed both on how Wikipedia is used by students and whether it is a valid learning support. The study carried out in [19] shows that students are quite satisfied with Wikipedia because they look up what they find, although they are cautious since there is no full guarantee of the accuracy of the information. After more than ten years, this approach seems to be established, both for students and anyone searching for information quickly and satisfactorily. In 2016, the vast majority of university students were using Wikipedia [20], while the use of Wikipedia still encounters resistance from lecturers who consequently negatively influence students’ approach to this instrument. What we propose in this paper is the use of Wikipedia to help course creators quickly configure courses according to different learning objectives. WikiTEL provides an authoring system that allows the integration with additional material selected by the teacher, making the configuration of learning paths fast and flexible.

Fidas et al. [21], as an example, describe some of the existing editing tools, oriented towards the end users, which are based on mobile applications. It is worth noticing that, although such tools have their own peculiarities, they share the idea of being mass oriented, giving anyone the opportunity to generate packages of contents. If, on the one side, these systems have the potentiality to allow easy generation of content, distributing the work to potentially numerous users, on the other side, such a high ease of use, necessary to support extensive targeting, conflicts with the possibility of defining sophisticated customization rules.

The approach pursued in Porteous et al.’s work [22,23], on the contrary, tries to maintain a high expressive capacity while supporting the author through the automatic or semi-automatic generation of the rules. By recognizing the difficulty of generating planning domains robust to dynamic changes, such as those required by interactive narrative, in particular, the mentioned works face the problem of proposing, to the content authors, possible extensions to existing domains, hence increasing their ability to manage unexpected events that can occur during the narration. By taking advantage of linguistic hierarchies, such as WordNet and ConceptNet, in particular, the system exploits hyponymy and hypernymy to propose to the authors the introduction of new types or, through antinomy, the introduction of new planning actions.

This article, in line with the latter works, tries to contribute to this research direction by keeping the high personalization expressiveness and complexity of the technology enhanced learning systems, integrating parts of the WCB system, for supporting the lesson editing with the automatic generation of the customization rules, with EXPLORAA, for the generation and the dispatching of highly personalized lessons.

3. What Is EXPLORAA?

The work described in this paper starts from a continuous intelligent tutoring system, called EXPLORAA [7], developed in the context of “Città Educatore” project (http://www. cittaeducante.it, accessed on 1 August 2021) (the name means “city that educates” in Italian). The original system, evolving from previous experiences (e.g., [24–26]), by relying on the idea of dynamically composing lessons through the use of automated planning, pursued the idea of providing a continuous and dynamic learning experience in both space and time. By guaranteeing flexible generation of interactive plots, automated planning represents a valid tool for supporting narrative [27]. Thanks to the use of automated planning, specifically, EXPLORAA allows to compose and dynamically execute personalized lessons. By starting from a static representation containing a high-level lesson track, intuitively, the customized lesson is planned and dynamically adapted and further personalized to the involved users. The use of the automated planning technology allows the creation a sufficiently extensive didactic experience, guaranteeing the reproduction of a large number of different situations.
which are, at the same time, characterized by a high variability of learning objects so as to increase the involvement level of the involved users. More specifically, the chosen automated planning technology, called timeline-based [9], represents the unifying element of the various modules, favoring the generation of different lessons that would be too complicated to obtain with a simple pre-compilation of stories while ensuring the dynamic adaptability of plans by promoting experiential learning.

From a high-level perspective, the general idea of the EXPLORAA system is described in Figure 1. It is possible to distinguish, from the figure, two classes of involved users: the students, i.e., a group of people, potentially, of any age, interested in using the learning services offered by the environment, and the teachers, i.e., users, with special privileges over the system, who have the opportunity to observe students, monitor and control the progress of the single educational experience and of the overall learning environment. Once the authoring phase is finished, the underlying timeline-based planner, through the planning process, takes care of keeping the profile of the involved users consistent with the lesson model, resulting in a solution (i.e., a plan) which represents a lesson customized to the profiles of the involved students. The arising plan will then be executed, bringing to the dispatching of learning objects to a lesson presentation module, consisting of a Web-based application and of an Android-based application. It is worth noting that the mobile application is also responsible for collecting information from the users, such as customization parameters, aimed at dynamically modifying the user’s state and, consequently, the content of the presented lesson. In particular, the application collects a set of user parameters, gathered through: (i) GPS/Wi-Fi/mobile networks for localization and geographical information, (ii) Bluetooth bracelets for psychological parameters, and (iii) questions administered to the users for any additional useful information (for a more detailed description of the EXPLORAA system refer, for example, to [25]).

![Figure 1. The EXPLORAA general idea.](image)

Particularly interesting, in this context, is the peculiar automated planning technology used for orchestrating the educational experience: timeline-based planning. The choice of timeline-based technology, specifically, was dictated by the need to reason on temporal aspects, requesting to dispatch information at proper times, but also by the need to dynamically react, through plan adaptation procedures, to uncontrollable changes coming from the users and, more generally, from the external environment. Timeline-based planning strongly relies on the concepts of rules and constraints, hence the use of the term “consistent” for the generation of a personalized learning object plan, and represents the unifying element of the various intelligent modules by ensuring the dynamic adaptability of plans with the aim to promote experiential learning.

The main data structure for timeline-based planning is, nomen omen, the timeline. In generic terms, a timeline is a function of time over a finite domain. Values on the timelines are extracted from a set of “assertions over a temporal intervals” called tokens.
Formally, a token is an expression of the form \( n(x_0, \ldots, x_i)@[s, e, \tau] \), where \( n \) is a predicate name, \( x_0, \ldots, x_i \) are the predicate's parameters (i.e., constants, numeric variables or object variables), \( s \) and \( e \) are temporal parameters belonging to the temporal domain \( \mathbb{T} \) (either real or discrete), such that \( s \leq e \) and \( \tau \) is a parameter (i.e., a constant or an object variable) representing the timeline on which the token apply. Roughly speaking, a token states that the assertion on the left of the "@" symbol holds from time \( s \) (start) to time \( e \) (end) on the timeline \( \tau \).

It is worth noticing that the tokens' parameters, including the temporal ones, are constituted, in general, by the variables of a constraint network \([28,29]\) and can hence be constrained so as to reduce their allowed values and bring the system to a desired behavior placing, for example, the tokens over the desired times. Moreover, those temporal overlapping tokens which assume the same value for their \( \tau \) variable undergo a further form of reasoning (i.e., scheduling) that depends on the value assigned to the \( \tau \) variable. This further form of reasoning introduces additional constraints aiming at completely avoiding the overlaps, in case the assumed value represents a state-variable, or allowing a limited number of overlaps which guarantees the compliance with the capacity constraint, in case the assumed value represents a reusable-resource.

The idea, pursued in EXPLORAA, consists in using such tokens for representing, in a whole, the profiles of the users, the pedagogical objectives and the planned learning objects. Coordination among these aspects, however, takes place through the concept of rule. Specifically, tokens are partitioned into two groups: facts and goals. The semantics, in this case, is borrowed from constraint logic programming \([30]\). Specifically, while the facts are considered inherently true, the goals must be achieved. The achievement of a goal can take place either through a unification with either a fact or another already achieved goal with the same predicate (i.e., the parameters of the current goal and the token with which is unifying are constrained to be pairwise equal), or through the application of a rule. Formally,

**Definition 1.** A rule is an expression of the form

\[
    n(x_0, \ldots, x_k)@[s, e, \tau] \leftarrow r
\]

where:

- \( n(x_0, \ldots, x_k)@[s, e, \tau] \) is the head of the rule, i.e., an expression in which \( n \) is a predicate name, \( x_0, \ldots, x_k \) are numeric variables or object variables, \( s \) and \( e \) are temporal variables, such that \( s \leq e \) representing the starting and the ending time of the assertion and \( \tau \) is an object variable representing the timeline on which the assertion apply;
- \( r \) is the body of the rule (or the requirement), i.e., either another token, a constraint among tokens (possibly including the \( x_0 \ldots x_k \) variables), a conjunction of requirements or a disjunction of requirements.

Rules define the causal relations that must be complied to in order for a given goal to be achieved. Roughly speaking, for each goal having the "form" of the head of a rule, the body of the rule (i.e., further tokens, constraints, conjunctions, and disjunctions) must also be asserted.

Finally, a timeline-based planning problem is a triple \( \mathcal{P} = (\mathcal{O}, \mathcal{R}, r) \), where \( \mathcal{O} \) is a set of typed objects, needed for instantiating the domains of the constraint network variables, \( \mathcal{R} \) is a set of rules and \( r \) is a requirement, i.e., either a token (whether a fact or goal), a constraint among tokens, a conjunction of requirements or a disjunction of requirements. In simple terms, starting from an initial partial plan containing the problem's requirement, the planning resolution procedure will deal with either applying the rules, in order to guarantee the achievement of goals (note that this process can introduce new (sub)goals into the current partial plan), or will demonstrate the goals' semantic equivalence, called unification, with other goals already achieved or with other facts (refer to \([31]\) for further details).
Thanks to the use of timeline-based planning, EXPLORAA is effective in addressing the problem of planning personalized lessons and dynamically adapting them at execution time. The definition of the planning problem through a formal language, however, tends to be an extremely slow process and prone at introducing errors. Hence, the idea of supporting this activity through the use of intelligent techniques.

4. Supporting the Authoring through the WIKITEL System

As can be easily deduced from the definition of the planning problem, the crucial and most difficult part concerns the definition of the rules, which, during the planning process, define the introduction of new tokens and the constraints between their parameters. In defining the rules, specifically, the teacher must take into account all the feasible evolutions of the lesson, considering all the possible learning objects for the students, their relationships, and how these are related to the state of the users. In particular, these possible evolutions are defined, ultimately, by applying the defined planning rules which, although allowing the problem to be easily broken down into subproblems and, therefore, help designing the lesson, are currently handwritten on a text file, without the support of any editor specifically built to take account of the application context. Since the use of authoring tools would certainly be able to reduce the workload assigned to course designers [6], we have developed a new tool that aims at tackling the editing problem by exploiting machine learning techniques, so as to be able to automate and, therefore, to facilitate, as much as possible, the task of designing the lessons, while maintaining the dynamic adaptation characteristics offered by the rule-based tutoring systems.

The pursued idea consists in extending the EXPLORAA system with a tool, shown in the Figure 2, capable of supporting the editing of lesson models, so as to significantly reduce the workload required by teachers at lesson design phase. The resulting system, called WIKITEL, aggregates the benefits of EXPLORAA, such as the capability of planning personalized lessons and dynamically adapting them at execution time, with WCB’s enhanced content creation capabilities. As can be seen from Section 3, the most complex part for defining a planning problem concerns the definition of the rules. The figure, in particular, shows a graphical user interface which, thanks to an intelligent support, aims at helping the content creator in defining the rules. The figure, in particular, shows a graphical user interface which, thanks to an intelligent support, aims at helping the content creator in defining the rules. On the left of the figure is shown a list of lesson models which, thanks to the WIKITEL system, can be easily edited. On the right, the learning objects associated with the selected lesson model are enumerated. Each of these learning objects, in particular, is associated with a planning rule which, starting from the content entered in the graphical interface, is automatically generated. Finally, the central graph shows the relationships between the learning objects of the selected lesson model. Note that these relations are associated to a lesson model and are, therefore, independent of the profiles of the students who will eventually follow the lesson. In other words, the lesson models contain the knowledge which allows the personalization thanks to the resolution of the planning problem. At a later stage, when the lesson model is sufficiently detailed to meet the pedagogical needs of the content creator, the teacher will select a lesson model, i.e., a set of rules for the planning problem, a set of students, whose profiles will define the initial state of the planning problem, and a set of a learning objects, that is, a set of goals for the planning problem. At this stage, the planning problem can be solved by the planner, which will eventually generate a solution constituting a personalized lesson for the chosen students.
The definition of the rules, in the same graphical interface, can be performed by the content creator by selecting the single learning objects. Figure 3, in particular, shows the information that can be entered in the system for defining the rule associated to a selected learning object. For each learning object, in particular, this information will be translated into a planning rule for the planner. The head’s rule, specifically, is represented by the learning object. The body, on the contrary, is defined on the basis of the information entered in the central part of the page. The top part, in particular, allows the insertion of some general information, such as the name of the selected learning object and the URL of the Wikipedia page. In the central part, the topics covered by the learning object are indicated. These topics are transformed into constraints for the planner that allow the planning of personalized lessons based on the interests of the students. Such topics are initially proposed to the content creator on the basis of the algorithm described in the Section 5, while still leaving to the author of the model the possibility to select them manually. The next space is dedicated to the duration, in seconds, of the learning object. This dimension is used by the planner to schedule the different tokens over time, preventing their temporal overlapping, thus providing the student with the time necessary to assimilate the topic without being pressured by new incoming learning objects. Once again this field can be edited by the teacher who defines the lesson model. The system, however, provides an estimate of the time required based on the length of the page. Finally, the related learning objects, which will characterize the sub-goals of the planning rule, are listed on the bottom of the page. Temporal constraints will place such sub-goals before the selected learning object, in case a bottom-up approach is selected, or after the selected learning object, in case a top-down strategy is selected.

Once these contents have been entered, the information for defining the rules is completely available. For each learning object, specifically, the rule that defines the conditions for its achievement will consist of a conjunction of: a constraint on the duration of the token and, for each of the related learning objects, a disjunction between (1) the presence of the related learning object, as a sub-goal, in the lesson, with an ordering constraint, before or after the head, depending on whether the rule is bottom-up or top-down; and (2) for each of the topics discussed within the sub-goal, a constraint on the user’s profile not to be interested in the specific topic dealt with. It is, therefore, clear that, in order to fully support the content creator, the above information must be entered as semi-automatically as possible. For this reason, therefore, we have equipped the system with intelligent techniques that propose contents to the content creator, still leaving the author the possibility to edit them.
A particularly interesting aspect in support of the editing of the lesson models concerns the way in which the related learning objects are proposed to the teacher. As shown on Figure 4, in particular, the most relevant related learning objects are selected and sorted according to their relevance, greatly simplifying to the author the task of selecting them. More specifically, the related topics are extracted from the links linked to the Wikipedia page analyzed. After the first process of selection of the page sections containing text useful for defining the main concept on the page, all the pages in the list are analyzed. This sorting is based on the value resulting from the cosine-similarity between the TF-IDF [32] vector of the query terms and the vector of TF-IDF of the terms of the i-th result. Both vectors are constructed after the stemming phase of the texts and the removal of all the stopwords, to normalize all the words contained therein. In literature, different algorithms exist with which it is possible to carry out stemming; in the WIKITEL platform Porter’s stemmer (https://www.nltk.org, accessed on 1 August 2021) was leveraged, which is the most used for the English language and the Snowball stemmer for the Italian language. The statistic value of the TF-IDF is then calculated: for each word in the query, the values of term frequency and inverse document frequency according to the following formulas:

$$tf(t,d) = \log(1 + f_{t,d})$$

$$idf(t,D) = \log(1 + \frac{N}{n_t})$$

In other words, the term $t$ is assigned a weight in the document $d$ which is:

- Higher when $t$ occurs many times in a small number of documents (resulting highly characterizing for those documents);
- As low as the term occurs few times in the document, or occurs in many documents;
- Extremely lower when the term appears in all documents.

There are several variations of the functions of TF-IDF, many of which are normalized versions of the original function.
Figure 4. Related topics are sorted according to their relevance.

For the calculation of the cosine similarity \([32]\), the query and each article are considered present in the set of results as vectors, according to the following formula:

\[
\cos(\vec{q}, \vec{d}) = \frac{\vec{q} \cdot \vec{d}}{|\vec{q}| |\vec{d}|} = \frac{\sum_{i=1}^{\|V\|} q_i d_i}{\sqrt{\sum_{i=1}^{\|V\|} q_i^2} \sqrt{\sum_{i=1}^{\|V\|} d_i^2}}
\]

(3)

where \(Q_i\) is the TF-IDF value of the term \(i\) in the query \(q\), while \(d_i\) is the value TF-IDF of the term \(i\) in the document \(d\). The value \(\cos(q, d)\) represents the cosine of the angle between the vectors \(\vec{q}\) and \(\vec{d}\) (see Figure 5). Each document is represented by a vector of real values, which constitute the TF-IDF weights \(\in \mathbb{R}^{\|V\|}\); we have, therefore, a vector space \(\|V\|\)-dimensional, where \(V\) represents the set of vectors, in which the terms constitute the axes of space and the documents of points (or vectors) in it.

Figure 5. Cosine similarity between vectors.

The space built in this way has a number very large in dimensions, even tens of millions if applied to an engine search, considering that the vectors are scattered and made up of many even values to zero.

5. User Modeling for a Personalized Experience

Cultural heritage disseminating to a targeted audience is a topic intersecting many domains: art history, marketing, cognitive studies, etc. Summing up a long path of research, it is worth stressing some main elements [33]:

...
The ability of any individual to understand and keep in memory information is proportional to his or her starting knowledge: the culture already owned the topic (the paradoxical consequence is that the communication will probably be more effective for people who need it less);

A useful way to approach the problem seems to be considering an information chain as a variously shaped sequence of elementary “narrative units”, to be defined according to the movie structure’s analysis by [34], as the elementary sequence of:

1. Somebody ...
2. ... who makes something happen (or: to whom something happens) ...
3. ... which leads to a change of the situation.

“This seems to be an acceptable starting point to focus a “bit” of narration as elementary sequence of events. At the same time, such a simplest form is, in cognitive terms, the very elementary narrative structure to be managed by a 1–2 year child mind [35].” (from [33]).

The success of storytelling applied to CH dissemination seems to depend on its effectiveness in both strengthening the cause-effect relationships (thus creating narrative units when they are not already present in users’ knowledge); and boosting the emotional impact: the inner link connecting a simple event to the universal human experience, making it closer to our life.

Given such a premise, the idea of building “knowledge packages” (paths of linked information) targeted to different kinds of users, has to be considered in relation to specific visitors “profiles”. There is a huge literature on how setting visitors classes, generally in relation to their expertise, behavior, attitude, interests [36,37]. In our case, we had to set up a sufficiently simple model to allow forms of automated reasoning in support to the definition of the lessons’ content, yet more complex than a merely knowledge-based “expert vs novice” or similar dichotomy [37], taking into account both personal expertise and interest to the matter, which are, as we have stated before, linked each other. We have, therefore, chosen to associate each user with a set of 48 topics (e.g., agriculture, anthropology, archaeology, architecture, etc.) that would represent their interests. These topics are selected by the user during the registration phase to the system, by means of the graphic interface shown in Figure 6. Additionally, in order to represent the users’ previous knowledge, we keep track of all the learning objects received by the student. This latter information is particularly relevant for avoiding sending the same learning objects to the same user several times in case the student takes several lessons with overlapping topics. Formally,

**Definition 2.** The user model, for a student $u$, is an expression of the form

$$UM_u(T, K)$$

where:

- $T$ is the set of topics of interest to the user;
- $K$ is the set of learning objects sent to the student, in the past, by the system.

The characterization of the users based on the topics of their interest allows to significantly facilitate the definition of the lesson models. The same characterization, in fact, can be performed on the Wikipedia pages, thus allowing to associate Wikipedia pages with user profiles. All the pages in Wikipedia are indeed associated with a category graph. Starting from a root node, all entities are tagged with one or more of these categories. It is, therefore, possible to define for each node two functions which return the lists of the categories higher than the entity under examination (super-categories) and of the lower categories (sub-categories).
Figure 6. The WIKITEL’s graphical interface allows users to specify their own topics of interest. The user of the figure, in particular, is interested in anthropology, archeology, astronomy, etc.

To customize the user experience, an algorithm has been developed which, at runtime, can explore the graph starting from the page in question. The algorithm goes up the entire graph and associates a list of macro-categories (nodes connected only to the root of the graph) to each topic to contextualize the user’s search. The initial version of the algorithm used to trace the graph performing a series of calls to the Wikipedia API, to recursively explore the super-categories of each node up to the root. However, in the first experiments, the execution times were excessive, taking 5–10 min per execution. To optimize these performances and make the system usable, we have decided to import locally all the Italian and English version of Wikipedia within a MongoDB instance. To do that, we started from the official dumps released by Wikimedia (https://dumps.wikimedia.org/, accessed on 1 August 2021), and through the dumpster-dive software (https://github.com/spencermountain/dumpster-dive, accessed on 1 August 2021) we created the collection in MongoDB (2 h for the Italian Wikipedia and 6 h for the English one). Finally, to optimize the exploration time of each level of the categories graph, we have imported the SQL tables “category”, “category-links”, and “page” of the Wikimedia dump. Through the exploitation of these tables the collection created in MongoDB was pre-processed, enriching it with information on the macro-categories and with information based on NLP techniques to further increase performance; specifically, the following fields have been inserted: - An already stemmed version of the plaintext; - The categories associated with each page; - The associated macro-categories; - A data structure with the plaintext stemmed and store as a bag-of-word divided by sentences; - Information on outgoing links to all connected pages (level two of the Wikipedia link graph starting from the selected page) to make predictions on the most promising pages to explore first.

In order to work on the latest version of Wikipedia, an automatic system has been developed that reproduces the steps described above once a month, when WikiMedia updates the dumps. Therefore, every month the system downloads the xml dump and sequentially execute the operations to obtain the updated and expanded MongoDB collection, as described previously.
6. Learning “Matera” with WIKITEL

The idea described in this article was also pursued in the light of a new project, CLEVERNESS, which aims to expand the ideas of EXPLORAA to create a system capable of providing cultural and social integration stimuli for older people. Given the current limitation of the fruition of artistic visit, due to the current crisis, cultural activities dedicated to frail older adults have begun to be subject to rethinking and reorganization, thanks, also, to the use of technology and remote teaching/social assistance. In this perspective, the fruition and learning of tourist places can benefit from technological tools to support active and personalized lessons. The WIKITEL system also embraces this vision and wants to be a first effort towards the creation of alternative tools to support personalized learning. In this section we illustrate the system described by placing it in a learning context of cultural heritage such as the city of Matera. Furthermore, the section contains an evaluation of the current status in two directions: (a) a quantitative evaluation of the time needed to create system contents; (b) a perspective on the possible innovative uses of the didactic contents created by the WIKITEL system.

6.1. The WIKITEL System at Work

To illustrate the operation of the system, we have instantiated the WIKITEL system to an example inspired by the need to prepare a lesson on the “Matera” city. In particular, as related topics, WIKITEL recognizes several relevant Wikipedia pages such as “Basilicata” (the Region in which Matera is located), “Cattedrale di Matera” (Matera’s main cathedral), and “Sassi di Matera” (the historic “Sassi” district) which, thanks to the editing interface, can easily be integrated into the lesson model. The editing process, at this point, can then proceed backwards, selecting the related topics and letting the editor help the author in the choice of other related topics, up to build a rich and varied lesson model like the one shown in the graph in Figure 7. The figure, more specifically, shows the related topics recognized by WIKITEL and selected by the author, providing a detail of the planning problem related to the notions related to a lesson about “Matera”. The nodes, in particular, denote all the possible learning paths that can be passed through by the different types of users, with the different learning objects $s_t$ associated for each node, with the different types of users $u_j$. From a planning perspective, in particular, a predicate symbol for each learning object has been defined, with a parameter $u$ representing the user to whom it is directed. By selecting a node of the graph it is possible to view (and, if necessary, modify) the recognized themes of interest of the selected topic. These themes, in particular, are translated into logical constraints within the planning problem, the satisfaction of which leads to the selection of personalized learning objects to be sent to individuals.

Similarly to the EXPLORAA system, in the WIKITEL case, the initial user model is described by a set of fact tokens in a timeline-based planning problem while the pedagogical objectives are described by a set of goal tokens. The application of the rules within the resolution process introduces additional tokens which represent the personalized learning objects to be sent to users. The presence of disjunctions within the rules, in particular, together with the unifications that will eventually take place with the tokens that describe the user’s status, will guarantee, though the automated planning problem resolution, the automatic production of a plan of personalized learning objects over time.

Figure 8 shows the solution of a planning problem constituting a personalized lesson, about the topic “Matera”, for a given student with specific interests. The upper part of the figure shows two timelines: a state-variable, containing the different tokens representing the personalized learning objects of a lesson arranged over time, and a reusable-resource, representing the student’s busy times. The latter timeline, in particular, can be used to prevent the system from planning activities at certain times of the day within which the student does not want to receive learning objects (e.g., meal times, night, etc.). The lower part of the figure, on the contrary, shows the causal graph resulting from the computation. This graph shows the different tokens that make up the planned personalized lesson and the different causal relationships that exist between them.
Finally, for the sake of completeness, the following code snippet shows (a part of) a rule, expressed in the \texttt{riddle} language (an initial description of the \texttt{riddle} language is provided at https://github.com/pstlab/oRatio/wiki/The-RIDDLE-Language, accessed on 1 August 2021), used for asserting the causal constraints between different learning objects on a specific lesson model. In particular, it is worth underlining that this rule
was automatically generated by the system starting from the information entered in the graphical interface during the definition of the lesson model.

```java
predicate LO_3(User u) : Interval { duration >= 6994.0; fact bt = new u.busy_time.Use(start: start, duration: duration, end: end, amount: 1.0); { goal lo6783 = new LO_6783(u: u); lo6783.start >= end; } or { !u.c_geograpia; } ... }
```

Specifically, the previous rule refers to the learning object #3, corresponding to the “Matera” Wikipedia page, and is applied to achieve the LO_3 goals whenever it is not possible to unify them with other tokens. Specifically, in the riddle syntax, the rule states that the predicate LO_3 has a parameter u of type User, indicating the user, among those participating in the lesson, who will receive the learning object, and inherits from the Interval predicate the start, the end, and the duration parameters. The body of the rule, within the curly brackets, defines a conjunction of requirements establishing, in the above case, that each LO_3 goal must have a duration greater or equal than 6994 s, must have a unitary consumption on the busy_time resource of user u with the same start, end and duration as those of the LO_3 goal and, thanks to the definition of a disjunction, either a subgoal of type LO_6783 for the same user u which must necessarily start after the end of LO_3, or the user u is not interested in geography. Intuitively, the #6783 learning object deals with the geography topic, hence the rule assigns the #6783 learning object only to those users which are interested in geography. The rule ends with further disjunctions, not shown for sake of space, which associate further arguments to the LO_3 goals.

6.2. On the WikITEL Current Performances

In order to provide an idea of the computation times required by the planning procedure, Figure 9 shows mean and standard deviation, among 5 different runs, of the resolution times, in seconds, compared to the number of learning objects present in the personalized lesson. It should be noted that a large part of the computation times are reserved for the management of scheduling constraints which, in our case, aim to prevent the overlapping over time of the different learning objects that compose the lesson. Such constraints are managed using domain independent scheduling techniques (i.e., [38]), which introduce the shown variability in the resolution times. Given the peculiarity of the problem, it is certainly possible to adopt more efficient and domain specific ordering techniques which, while maintaining the constraints imposed by the rules, establish precedence constraints between the learning objects. From the graph, nonetheless, it is possible to see that, although the resolution times tend to slightly grow exponentially, with the increase in the number of learning objects present in the personalized lesson, these remain perfectly compatible with a practical use of the system.

The requests in the first version of the system to calculate the macro categories uses a series of recursive calls to the wikimedia api. For example, to associate a macrocategory to the entity “Matera” the system takes 250 s.

To speed up the computation, we have moved the query system on DB, and the answers are instantaneous (0.5–1 s).

We launched the execution only starting from a subset of the nodes of the Wikipedia graph reaching up to eight “jumps” (link) around the main topic “Palombaro Lungo”. The work completed on the collection takes 1.7 days for this prototype.
7. Discussing the Uses of WikiteL

The WikiteL structure, as conceived and shaped by the working team and described above, was mainly focused in order to fit learning requirements, also in consideration of the COVID-19 emergency and the need of improving on line resources managing. Nevertheless, on the one hand, the pandemic situation fostered the development of technologies whose usability will keep their evidence also in a normalized context, on the other, the flexibility of WikiteL technology, as a ‘serendipity’ side-effect, appeared to us as a feature able to adapt the tool to a wide range of working situation, spreading well over the original learning dimension. Taking into account the current professional contexts, in which interconnections among cultural fields are constantly increasing their relevance, the WikiteL application may be defined as a real ‘cultural cross-over’: a tool to focus disciplinary links and help in deepening the knowledge components of cultural heritage. To illustrate such features we can deal with the example of “Matera” (a small city known all over the world for the historic “Sassi” district, which make it one of the oldest still inhabited cities in the world), a cultural context not too wide (as the great Italian art capitals, such as Rome, Florence, etc.), but sufficiently dense of cultural items and conceptual paths, many different scenarios may be drawn as potential chances for the use of such a tool:

Scenario #1
Diana is a tourist guide. She lives and works in Matera. She is originally an art historian, but she is used to lead different kind of groups, also in connection with corporate initiatives of different professionals or specific targets. Therefore, she has worked with leading groups of engineers, bank employees, generic tourists, old people, and so on. WikiteL will give her an important help in preparing each time a tour tailored on each group, focusing the topic more interesting for her crew.

Scenario #2
Daniel is an high school teacher. WikiteL will help him to create lessons on specific topics identifying a range of subjects and giving each student the chance to study Matera’s cultural environment deepening different domains identified by the system, according to his or her personal interests. Thus, each student will shape a different report, even about the same monument.
Scenario #3
Nora is a freelance journalist. She is regularly charged to write articles on different subjects for various journals, and every time she starts working she needs to approach the given subject focusing the aspects more attractive for each different customer. WikiTEL will make her able to identify each time the topics to enhance, to make her article about Matera’s art history tailored for the journal she’s dealing with (and to write and sell different articles for different customers, on the same cultural topic).

Scenario #4
Vito is a politician. He has just been elected as member in Matera’s County Council, and he is setting up a commission to draw a law concerning new rules for building constructions and restoring the existing ones laying upon the Matera’s Palombaro Lungo area, in order to avoid damaging the ancient structure and preserve the original features. Thanks to WikiTEL, he will be able to identify the topics involved and focus all the points of view, as to select experts from the relative fields, to compose an adequate commission for writing up the law.

These are just a few hypothetical examples of the WikiTEL contexts of action. Generally speaking, the tool allows to reach in a short time, from a culturally relevant subject, a satisfactory overview of a series of connections and nodes, to be used in different perspectives in terms of learning-building, planning, creative working, etc.

8. Conclusions
The main contribution in this work is to provide a single environment for the development of learning paths and the fruition of cultural heritage in an adaptive and personalised way. WikiTEL is a platform that aims to be used by creators of cultural routes, exploiting the knowledge base of Wikipedia, but also enriching the content with other materials, films, or other web pages of interest.

With the support of WikiTEL, the course creator can quickly create a course according to the predefined learning objectives, thus reducing the workload of course designers.

WikiTEL also offers a customised training path based on the interests of the users, providing training stimuli at the right time. Adaptivity is supported by the categorization of Wikipedia resources. The training path is generated by a timeline planner.

The exploration of Wikipedia takes place locally. Thanks to dumps synchronized with Wikimedia updates, the generation of the Wikipedia sub-graph of interest to the course designer has reduced execution times.

Techniques of this type are fundamental in the current context: the experience of distance learning, forced by the pandemic, has shown that the IT tools for the delivery of courses in synchronous or asynchronous mode have many flaws. For example, in many distance courses, the percentage of students dropouts is very high due to less involvement and, more often, due to frustration. These techniques aim at creating courses that follow the student and can change the stimuli produced to adapt the learning path to their needs, making him or her feels always involved. At the same time, the features of the system make allow a quick systematic arrangement of cultural contents in conceptual nodes, which will be useful in a wide range of activities, still far from being completely explored.

WikiTEL is still in a preliminary stage and the interface must be improved in order to allow a simple and intuitive use even to non-experts in the use of current technologies. From the user modeling point of view, massive experimentation can highlight which aspects of the model need to be better specified. At the moment we have mainly specified the user’s interests, but it can be detailed to highlight gender, ethnicity, religion, language, etc. The system must also offer accessible and usable pathways, as outlined in [39].

Author Contributions: Conceptualization, all authors; methodology, all authors; software, R.D.B. and C.D.M.; writing—original draft preparation, R.D.B.; writing—review and editing, all authors. All authors have read and agreed to the submitted version of the manuscript.
Funding: This research is supported by CNR under FOE2020 funds to the strategic project “Technologies to support the most vulnerable groups: young and old – CLEVERNESS”. A.C. and R.D.B. are also funded by CTEMFT project “Casa delle Tecnologie a Matera”.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations
The following abbreviations are used in this manuscript:

ExPLoRAA Experiential learning for active ageing
WCB Wiki course builder
WikiTEL Wikipedia-based technology-enhanced learning

References
1. Pavlova, L.; Draganov, L.; Luchev, D.; Paneva-Marinova, D.; Szántó, G.; Mártus, L.; Szkaliczki, T. Technology-enhanced Learning for Cultural Heritage. *Digit. Present. Preserv. Cult. Sci. Herit.* 2015, *V*, 293–301.
2. Westera, W. Technology-Enhanced Learning: Review and prospects. *ACM Comput. Classif. Syst.* 2010, *4*, 159–182.
3. Pozzi, F.; Alivizatou, M.; Dagnino, F.; Ott, M. Going Beyond Preservation: How to Support Technology-Enhanced Learning in ICH Education. *Int. J. Herit. Digit. Envi 2015*, *4*, 21–40. [CrossRef]
4. Gruber, M.; Glahn, C. E-learning for arts and cultural heritage education in archives and museums. In *Applying Digital Media to Culture*; Hasebrook, J., Muhr, G., Schrader, A., Eds.; IOS Press: Amsterdam, The Netherlands, 201.
5. KOK, A. Understanding the Technology Enhanced Learning Environments from A Cognitive Perspective. *Int. Educ. Stud.* 2009, *2*, 3–9. [CrossRef]
6. Alevren, V.; McLaren, B.M.; Sewall, J.; Koedinger, K.R. The Cognitive Tutor Authoring Tools (CTAT): Preliminary Evaluation of Efficiency Gains. In *Intelligent Tutoring Systems*; Ikeda, M., Ashley, K.D., Chan, T.W., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 61–70.
7. De Benedictis, R.; Cortellessa, G.; Fracasso, F.; Cesta, A. Technology-enhanced learning to support active ageing. *Form@Re-Open J. Per La Form. Rete 2019*, *19*, 301–311. (In Italian)

Ghallab, M.; Nau, D.; Traverso, P. *Automated Planning: Theory and Practice*; Morgan Kaufmann Publishers Inc.: Burlington, MA, USA, 2004.

Muscettola, N. HSTS: Integrating Planning and Scheduling. In *Intelligent Scheduling*; Zweben, M., Fox, M.S., Eds.; Morgan Kaufmann: San Francisco, CA, USA, 1994.

Cesta, A.; Cortellessa, G.; De Benedictis, R.; De Medio, C.; Fracasso, F.; Limongelli, C. Toward Automated Courseware Production for the ExPLoRAA Learning Environment. In *Proceedings of the A*PIA 2019-Advances in Artificial Intelligence-XVIIIth International Conference of the Italian Association for Artificial Intelligence, Proceedings, LNCS 11946, Rende, Italy, 19–22 November 2019; pp. 528–541.

Limongelli, C.; Lombardi, M.; Marani, A. Towards the Recommendation of Resources in Coursera. In *Intelligent Tutoring Systems: 13th International Conference, ITS 2016*. Proceedings; Springer International Publishing: Cham, Switzerland, 2016; Volume 9684, p. 461.

De Medio, C.; Gasparetti, F.; Limongelli, C.; Scairrone, F. Automatic Extraction and Sequencing of Wikipedia Pages for Smart Course Building. In *Proceedings of the 21st International Conference Information Visualisation (IV)*, London, UK, 11–14 July 2017; pp. 378–383.

Cesta, A.; Cortellessa, G.; De Benedictis, R.; De Medio, C.; Limongelli, C.; Scairrone, F.; Tassarotti, G.; Palombini, A. Personalizing Technology-Enhanced Learning for Cultural Visits. In *Proceedings of the Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, Genoa, Italy, 14–17 July 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 333–339. [CrossRef]

Gruber, M. The Role of E-Learning in Arts and Cultural Heritage Education. In *Kreativität und Innovationskompetenz im Digitalen Netz*; Hornung-Prähauser, V., Luckmann, M., Eds.; Salzburg Research: Salzburg, Austria, 2009 pp. 343–350.

Mohammed, P.; Mohan, P. Breakthroughs and Challenges in Culturally-Aware Technology Enhanced Learning. In *Proceedings of Workshop on Culturally-Aware Technology Enhanced Learning in Conjunction with EC-TEL*. 2013. Available online: http://phaedramohammed.com/papers/cultTEL2013-mohammed-mohan.pdf (accessed on 1 August 2021);

Riches Resources. Museum Education with Digital Technologies: Participation and Lifelong Learning. 2016. Available online: https://ec.europa.eu/futurium/en/system/files/gd/rch_thinkpapers_06.pdf (accessed on 1 August 2021).

García Fernandez, J.; Medeiros, L. Cultural Heritage and Communication through Simulation Videogames–A Validation of Minecraft. *Heritage 2019*, *2*, 2262–2274. [CrossRef]
18. Not, E.; Petrelli, D. Empowering cultural heritage professionals with tools for authoring and deploying personalised visitor experiences. *User Modeling User Adapt. Interact.* 2019, 29, 67–120. [CrossRef]
19. Lim, S. How and why do college students use Wikipedia? *J. Am. Soc. Inf. Sci. Technol.* 2009, 60, 2189–2202. [CrossRef]
20. Selwyn, N.; Godart, S. Students’ use of Wikipedia as an academic resource—Patterns of use and perceptions of usefulness. *Internet High. Educ.* 2016, 28, 28–34. [CrossRef]
21. Fidas, C.; Sintoris, C.; Yiannoutsou, N.; Avouris, N. A survey on tools for end user authoring of mobile applications for cultural heritage. In Proceedings of the 2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA), Corfu, Greece, 6–8 July 2015; pp. 1–5.
22. Porteous, J.; Lindsay, A.; Read, J.; Truran, M.; Cavazza, M. Automated extension of narrative planning domains with antonymic operators. In Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems (AAMAS’15), Istanbul, Turkey, 4–8 May 2015; International Foundation for Autonomous Agents and Multiagent Systems: Richland County, SC, 2015; pp. 1547–1555.
23. Porteous, J.; Ferreira, J.F.; Lindsay, A.; Cavazza, M. Extending narrative planning domains with linguistic resources. In Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, 9–13 May 2020.
24. Cesta, A.; Cortellessa, G.; De Benedictis, R.; Fracasso, F. Active Aging by Continuous Learning: A Training Environment for Cultural Visits. In *Ambient Assisted Living*; Leone, A., Caroppo, A., Rescio, G., Diraco, G., Siciliano, P., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 521–533.
25. Cesta, A.; Cortellessa, G.; De Benedictis, R.; Fracasso, F. ExPLoRAA: An Intelligent Tutoring System for Active Ageing in (Flexible) Time and Space. In Proceedings of the 4th Italian Workshop on Artificial Intelligence for Ambient Assisted Living, Trento, Italy, 20–23 November 2018; CEUR-WS; Volume 2333, pp. 92–109.
26. Cesta, A.; Cortellessa, G.; De Benedictis, R. Training for Crisis Decision Making—An Approach Based on Plan Adaptation. *Knowl. Based Syst.* 2014, 58, 98–112. [CrossRef]
27. Young, R.M. Notes on the use of plan structures in the creation of interactive plot. In *Proceedings of the AAAI Fall Symposium on Narrative Intelligence*; The AAAI Press: Menlo Park, CA, USA, 1999; pp. 164–167.
28. Lecoutre, C. *Constraint Networks: Techniques and Algorithms*; Wiley-IEEE Press: Hoboken, NJ, USA, 2009.
29. Dechter, R. *Constraint Processing*; Elsevier Morgan Kaufmann: San Francisco, CA, USA, 2003.
30. Apt, K.R.; Wallace, M.G. *Constraint Logic Programming Using ECL/IPSP*; Cambridge University Press: New York, NY, USA, 2007.
31. De Benedictis, R.; Cesta, A. Lifted Heuristics for Timeline-based Planning. In Proceedings of the ECAI-2020, 24th European Conference on Artificial Intelligence, Santiago de Compostela, Spain, 29 August–8 September 2020.
32. Baeza-Yates, R.A.; Ribeiro-Neto, B. *Modern Information Retrieval*; Addison-Wesley Longman Publishing Co., Inc.: Boston, MA, USA, 1999.
33. Palombini, A. Storytelling and telling history. Towards a grammar of narratives for Cultural Heritage dissemination in the Digital Era. *J. Cult. Herit.* 2017, 24, 134–139. [CrossRef]
34. Casetti, F.; Di Chio, F. *L’analis del Film*; Bompiani: Milano, Italy, 1990.
35. McKeough, A. Narrative knowledge and its development: Toward an integrative framework. *Issues Educ. Contrib. Educ. Psychol.* 1996, 2, 73–81.
36. Almeshari, M.; Dowell, J.; Nyhan, J. Using Personas to Model Museum Visitors. In Proceedings of the PATCH Workshop, Larnaca, Cyprus, 9–12 June 2019.
37. Walsh, D.; Clough, P.; Foster, J. User Categories for Digital Cultural Heritage. In Proceedings of the ACHS Workshop, Newark, NJ, USA, 22 June 2016.
38. Cesta, A.; Oddi, A.; Smith, S.F. A Constraint-Based Method for Project Scheduling with Time Windows. *J. Heuristics* 2002, 8, 109–136. [CrossRef]
39. The Ethics of Artificial Intelligence: Issues and Initiatives; Scientific Foresight Unit (STOA) European Parliamentary Research Service. Available online: https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2020)634452 (accessed on 1 August 2021)