Learning Design with an AI Assistant

Blaženka Divjak(✉), Barbi Svetec, Petra Vondra, Josipa Bađari, and Darko Grabar

Faculty of Organization and Informatics, University of Zagreb, Pavlinska 2, Varaždin, Croatia
{blazenka.divjak,barbi.svetec,petra.vondra,josipa.badjari, darko.grabar}@foi.unizg.hr

Abstract. This paper presents the ongoing experimental work related to the design of an AI assistant using API calls to an LLM to support educators in creating sound learning design (LD): LeDA. LeDA is one of the innovative features of a collaborative LD tool *Balanced Design Planning*, founded on the principles of student-centeredness and constructive alignment, that has so far been used by more than 2000 users in over 40 countries. We explain LeDA's background in research and international projects, outline the software development, and present the results of ongoing validation. LeDA has been validated in 34 higher education courses, helping 22 teachers to generate, customize, and optimize LDs. Data on the use of LeDA has been automatically collected in the tool, user feedback gathered via integrated forms, and through a questionnaire administered among the teachers. In the current total of 664 interactions with LeDA, the rate of accepting AI's suggestions without modification is around 40%, with the majority relating to creating teaching and learning activities (TLAs). User feedback suggests that teachers recognized LeDA's potential in generating learning outcomes, topics, units, and TLAs, while they were less satisfied with the assistance in integrating innovative pedagogies. Generally, it was reported that LeDA is more efficient in generating solutions on a lower conceptual level, especially if it is provided with the user's comprehensive input. The findings showed that LeDA was easy to use, but pointed to possible adjustments, which would contribute to higher usability.

Keywords: Learning Design · AI · AI assistant · Learning Analytics · API · LLM

1 Introduction

In the last couple of years, the rapid advancement of generative artificial intelligence (GenAI) has opened new challenges and possibilities in education. One of the possibilities refers to supporting educators in designing (more) effective learning experiences, including through provision of evidence-base for meaningful decision-making, as well assistance in the design of curricula and teaching and learning activities (TLAs). This means that educators have to learn how to use GenAI constructively, exploiting its benefits, but at the same time being aware of its limitations and their own responsibility to act as critical thinkers and role models in terms of ethics and academic integrity.

A. I. Cristea et al. (Eds.): AIED 2025, LNAI 15877, pp. 207–220, 2025.
https://doi.org/10.1007/978-3-031-98414-3_15

In this paper, we present the ongoing experimental work related to an AI assistant developed to support educators in creating sound learning design (LD): LeDA. LeDA is a part of the innovative, collaborative, free-to-use LD tool *Balanced Design Planning* (BDP) (learning-design.eu), which is based on the principles of student-centeredness and constructive alignment and currently has over 2000 users across 40 countries.

After briefly explaining the overall BDP concept and tool, we outline the development of LeDA and demonstrate its main features. We present the results of the pilot and validation exercise, which included 34 higher education (HE) courses and 22 HE teachers. The pilot identified LeDA's strengths and weaknesses and informed the recommendations for more effective work with and optimization of an AI assistant in LD, which could be of use in the development of other similar tools in the evolving landscape of digital pedagogy and contribute to the overall discourse on AI in education.

2 Background

2.1 Potentials and Challenges of AI in Education

GenAI has a potential to transform education as we know it [7]. On the one hand, it can support teachers in informed decision-making during instructional design, based on learning analytics [20, 22], and help them make teaching and learning more personalized [16]. For example, AI can be used to track and analyze students' learning performance in real-time to develop a student learning profile. Teachers can use this data to suggest personalized LD and methods to meet different needs and interest of students [17]. GenAI can also decrease educators' workload, help them deal with challenges and enhance LD [5, 16]. On the other hand, GenAI causes concerns about ethics, academic integrity, and unfair practices [4]. Some studies [5] have pointed out the weaknesses of using GenAI for course design, in terms of lacking understanding of the course context, generating unfeasible activities and suggesting resources and materials that do not exist.

To make meaningful use of GenAI in education, including in LD, it is necessary to balance its use with critical thinking and creativity, and avoid overreliance on AI tools [5]. In fact, it has been pointed out that the most significant AI effects occur when machines complement humans and collaboration takes place [13, 26]. In this paper, we present an example of an AI assistant supporting such educator-AI collaboration in LD, mindful of the importance of designers' domain knowledge and critical approach to AI's outputs, aiming to support responsible integration of AI in educational practices.

2.2 Learning Design

There are different definitions of learning design (LD), but in most basic terms, it refers to the order of TLAs, with respective resources and student support [19], which teachers and students should do so that students would acquire the intended learning outcomes [18]. LD guides educators, helping them make informed decisions when designing TLAs, in line with a particular pedagogical approach [6], to ensure efficient teaching and learning [2]. LD should be learner-centered and correspond with student needs [2].

As a field of practice and research, LD emerged in the 2000s [19], so it stresses the use of technology [2]. Recently, there has also been an increase in exploring the possibilities for synergies and integration of LD with learning analytics [19, 23]. Different approaches and tools have been developed to support LD (e.g., ABC LD by University College London; OULDI by The Open University UK). In this paper, we present an innovative approach that builds on the previous ones, on contemporary research, and the latest technological developments, to support educators in development of LD that is not only pedagogically sound but also data-informed, innovative and creative.

2.3 Learning Design Concept, Tool, and AI Assistant

Motivated by the needs expressed by HE teachers in several HE institutions around Europe, the University of Zagreb (UNIZG) Faculty of Organization and Informatics' Laboratory for Learning Analytics has developed an innovative LD concept and tool: the BDP (learning-design.eu). The BDP concept and tool have been explained in several publications [10, 11, 12, 8, 24]. The tool has undergone several cycles of development in line with the design cycle methodology [25]. The BDP supports educators in developing student-centered LD, which is strongly based on learning outcomes and constructive alignment between learning outcomes, TLAs and assessment [3], and student workload. It encourages the use of innovative pedagogies and design analytics.

The BDP approach starts with basic course information. It continues with course learning outcomes, their levels [1] and prioritization [8, 10]. In the following steps, LD is developed to the level of topics, units and TLAs, linked to the intended learning outcomes. TLAs are planned in detail, including the corresponding workload, assessment, and learning type (acquisition, discussion, investigation, practice, production, assessment). In every step, users have access to detailed design analytics to ensure the quality of LD. The BDP tool enables direct export of course LDs to the Moodle LMS.

This paper focuses on one of the latest developments of the BDP tool: an AI assistant in LD which uses API calls to an LLM (GPT-4o): LeDA. LeDA follows the BDP theoretical concept [10] and is built around two principles. First, LeDA enables interactivity, cooperating with an educator through written text, providing responses to clearly formulated prompts. Second, LeDA provides assistance, as it depends on educators' input, needs and requests, but does not make decisions autonomously. Similar to the overall development of the BDP tool, the development of LeDA has included the three phases of the design cycle [25] (Fig. 1) and followed human-centered design [15].

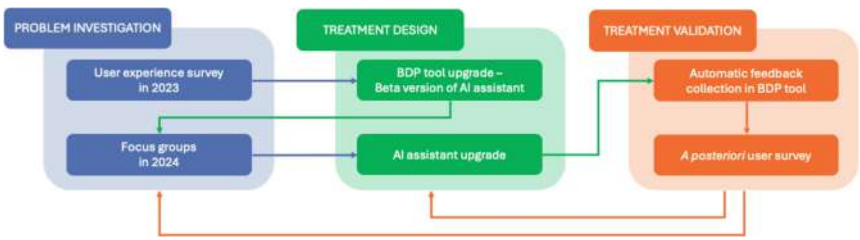

Fig. 1. Development of the AI assistant as part of the LD tool design cycle

In the **problem investigation** phase, the requirements for the AI assistant were determined as part of the process of validation of the BDP LD tool. This phase included:

(1) *A user experience survey* conducted in 2023, as part of the *Innovating Learning Design in HE* (iLed) project, with 53 teachers from HEIs in four European countries [11]. The survey showed that teachers would appreciate the help of a chatbot in the formulation of learning outcomes, support in ensuring constructive alignment, help with design analytics interpretation, and real-time guidance in LD. This motivated the development of the beta-version of LeDA. (2) *Focus groups on AI algorithms to support LD* held in 2024, as part of the *Trustworthy Learning Analytics and Artificial Intelligence for Sound Learning Design* (TRUELA) project, with 18 experts from Europe and Africa. The focus groups (held after the preliminary treatment design) were presented with the beta-version of LeDA and provided preliminary highly positive feedback, but also included ideas for further development, related to AI's help in design analytics interpretation, integration of innovative pedagogies, and content creation.

In the **treatment design** phase, a technical upgrade of the BDP tool was done resulting in the beta version of LeDA (Fig. 2). An important aspect was the preparation of context-based prompts. Each prepared prompt consists of five parts: GPT role (learning designer), Context (course details), Query (desired output), Rules & Constraints (BDP tool related definitions/rules/constraints) and Output Specification (format of output that responds to BDP course format). An example of a prompt is provided below.

```
{
"model": "gpt-4o",
"messages": [
    {"role": "system",
    "content": "You are acting as a learning designer, supporting teachers/educators in designing their courses.\n\nHere are some definitions for key terms:\n\n- **Learning outcomes** are statements of what a learner knows, understands and is able to do on completion of a learning process. Each learning outcome has a corresponding level of Bloom's taxonomy:\n\n 1. **Remembering** – Exhibit memory of previously learned material by recalling facts, terms, basic concepts, and answers. \n   - *Verbs:* Choose, Define, Find, List, Match, Name, Recall, Relate, Select, Describe\n 2. **Understanding** – Demonstrate understanding of facts and ideas by organizing, comparing, translating, interpreting, giving descriptions, and stating main ideas. \n   - *Verbs:* Classify, Compare, Contrast, Demonstrate, Explain, Illustrate, Interpret, Outline, Relate, Rephrase, Summarize\n 3. **Applying** – Solve problems to new situations by applying acquired knowledge, facts, techniques and rules in a different way. \n   - *Verbs:* Apply, Build, Choose, Construct, Develop, Experiment with, Identify, Interview, Make use of, Model, Organize, Plan, Select, Solve\n 4. **Analyzing** – Examine and break information into parts by identifying motives or causes. Make inferences and find evidence to support generalizations. \n   - *Verbs:* Analyze, Assume, Categorize, Classify, Compare, Distinguish, Divide, Examine, Inspect, List, Simplify, Survey, Test for\n 5. **Evaluating** – Present and defend opinions by making judgments about information, validity of ideas, or quality of work based on a set of criteria. \n   - *Verbs:* Assess, Choose, Compare, Conclude, Criticize, Decide, Deduct, Defend, Determine, Estimate, Evaluate, Explain, Interpret, Judge, Justify, Mark, Measure, Prioritize, Prove, Recommend, Select\n 6. **Creating** – Compile information together in a different way by combining elements in a new pattern or proposing alternative solutions. \n   - *Verbs:* Adapt, Build, Choose, Combine, Compose, Construct, Create, Design, Develop, Elaborate, Estimate, Formulate, Improve, Invent, Plan, Predict, Propose, Solve, Test a theory\n\nYou MUST respond in the **Croatian language**, even if existing content is in a different language.\n\nYou will be asked to suggest **units for a topic** based on provided course and topic information, which will be given in JSON format. \n\nThe topic you're working with has **ID 8281**.\n\nYou should:\n- Suggest up to **8 units** within defined topics\n- Specify the **learning outcomes** each unit contributes to\n- Indicate the **percentage of contribution** to those outcomes\n- Define the **level of Bloom's taxonomy** for each\n- Provide a **summary explanation** of why you're suggesting those units\n\n**Units** must include:\n- A **title**\n- A **description**\n\nFormat your response as **JSON** with the following structure:\n\n```json\n{\n  \"suggestions\": [\n    {\n      \"unit\": \"Unit title\",\n      \"description\": \"Unit description\"\n    },\n    ...\n  ],\n  \"explanation\": \"Short explanation of why these units were suggested\"\n}\n```\n\n**Constraints:**\n- The topic is the highest level part of a course, consisting of units and TLAs related to one comprehensive learning topic (e.g., one chapter).\n- A unit is a mid-level component of a topic, consisting of TLAs that are content-related and sequenced according to a pedagogical approach.\n- The topic with ID 8281 already has defined units — avoid overlaps when suggesting new ones."
    },
    {
    "role": "user",
    "content": "**Course data (as JSON)**"
    }
    ]
}
```

With this upgrade, LD is supported by an AI assistant able to: 1.1. Generate learning outcome(s), 1. 2. Generate learning outcome(s) according to selected prevailing level of Bloom's taxonomy (Fig. 3), 2.1. Generate topic(s), 2.2. Generate topics and their contribution (%) to learning outcome(s) achievement, 3.1. Generate unit(s) within a selected topic, 3.2. Generate units within a topic according to selected specific teaching and learning approach, 4.1. Generate TLAs within the selected unit (Fig. 4), 4.2. Generate TLAs supporting the prevailing learning type within the selected unit.

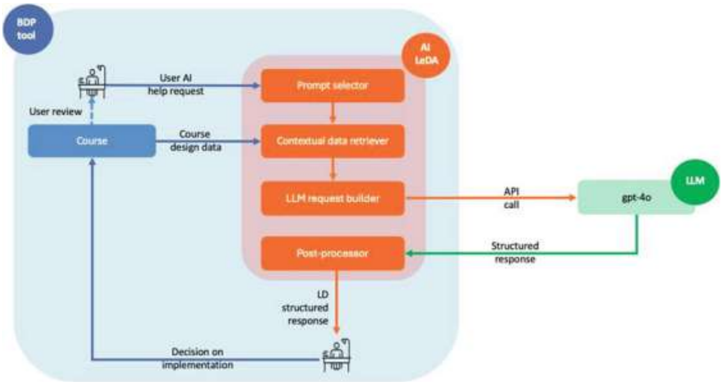

Fig. 2. LeDA's architecture

Importantly, on each level (e.g., topic) proposals are generated considering the content of higher levels of design (e.g., learning outcomes), in line with the principles of LD: (a) Learning outcomes are generated based on course data (description, keywords, target audience, course type, education level, workload in hours, number of course participants, and delivery mode). (b) Topics are generated based on course data and learning outcomes. (c) Units are generated based on course data, learning outcomes, topics and how much (%) the topic contributes to a specific learning outcome. (d) TLAs are generated based on course data, learning outcomes, topics and units.

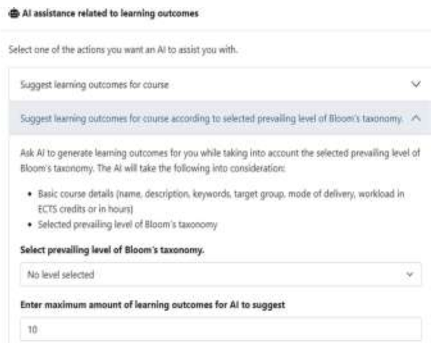

Fig. 3. Query for LeDA: suggest Los

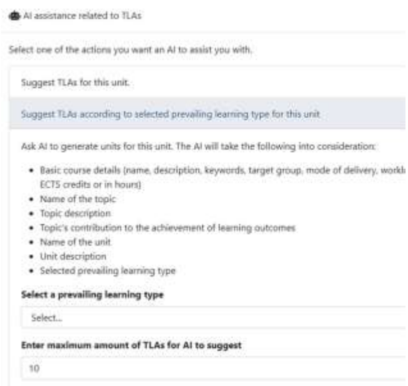

Fig. 4. Query for LeDA: suggest TLA

While the prompts are pre-developed, educators can also direct the AI assistant further. This means they can determine the number and some characteristics of the LD elements they wish to be suggested by LeDA, including the preferred pedagogical approach and required learning types. Furthermore, the AI-assisted LD process is designed to encourage educators' critical thinking and interventions into AI's suggestions. Therefore, besides the suggestion of an LD element (e.g., a TLA), LeDA also provides an educator with a disclaimer reminding them to approach LeDA's suggestions critically. To support human critical thinking, AI's suggestions can be either automatically implemented in course design or a prompt is reiterated and accepted when the AI's suggestion meets a satisfactory LD quality. There is a dashboard displaying LD analyses, including the acceptance of AI's suggestions in different elements of LD.

In terms of architecture, LeDA comprises four functional components as follows:

Prompt Selector (PS): Upon receiving a help request from the user (e.g., for new unit suggestions), the PS identifies the appropriate prompt template. These templates are crafted in advance to align with BDP knowledge base.

Contextual Data Retriever (CDR): CDR fetches relevant course design data from the BDP tool, such as existing units, learning outcomes, and topics. This contextual information is necessary to ground the LLM's response in the current state of the course, thereby avoiding redundancy and ensuring alignment with pedagogical intentions.

LLM Request Builder (RB): RB synthesizes the selected prompt template and contextual data into a unified API request. This includes formatting the input according to the expected structure (e.g. JSON), setting parameters for model behavior, and ensuring that instructions are language-specific.

Post-Processor (PP): Once the LLM returns a structured response, PP validates and adapts the output for the BDP environment. This includes formatting adjustments, semantic validations, and error handling. PP ensures that the final output is pedagogically coherent, correctly structured, and ready for user review within the BDP interface.

In the **treatment validation** phase, users' feedback was collected. The pre-testing was done on several selected courses, followed by a wider pilot and validation. The feedback was collected with two aims: to provide input for the next upgrade of LeDA, and to provide a basis for recommendations for educators and meaningful use of LeDA. We report on the treatment validation, i.e. the pilot, in the following sections.

3 Methodology

In the treatment validation phase (Fig. 5), a pilot was conducted at UNIZG, encompassing a total of 34 courses and 22 HE teachers (users) in the interdisciplinary field, mainly Information Technology (IT). Most of the courses are delivered at undergraduate and graduate study levels (26), but there are also courses from professional studies (2), postgraduate studies (2) and a lifelong learning training (4). For most of the courses, LD had already been developed in the BDP tool without AI assistance. Now, the users were asked to redesign their courses with LeDA's support, so that the LDs created without and with AI assistance could be analyzed. Several users could work on one course, and each user had a possibility to (re)design more than one of the courses they teach. All the users included in the pilot had previous experience with LD in the BDP tool and were motivated for educational innovation.

The users' feedback on the LeDA functionality was collected in two ways: **(a) Automatic collection in the LD tool.** The BDP tool enables collection of user feedback immediately after each interaction with LeDA. This includes a numerical grade (1 - 5) and qualitative written comments. **(b)** ***A posteriori*** **user survey.** A questionnaire was distributed among the HE teachers who took part in the pilot. It included 24 questions, divided into four groups: information about users and courses, general usefulness of the AI assistant, usefulness of the AI assistant in LD aspects, and acceptance of AI in education. The questions were either multiple-choice or open-ended (qualitative).

With this pilot, we aimed to respond to the following research questions (RQ):

RQ1: In what aspects does the AI assistant (LeDA) enhance the efficiency and quality of learning design for educators and what aspects present a challenge?

RQ2: What are the recommendations for users and developers that could contribute to the more efficient use of LeDA?

Fig. 5. Research cucle

4 Results

4.1 LeDA System Data

By 5 February 2025, a total number of interactions with LeDA (i.e. users asking for LD suggestions) was 664. According to quantitative data, based on the numerical grades awarded by the users (HE teachers) in real time, the acceptance rate of LeDA's responses without any user modification was 38%. The highest proportion of the accepted responses refers to creating units (48%), followed by TLAs (40%) and topics (34%), while the acceptance of learning outcomes comes last (24%). As for LeDA's usefulness, the mean mark (1 – 5 scale) awarded by the users was 3.

When it comes to the qualitative comments given by HE teachers during design with LeDA, the feedback is categorized based on different LD elements:

Learning Outcomes. In general, users found LeDA's suggestions related to learning outcomes somewhat useful but inconsistent in quality. While some suggestions were rated positively, most were either rejected or required modification. Users preferred learning outcomes that were better contextualized and tailored to course content. The low acceptance rate (24%) indicates that most suggestions were not fully satisfactory. The main groups of users' comments indicated that suggested learning outcomes were either too high or too low on the Bloom's taxonomy scale and required better calibration, did not fully align with the course description or not entirely relevant to the course. Moreover, they thought the proposals were not well-phrased to correspond to the standard learning outcome format or were too general and needed more specificity.

Topics. In general, users indicated that AI-generated topics were moderately relevant but sometimes redundant. Some topics repeated content instead of introducing new ideas and users wanted more diverse and specific topic suggestions. The acceptance rate of 34% indicates that some suggestions were useful, but many required users' interventions. The main groups of users' comments indicated that suggested topics overlapped too much, were not always relevant to the course or lacked practical application, were sometimes repetitive, making it difficult to structure diverse course content, and could be too generic, lacking depth or specificity.

Units. Users generally found the AI-generated units more useful than topics and learning outcomes. They found that in most cases unit suggestions were well-structured and aligned with learning outcomes. However, some suggested units lacked depth or connection to the broader course structure. Nearly half of the AI-generated units (48%) were immediately accepted. The main groups of users' comments indicated that suggested units were appreciated, confirming AI's ability to generate useful ideas, but could be too broad, generic, lack innovation, or cover entire courses instead of specific topics; sometimes they were not sufficiently aligned with learning outcomes or topics.

TLAs. According to users' general perspective, LeDA was partially effective in suggesting appropriate TLAs. The 40% acceptance rate indicates that some activities were useful, but many needed user improvements. The main groups of users' comments indicated that suggested TLAs were useful but needed refinements to fit the course structure; they were sometimes too generic and not well-adapted to specific course needs, not applicable to the learning context or lacked practical implementation strategies.

4.2 *A Posteriori* Survey Data

A total of 22 HE teachers participated in the survey, from different fields: pedagogy (1), mathematics (4), IT (8), economy (6), interdisciplinary (9). We present the highlights of results, with percentages rounded to the nearest integer due to the sample size.

A significant portion were novel HE teachers, with 1 – 5 (36.4%) or 5 - 10 years of teaching experience in HE (5%), and mid-career teachers with 10 - 20 years of experience (36%), followed by highly experienced teachers, with 20 - 30 (14%) or more than 30 years (9%) of teaching experience. In terms of their previous experience with BDP LD, the total range was from 0 to 12 LDs: a majority had created between 1 and 5

LDs (73%); several had no previous experience with the BDP tool (14%). For a majority (64%), this was the first time to work with an educational chatbot.

As presented in Fig. 6, a majority of the users (68%) were (very) satisfied with their LDs prepared previously without AI assistance. Interestingly, a vast majority (86%) felt improvements in LD would be welcome, and a majority (73%) would apply the LD designed with LeDA at least to some extent if they were authorized to do so.

As shown in Fig. 7, all survey participants agreed that LeDA is easy to use (completely or mostly agree). A great majority (82%) recognized LeDA as useful for their work in teaching and research and almost three quarters (73%) plan to use it in future. In qualitative comments, the users pointed out it would be useful if it were possible to partially accept LeDA's suggestions, combine them with pre-existing LD, provide LeDA with additional instructions, and use AI's help to improve the existing LD. Some also asked for more insight into AI's reasoning. They also thought it would be useful to train LeDA on quality assured examples of sound LD.

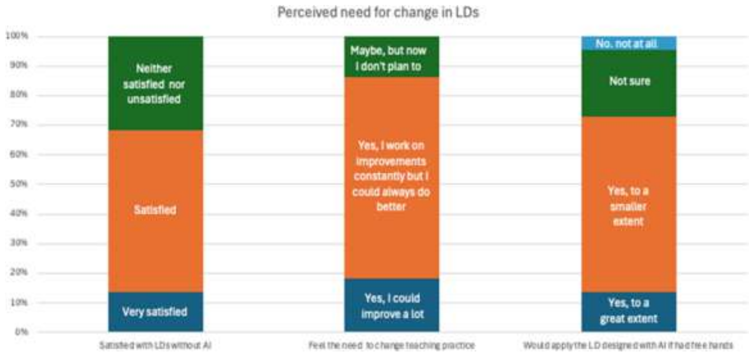

Fig. 6. Perceived need for change in LDs

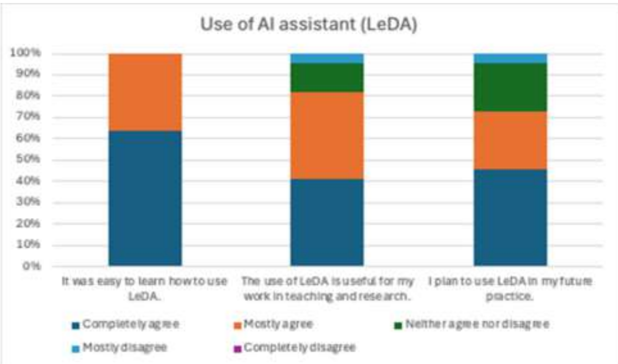

Fig. 7. Use of AI assistant

In terms of the usefulness of LeDA in different aspects of LD (Fig. 8), most of the users (68%) found it very useful or mostly useful in designing TLAs and planning topics related to learning outcomes. A bit smaller portion (64%), but still a majority, evaluated LeDA as very useful or mostly useful in helping them to improve learning outcomes. The lowest usefulness was found in terms of support for integrating innovative pedagogies in LD, which was found very/mostly useful by a half of the teachers (55%). Many users still have reservations about this, as one-third of them (32%) were not sure. In their qualitative comments, users reported that LeDA enticed them to make changes in LD (even if they did not accept its original proposal), especially when it comes to embracing a more creative approach to the design of TLAs. In terms of lessons learnt, several users pointed out that LeDA was a useful tool for improvements and a source of ideas, which should be taken critically and adjusted as needed.

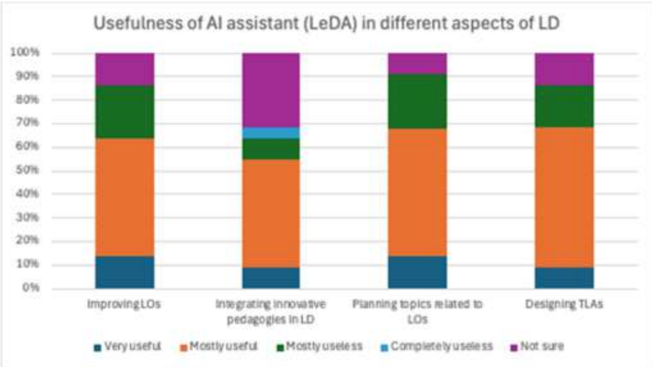

Fig. 8. Usefulness of the AI assistant in different aspects of LD

A vast majority of the users thought AI can have a positive influence on their teaching and research, especially when it comes to improving some aspects of their work or saving time (68%) and helping them be more efficient (27%). A majority (68%) considered themselves (mostly) ready to meaningfully apply AI in practice. However, the most frequent grade they assigned to their current success in integrating AI in teaching (1 – insufficient, 5 – excellent) was 3 (41%), followed by 2 (27%) and 4 (23%), while only a minority chose 5 (9%). As for concerns about AI in education, in qualitative comments, the users stressed the ethical use of AI by users and students, issues with assessment, lack of critical approach to solutions proposed by AI, and the unpreparedness of institutions for fast development of AI. They also pointed out that students could use AI more than teachers and become more skilled users, at the same time not developing critical thinking and possibly developing dependence on AI. One user warned that teachers and students can accept AI tools that are not scientifically grounded.

5 Discussion

5.1 RQ1: Benefits and Challenges of LeDA

Our results from the two data sources were highly aligned; we provide them with a link to recommendations for users (RU) and developers (RD). The results show that the teachers were generally satisfied with the AI assistant, and especially some aspects of

AI assistance. In principle, the users found that LeDA is more effective on lower levels of LD (units, TLAs), which is reflected in the higher acceptance rate of AI's suggestions without any modification (48% and 40% respectively). They were less content with AI assistance on the higher levels of LD (learning outcomes, topics), which are more abstract and overarching, as reflected in the acceptance rate (24% and 34% respectively). The reasons for that might be three-fold. First, LeDA uses input from a specific LLM (GPT-4o), but also from basic descriptions and features of a course as provided by a teacher. If this descriptive part is not informative enough, LeDA infers from the pre-training data of GPT-4o and generates outputs that may lack specificity [21]. Consequently, teachers can be reluctant to use its suggestions without any modification. (RU1 and RD1) Second, higher-level elements of LD are normally parts of study program curricula [14], meaning that teachers have less flexibility to change them outside the regular (often lengthy and complex) QA procedures. Therefore, it would be commendable to start LD from the study program level. (RU2 and RD2) Third, the current possibilities provided by LeDA do not enable teachers to choose some (but not all) of the proposed learning outcomes and topics. This calls for more flexibility in the tool (RD3).

When it came to units and TLAs, teachers were highly satisfied, but also pointed to some aspects for improvement. First, the teachers noted that there is a need for consistent alignment between units, topics and learning outcomes. This implies that teachers need to be experts in pedagogy and LD, which would enable them to meaningfully evaluate the AI results, taking into account the design analytics. It suggests that more coherent results could be achieved if the AI assistant was trained on examples of good LDs. This could be achieved by applying the Retrieval-Augmented Generation (RAG) which includes the development of a structural knowledge base on LD. This could enable higher specificity in unit and TLA descriptions, more context-specific suggestions and avoiding generic or overly broad suggestions. (RU3 and RD4) Moreover, the teachers noted that LeDA could be more innovative in proposing TLAs. While this could be further improved on the technical level, for example by introducing a storyboard and course visualization, it should be stressed that meaningful use of AI means that teachers are still conductors and must approach AI's outputs critically. (RU4, RU5 and RD5) Finally, the teachers were concerned about AI's reasoning in general, and asked for more insight into how LeDA functions and creates suggestions (RU6 and RD6).

5.2 RQ2: Recommendations for Users and Developers

Based on the challenges identified in relation to RQ1, we provide a list of recommendations for users and developers that could contribute to the quality of LeDA's output and its more meaningful use, so as to increase the efficiency of teaching and learning.

Recommendations for users (RU):

RU1. Be detailed, focused and informative in the basic description of the course LD. Enhance AI's understanding of course context to generate more relevant learning outcomes and topics.

RU2. Consider the overarching study program perspective when designing a course.

RU3. Work on your data literacy and skills enabling you to make sense of design analytics and learning analytics.

RU4. Be aware of your conductor / decision-making role in LD and its orchestration.

RU5. Be aware that an AI assistant is not a person but a very capable, although not unmistakable machine.

RU6. Work on your AI literacy as a basis for a critical approach to AI's suggestions.

Recommendations for developers (RD):

RD1.To support better AI's suggestions related to higher levels of LD, the BDP tool has recently been upgraded with additional fields to fill in information about the course, e.g. target groups of learners, level of education, type of the course.

RD2. To support the study program perspective, the BDP tool should enable linking courses to study program learning outcomes and workload.

RD3. It would be useful to enable users to select the preferred learning outcomes among AI-generated suggestions.

RD4. The possibilities to train LeDA on examples of LD with a quality mark should be explored. This is in relation with the application of RAG which includes development of a structural knowledge base on LD.

RD5. It would be valuable to enable course visualization and a storyboard, to start LD from the fundamental idea of the course, and that way support more creativity in LD.

RD6. Additional explanations of AI's reasoning and a help section should be introduced to support users in critical evaluation of the suggestions.

5.3 Limitations

The main limitation of this pilot is that it was conducted with a limited number of users, from a limited number of courses, on one HEI. Examining user experience with more users in different educational contexts might provide additional insights. Moreover, the teachers included in this pilot had previous experience with the BDP concept and tool, and belonged to a HEI which is mainly in the field of IT, both of which might have made them more critical towards the LeDA's suggestions.

6 Conclusion

We presented one of the innovative functionalities of the BDP learning design (LD) concept and tool, namely, the LD AI assistant (LeDA). To validate the usefulness of LeDA, as well as to identify the key challenges and recommendations for more effective use of AI's assistance by educators and for further technical upgrades by developers, we analyzed LeDA system data and user feedback. We found that users generally appreciated the assistance provided by LeDA, especially when it comes to lower levels of LD (units and TLAs), while they were more critical about LeDA's outputs at higher levels (learning outcomes and topics). We recommended that educators should be more concrete in providing course descriptions, aware of their responsibilities as the conductors of LD, work on their data and AI literacies, and consider the broader study program perspective. We should all be aware that the dignity and conductor role of humans should always be preserved, as AI can be a powerful assistant but is still a machine.

Acknowledgements. This study has been conducted with the support of the iLed (Erasmus+, financed by the European Union, 2022–1-HR01-KA220-HED-000085818) and TRUELA (financed by the Croatian Science Foundation, IP-2022–10-2854) project.

References

1. Anderson, L.W., Krathwohl, D.R.: A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives, Complete Longman, New York (2001)
2. Bennett, S., Agostinho, S., Lockyer, L.: Technology tools to support learning design: Implications derived from an investigation of university teachers' design practices. Comput. Educ. **81**, 211–220 (2015). https://doi.org/10.1016/j.compedu.2014.10.016
3. Biggs, J.: Enhancing teaching through constructive alignment. High. Educ. **32**(3), 347–364 (1996). https://doi.org/10.1007/BF00138871
4. Chang, D.H., Lin, M.P.-C., Hajian, S., Wang, Q.Q.: Educational design principles of using AI Chatbot that supports self-regulated learning in education: goal setting, feedback, and personalization. Sustainability **15**, 12921 (2023). https://doi.org/10.3390/su151712921
5. Choi, G.W., Kim, S.H., Lee, D., et al.: Utilizing generative ai for instructional design: exploring strengths, weaknesses, opportunities, and threats. TechTrends **68**, 832–844 (2024). https://doi.org/10.1007/s11528-024-00967-w
6. Conole, G.: Designing for Learning in an Open World. Springer New York, New York, NY (2013). https://doi.org/10.1007/978-1-4419-8517-0
7. Cooper, G.: Examining science education in ChatGPT: an exploratory study of generative artificial intelligence. J. Sci. Educ. Technol. **32**, 444–452 (2023)
8. Divjak, B., Grabar, D., Svetec, B., Vondra, P.: Automating the comparison of learning design and delivery using course scaffolding in moodle. In: Novitzká, V., Szakál, A. (eds.) INFORMATICS 2024 PROCEEDINGS, pp. 54–59. IEEE, Poprad (2024)
9. Divjak, B., Grabar, D., Svetec, B., Vondra, P.: Balanced Learning Design Planning. J. Inf. Organizational Sci. **46**(2), 361–375 (2022). https://doi.org/10.31341/jios.46.2.6
10. Divjak, B., Kadoic, N., Zugec, B.: The use of decision-making methods to ensure assessment validity. In: 2021 IEEE Technology & Engineering Management Conference - Europe (TEMSCON-EUR). IEEE, May 2021, pp. 1–6. https://doi.org/10.1109/TEMSCON-EUR52034.2021.9488580
11. Divjak, B., Rienties, B., Bađari, J., Grabar, D., Horvat, D., Vondra, P.: Enhancing learning design through user experience research: insights from a survey in four european countries. In: Proceedings of the Central European Conference on Information and Intelligent Systems, Dubrovnik, Croatia, pp. 213–221 (2023)
12. Divjak, B., Svetec, B., Horvat, D., Kadoić, N.: Assessment validity and learning analytics as prerequisites for ensuring student-centred learning design. Br. J. Edu. Technol. **54**(1), 313–334 (2023). https://doi.org/10.1111/bjet.13290
13. Duan, Y., Edwards, J.S., Dwivedi, J.K.: Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. Int. J. Inf. Manage. **48**, 63–71 (2019). https://doi.org/10.1016/j.ijinfomgt.2019.01.021
14. European Higher Education Area: Rome Ministerial Communiqué 2020 (2020). https://ehea.info/Upload/Rome_Ministerial_Communique_Annex_III.pdf
15. Giacomin, J.: What Is Human Centred Design? Des. J. **17**(4), 606–623 (2014). https://doi.org/10.2752/175630614X14056185480186

16. Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernandez-Leo, D., Rienties, B.: The promise and challenges of generative AI in education. Behav. Inf. Technol., 1–27 (2024). https://doi.org/10.1080/0144929X.2024.2394886
17. Kim, J.: Leading teachers' perspective on teacher-AI collaboration in education. Education and Information Technologies 29 (2023). https://doi.org/10.1007/s10639-023-12109-5
18. Koper, R., Olivier, B.: Representing the learning design of units of learning. Educ. Technol. Soc. **7**(3), 97–111 (2004)
19. Lockyer, L., Heathcote, E., Dawson, S.: Informing pedagogical action. Am. Behav. Sci. **57**(10), 1439–1459 (2013). https://doi.org/10.1177/0002764213479367
20. Mangaroska, K., Giannakos, M.: Learning analytics for learning design: a systematic literature review of analytics-driven design to enhance learning. IEEE Trans. Learn. Technol. **12**(4), 516–534 (2018)
21. Meyer, J.G., Urbanowicz, R.J., Martin, P.C.N., et al.: ChatGPT and large language models in academia: opportunities and challenges. BioData Mining **16**, 20 (2023). https://doi.org/10.1186/s13040-023-00339-9
22. Muljana, P.S., Luo, T.: Utilizing learning analytics in course design: voices from instructional designers in higher education. J. Comput. High. Educ. **33**(1), 206–234 (2021)
23. Pishtari, G., et al.: Learning design and learning analytics in mobile and ubiquitous learning: a systematic review. Br. J. Edu. Technol. **51**(4), 1078–1100 (2020). https://doi.org/10.1111/bjet.12944
24. Rienties, B., Balaban, I., Divjak, B., Grabar, D., Svetec, B., Vondra, P.: Applying and Translating Learning Design and Analytics Approaches Across Borders. In: Springer (2023), pp. 35–53. https://doi.org/10.1007/978-3-031-27646-0_3
25. Wieringa, R.J.: Design Science Methodology for Information Systems and Software Engineering. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-43839-8
26. Wilson, H.J., Daugherty, P.R.: Human + Machine: Reimagining Work in the Age of AI. Harvard Business Review Press, Boston, MA (2018)