

Integrating Learning Context and Explainability in Educational Recommender Systems Using Markov Decision Process over Knowledge Graphs

DISSERTATION

zur Erlangung des Grades eines Doktors der Ingenieurwissenschaften

vorgelegt von

M.Sc. Hasan Abu Rasheed

eingereicht bei der Naturwissenschaftlich-Technischen Fakultät

der Universität Siegen

Siegen 2025

Betreuer und erster Gutachter
Prof. Dr.-Ing. habil. Madjid Fathi
Universität Siegen

Zweiter Gutachter
Prof. Dr.-Ing. habil. Roman Obermaisser
Universität Siegen

Tag der mündlichen Prüfung
23.09.2025

Acknowledgment

I would like to express my deepest gratitude to my first supervisor and examiner, Prof. Dr.-Ing. habil. Madjid Fathi, my second examiner, Prof. Dr.-Ing. habil. Roman Obermaisser, and my supervisor and mentor, Dr. Christian Weber, whose invaluable guidance, encouragement, mentorship, and expertise have been instrumental in shaping this thesis. Their constructive feedback and continuous support throughout the research process have been one of the main factors in refining my ideas and ensuring the success of this work.

My sincere thanks also extend to my colleagues at the Institute for Knowledge-Based Systems and Knowledge Management, with whom I had the great pleasure of working. I am grateful for their insightful discussions, camaraderie, and collaboration during this journey. Prof. Alexander Holland, Dr.-Ing. Mareike Dornhöfer, Dr.-Ing. Johannes Zenkert, M.Sc. Mubaris Nadeem, M.Sc. Parvin Ghadamighalandari, and Dr. rer. nat. André Klahold, have all been counter stones in experience I have built throughout the last years. A special acknowledgment is extended to the MIGS group for their remarkable support and team spirit, which created an inspiring and motivating environment for research and collaboration since the very first days of building the group.

Finally, I would like to thank my family, my parents, brother, sisters, and their families, for their unconditional love, encouragement, and patience throughout the years of my academic pursuit. Their trust and belief in me have been a source of strength during challenging times, and their support has made this achievement possible.

Abstract

Human learning is a complex and multi-dimensional process, governed by a wide range of factors that describe the individual differences between learners, and the contexts in which learning happens. Since these differences influence how learners respond to the learning content and activities, it is essential for an effective learning process to consider these factors when generating learning recommendations. Context-aware recommender systems (CARS) have offered a promising solution for tailoring learning experiences to the specific learning contexts. However, existing CARS often fall short in comprehensively integrating complex contextual data into their reasoning. Moreover, complex CARS face challenges in providing transparent and explainable recommendations to learners and educators, especially when they have no technical background, which hinders CARS's effectiveness and acceptance among these educational stakeholders.

This thesis addresses these challenges by developing a novel method for building **Context-Aware Recommendations and Explainability through Knowledge Graphs (CARExKG)**. Structurally, the CARExKG method employs knowledge graphs to represent contextual learning factors and their interdependencies, capturing the dynamic interplay between various contextual variables in complex learning settings. Algorithmically, CARExKG employs a Markov decision process over knowledge graphs, featuring a context-sensitive reward function tailored to enable the RS to generate contextualized learning paths across various learning settings.

In order to ensure human oversight, reduce stakeholders' resistance to recommendations, and encourage collaborative human-AI decision-making, CARExKG further incorporates an explainability framework, utilizing large language models and expert-input from pedagogy specialists, to generate user-centric explanations, supporting learner's understanding of the reasoning behind generating the recommended path, and enhancing their decision-making ability and ownership of their educational journey.

To evaluate the proposed method, a set of experiments was designed to measure the effectiveness of the knowledge graphs, the recommendation algorithm, and the explainability framework. CARExKG is then evaluated as a complete system through a real-world user study with nursing staff in two elderly homes, where a complex learning scenario was constructed to mimic the multi-dimensional challenges they face in their profession. Evaluation results have demonstrated the effectiveness of the proposed approach for constructing the knowledge graph and the reasoning of the RS. Results also show the ability of the CARExKG method to improve learner satisfaction and outcomes in vocational education and training settings. Experiment findings underscore the importance of the interdisciplinary approach followed in designing CARExKG, which combines artificial intelligence, educational technology, and pedagogy to create adaptive, explainable, and learner-centered educational tools.

Zusammenfassung

Menschliches Lernen ist ein komplexer und multidimensionaler Prozess, der von einem breiten Spektrum von Faktoren bestimmt wird, die die individuellen Unterschiede zwischen den Lernenden und die Kontexte, in denen Lernen stattfindet, beschreiben. Da diese Unterschiede Einfluss darauf haben, wie die Lernenden auf die Lerninhalte und -aktivitäten reagieren, ist es unerlässlich, diese Faktoren bei der Erstellung von Lernempfehlungen zu berücksichtigen. Context Aware Recommender Systems (CARS) bieten eine vielversprechende Lösung für die Anpassung von Lernerfahrungen an den jeweiligen Lernkontext. Die bestehenden CARS sind jedoch oft nicht fähig, komplexe kontextbezogene Daten umfassend in ihre Überlegungen einzubeziehen. Darüber hinaus stehen komplexe CARS vor der Herausforderung, transparente und erklärbare Empfehlungen zu bieten, vor allem, wenn Lernenden und Lehrenden keinen technischen Hintergrund haben, was die Effektivität und Akzeptanz von CARS für sie beeinträchtigt.

Diese Dissertation befasst sich mit diesen Herausforderungen, indem sie eine neuartige Methode zur Erstellung kontextbezogener Empfehlungen und Erklärbarkeit durch Wissensgraphen (CARExKG) entwickelt. Strukturell verwendet CARExKG Wissensgraphen, um kontextuelle Lernfaktoren und deren Abhängigkeiten darzustellen und das dynamische Zusammenspiel zwischen verschiedenen kontextuellen Variablen in komplexen Lernumgebungen zu erfassen. Algorithmisch verwendet CARExKG einen Markov-Entscheidungsprozess über die Wissensgraphen, mit einem speziellen Design einer kontextsensitiven Reward-Funktion, die es dem RS ermöglicht, kontextualisierte Lernpfade in verschiedenen Lernumgebungen zu generieren.

Um die menschliche Kontrolle zu gewährleisten, den Widerstand der Beteiligten gegenüber den Empfehlungen zu verringern und die kollaborative Mensch-KI-Entscheidungsfindung zu fördern, beinhaltet CARExKG darüber hinaus ein Erklärungskonzept, das große Sprachmodelle und den Experteninput von Pädagogen nutzt, um nutzerzentrierte Erklärungen zu generieren, die das Verständnis der Lernenden für die Argumentation hinter der Generierung des empfohlenen Pfads unterstützen und ihre Entscheidungsfähigkeit und Eigenverantwortung für ihre Bildungsreise verbessern.

Um CARExKG zu evaluieren, wurde eine Reihe von Experimenten entwickelt, um die Effektivität der Wissensgraphen, des Empfehlungsalgorithmus und des Erklärbarkeitssystems zu messen. CARExKG wurde dann als komplettes System durch eine Nutzerstudie mit Pflegepersonal in zwei Altenheimen evaluiert, in der ein komplexes Lernszenario konstruiert wurde, um die multidimensionalen Herausforderungen, mit denen sie in ihrem Beruf konfrontiert sind, zu imitieren. Die Evaluierungsergebnisse haben die Effektivität des vorgeschlagenen Ansatzes für die Konstruktion des Wissensgraphen und die Logik des RS gezeigt. Die Ergebnisse zeigen auch die Fähigkeit der CARExKG-Methode, die Zufriedenheit der Lernenden und die Ergebnisse in der beruflichen Weiterbildung zu verbessern.

Table of Contents

List of Figures.....	xv
List of Tables	xix
1 Introduction	1
1.1 Problem statement and research motivation	3
1.2 Research questions and objectives	5
1.3 Conceptual structure of the proposed solution to thesis RQs	8
1.4 Structure of the thesis.....	12
1.4.1 Chapter 2: Theoretical Foundations	12
1.4.2 Chapter 3: Context Indication in Personalized and Adaptive Learning Recommendations. A Pedagogical and Technical Perspective.....	13
1.4.3 Chapter 4: Knowledge Graph Construction for Context-Aware Personalized Learning	14
1.4.4 Chapter 5: Markov Decision Process (MDP) over Knowledge Graphs for Learning Path Recommendation System	14
1.4.5 Chapter 6: A Framework for Explainable Recommendation Systems over Knowledge Graphs	15
1.4.6 Chapter 7: Explainability for Learning: Supporting Learner’s Agency with Interactive, Conversational, and KG-based Explainability.....	15
1.4.7 Chapter 8: Evaluation Strategy, Experiments and Results	15
1.4.8 Chapter 9: Conclusion	16
2 Theoretical Foundation	17
2.1 What is learning?.....	17
2.1.1 Learning in formal education	19
2.1.2 Learning in vocational education and training	19
2.2 Learning personalization.....	20
2.3 Learning personalization through recommendation systems.....	21
2.4 Context-aware recommender systems (CARS)	22
2.4.1 Definition of “Context”	22
2.4.2 Context capturing in CARS	24
2.5 Graph-based CARS.....	27
2.5.1 Knowledge graphs	28
2.5.2 Knowledge graphs in recommender systems	33
2.5.3 Educational knowledge graphs	35
2.5.4 Reasoning over knowledge graphs	36
2.6 Foundations of Markov decision processes	41

2.6.1 Knowledge graphs as MDP environments	42
2.6.2 Reinforcement learning for MDP-based knowledge reasoning.....	43
2.7 Educational CARS and pedagogical alignment	45
2.7.1 Challenges for CARS in education.....	47
2.8 Explainable AI in Education	48
2.8.1 Adding explainability in educational recommender systems	50
2.8.2 XAI application in educational recommender systems	52
2.8.3 Knowledge graph-based explainability	54
2.8.4 Limitations and research challenges	56
2.8.5 Ethical implications of explainability in education	56
3 Context Indication in Learning Recommendations: Perspectives from Technology and Pedagogy	61
3.1 Introduction	61
3.2 Review methodology	63
3.2.1 Literature search strategy and data collection	63
3.2.2 Selection Criteria	64
3.2.3 Literature Screening Process	66
3.3 Analysis of the Surveyed Literature.....	66
3.3.1 Contextualization Domains of Application	67
3.3.2 Learning Theories in Context-aware Approaches	68
3.3.3 Context Dimensions	69
3.4 Results and Discussion.....	72
3.4.1 Context-factor Origins	72
3.4.2 Profile-based Context Dimensions	77
3.4.3 Context Capturing and Contextual Factors	79
3.5 Related and similar reviews	85
3.6 Context representation for the knowledge graph construction	86
4 Knowledge Graph Construction for Context-Aware Personalized Learning.....	91
4.1 Knowledge graph design, construction, and requirements	92
4.2 Expert integration in KG construction	94
4.3 Use case 1 - Lessons-learned recommendations for chip-design process engineers	96
4.3.1 Use case description	96
4.3.2 KG construction methodology.....	96
4.3.3 Text mining pipeline for KG construction	98
4.3.4 Graph-based Search, information retrieval, and explainability	102
4.4 Use case 2 - Learning path recommendation for staff training in elderly- care domain.....	104
4.4.1 use case description	104

4.4.2 KG construction methodology.....	105
4.4.3 Modeling context in the knowledge graph	109
4.4.4 Text mining and analysis for semantic relation extraction.....	110
4.4.5 Enhancement with large language models	111
4.4.6 Explainability components	112
4.5 Use case 3 - Higher education module scheduling for learning personalization	113
4.5.1 Use case description	113
4.5.2 KG construction methodology.....	114
4.5.3 Automated pipeline for content extraction and classification	116
4.6 Discussion	118
5 Markov Decision Process (MDP) over Knowledge Graphs for Learning Path Recommendation Systems.....	119
5.1 Recommendation problem formulation	120
5.2 Graph path-finding algorithm	120
5.3 Graph path ranking.....	123
5.3.1 The knowledge graph as the MDP learning environment	123
5.3.2 Implementing State-Action Value Updates	125
5.3.3 Algorithm Overview	127
5.4 Discussion	128
6 A Framework for Explainable Recommendation Systems over Knowledge Graphs.....	133
6.1 Introduction	133
6.2 TExKG: a transferable, graph-based, explainability framework for RSs	135
6.2.1 Knowledge graph construction	137
6.2.2 Graph-based information retrieval for the RS	139
6.2.3 Graph-based explainability	139
6.2.4 Contextualized implementation of the framework	140
6.3 Discussion	143
7 Explainability for Learning: Supporting Learner’s Agency with Interactive, Conversational, and KG-based Explainability	145
7.1 Conversational explainability.....	147
7.2 Methodology for utilizing LLMs	148
7.3 LLM prompt contextualization	151
8 Evaluation Strategy, Experiments, and Results	157
8.1 Knowledge graph evaluation	157
8.1.1 Experiment 1: Lessons learned information retrieval.....	158
8.1.2 Experiment 2: Personalized learning in VET	161

8.1.3 Experiment 3: Personalized learning in higher education	165
8.2 Graph-based recommender system	167
8.2.1 Experiment 4: Personalized Nursing VET	167
8.3 TExKG framework.....	171
8.3.1 Experiment 5: evaluation use cases	171
8.4 Conversational explainability platform.....	183
8.4.1 Experiment 6: User acceptance in nursing VET	183
8.4.2 Experiment 7: higher education user studies	185
9 Conclusion.....	191
9.1 Future work	193
List of Publications	195
Bibliography	199
Appendix A: Post test questions of the user study (conducted in German language).....	223

List of Figures

Figure 1. Conceptual design of the solution for the research questions.	10
Figure 2. strategies for integrating context into recommendations (Adomavicius et al., 2022).	25
Figure 3. Multidimensional model for the User $\times$ Item $\times$ Time, where “Time” represents the context factor (Adomavicius et al., 2022).....	27
Figure 4. The architecture of the user-event graph embedding learning (UEG-EL) framework as proposed by (D. Liu et al., 2022).	27
Figure 5. An example Knowledge Graph structure (Bhardwaj, 2022).	29
Figure 6. Explanation approaches for a black-ox model (Barredo Arrieta et al., 2020).	52
Figure 7. Algorithmic and user-centered explainability (Afchar et al., 2022).....	54
Figure 8. Most common terms appearing in titles and abstracts of retrieved search results.	64
Figure 9. Search, screening, and selection process of the papers included in the review (Abu-Rasheed, Weber, & Fathi, 2023).	65
Figure 10. Number of papers covering "user context" (orange) and "learning context" (blue) as parts of the total paper sum per year (grey) (Abu-Rasheed, Weber, & Fathi, 2023).	66
Figure 11. Domains of application for the context-aware methods and systems in the surveyed literature (Abu-Rasheed, Weber, & Fathi, 2023).	68
Figure 12. Learning theories addressed by the proposed solution in the surveyed literature (Abu-Rasheed, Weber, & Fathi, 2023).....	69
Figure 13. Proposed classification scheme of context factors based on their representation in learning platforms (Abu-Rasheed, Weber, & Fathi, 2023).	78
Figure 14. Mapping frequent contextual factors based on their relevance to pedagogy and technology. The size of the factor reflects its frequency in the surveyed literature (Abu-Rasheed, Weber, & Fathi, 2023).	81
Figure 15. Multidimensional knowledge graph concept for use case 1.....	98

Figure 16. Domain-specific text mining pipeline for multi-source documents.	99
Figure 17. Construction of the multidimensional knowledge graph from the three node types (A), and the defined types of nodes and the relations among nodes of the same type and those of different types (B).	101
Figure 18. Graph-based search and explanation strategy.	103
Figure 19. Overall KG design and utilization in explainable lessons-learner recommendation.....	103
Figure 20. Hierarchical and network models for the learning materials in the use case (Buchmann et al., 2024).	106
Figure 21. First three levels of the KG structure, representing the role of semantic relations in creating the KG from the hierarchical one (Abu-Rasheed, Dornhöfer, et al., 2023).	109
Figure 22. Text mining pipeline (TMP) and semantic similarity calculation for title and description properties of the learning objects.	111
Figure 23. Overall KG design and utilization in creating explainable learning path recommendations for nursing VET.....	113
Figure 24. The higher education ontology with user-, curriculum-, and domain-models.	115
Figure 25. Pipeline components for the transcription, extraction, classification, and KG construction, based on a human-AI collaborative approach.	117
Figure 26. Overall system for problem-based learning path recommendation (Abu-Rasheed et al., 2024).	117
Figure 27. KG dynamic and static layers (Abu-Rasheed, Weber, Dornhöfer, et al., 2023).	121
Figure 28. Integration of contextual factors in the state-action value function	125
Figure 29. Proposed framework for transferrable, domain-oriented, explainable RS, based on knowledge graphs. Domain-specific elements are shown in white, while domain-agnostic algorithms are shown in blue (Abu-Rasheed, Weber, Zenkert, Dornhöfer, et al., 2022).	136

Figure 30. Example of integrating expert defined rules and domain requirements in the creation of graph nodes and relations (Abu-Rasheed, Weber, Zenkert, Dornhöfer, et al., 2022).....	138
Figure 31. Visual and textual explanations generated from the information and structure of the knowledge graph.....	140
Figure 32. Framework implementation in the emergency response use case (Abu-Rasheed et al., 2024).....	141
Figure 33. Framework implementation in the elderly care use case (Abu-Rasheed et al., 2024).	142
Figure 34. Framework implementation in the mental health use case (Abu-Rasheed et al., 2024).....	143
Figure 35. Proposed interface with the complete view of explainability components (Abu-Rasheed, Weber, et al., 2024a).....	146
Figure 36: Chatbot system architecture within the Angular Web-App, and its connections to the KG, recommendation algorithm, and the LLM (Abu-Rasheed et al., 2024).	150
Figure 37: Chatbot interface for user-chatbot sessions and mentor requests (Abu-Rasheed et al., 2024).....	150
Figure 38: Group-chat session, including the user (green), mentor (red) and chatbot (blue) (Abu-Rasheed et al., 2024).	151
Figure 39: Constructing the prompt's context from multiple information sources (Abu-Rasheed, Weber, et al., 2024b).....	152
Figure 40. Quantitative comparison of search results (Abu-Rasheed, Weber, Zenkert, Krumm, et al., 2022).....	159
Figure 41. Effect of LNs and structural relations on the search results and graph connectivity (Abu Rasheed et al., 2021).....	161
Figure 42. Semantic-relation textual-similarity scores (Purple), compared to the Journeys (Blue and Orange) connected by those relations through the text mining pipeline (Abu-Rasheed, Dornhöfer, et al., 2023).....	162
Figure 43. Groups comparison results with and without recommendation (Abu-Rasheed et al., 2024).....	169

Figure 44. Group comparisons on the four evaluation dimensions in pre- and post-tests (Abu-Rasheed et al., 2024).	170
Figure 45. BLEU and Rouge-L scores for the search results. The figure represents the scores for the short search queries. The average BLEU score for the short search queries (on the horizontal axis) reached 37.2%, and the average Rouge-L score was 25.1%.	173
Figure 46. BLEU and Rouge-L scores for the search results. The figure shows the scores of the long search queries. The average BLEU score for the long search queries (on the horizontal axis) reached 38.3%, and the average Rouge-L score was 27.4%.	174
Figure 47. Direct relations between user profiles and relevant job postings. Each relation is defined based on the similarity score between the two nodes.	176
Figure 48. The use of direct and transitive relations in the knowledge graph. Here, jobPosting 63 is recommended to the userProfile_15, since it is relevant to the directly connected jobPosting_45.	176
Figure 49. A sample of the overall calculated similarity scores between user profiles and job postings. Similarities above a predefined threshold are used to create corresponding relations in the knowledge graph. Negative similarity scores are caused by the angle value between document vectors, which results in a negative cosine value of that angle when calculating the cosine similarity.	177
Figure 50. Two templates that are used to generate the textual explanations. The first template is designed for direct recommendations, where only the link between the user profile and the job posting is explained. The second template includes information about the in	179
Figure 51. BLEU and Rouge-L scores for a sample of recommended job postings to two user profiles.....	180
Figure 52. Groups 2 and 4 (reversed) with recommendation, comparison with and without explanations (Abu-Rasheed et al., 2024).	185
Figure 53. Confusion matrix for the intent classifier (Abu-Rasheed et al., 2024).	186
Figure 54. User evaluation of the chatbot menus, group chat, answer quality, and speed (Abu-Rasheed et al., 2024).	187
Figure 55. Average user evaluation of the chatbot responses for each of the eight scenarios (Abu-Rasheed et al., 2024).	188

List of Tables

Table 1. Frequent contextual factors, their context-origin, profiling categories, and examples of papers addressing them in the surveyed literature (Abu-Rasheed, Weber, & Fathi, 2023)	74
Table 2. Related reviews in the literature within the review period (Abu-Rasheed et al., 2023)	86
Table 3. Number of relations introduced by different types of nodes	160
Table 4. Number of nodes in the graph.....	160
Table 5. KG quality evaluation against hierarchical model. Preferred value-trends in VET and TEL use-case are explained.....	163
Table 6. Precision (p), Recall (R), and F1 measure values for Topics and Sub-Topics of the evaluation modules	166
Table 7. Evaluation results of the proposed framework on use case 1.	174
Table 8. Precision, Recall and F1 scores of the customized NER, with regard to the predefined classes	180
Table 9. Explainability comparison based on the MEP scores.	182
Table 10. Precision, Recall, and F1 measure for each intent class	187

1 Introduction

Learning, whether in formal educational settings or vocational education and training (VET), occurs within specific contexts that shape its processes and outcomes. The context of learning is influenced by various factors, including the learner's environment, objectives, prior knowledge, preferences, and even the socio-cultural dynamics that influence their learning trajectory (Hemmler & Ifenthaler, 2022). It is not new that the overarching goal of education and training is to enable learners to apply acquired knowledge and skills to solve real-world problems, complete tasks, and achieve professional or personal objectives. However, real-world tasks are inherently multidimensional and interdisciplinary, requiring an integration of knowledge and competencies across various domains. For instance, in nursing, a professional must combine clinical knowledge, empathy, communication skills, and quick decision-making to address patient needs effectively. This complexity underscores the need for learning systems that can adapt to diverse, context-dependent challenges.

In today's rapidly evolving educational landscape, there is a growing recognition that traditional one-size-fits-all approaches to education and training are insufficient to meet the diverse needs of learners (Verbert et al., 2012). Despite the diversity of learners and their contexts, conventional education and training programs typically offer standardized curricula that fail, in many cases, to account for the unique needs, goals, and contexts of individual learners. They typically offer a standardized curriculum for all learners, despite the fact that they may work in different contexts and face diverse challenges in their respective fields (Manouselis et al., 2011). This disconnection between standardized curricula and the varied contexts in which learners apply their knowledge has led to an increased focus on personalization in education and vocational training. In vocational education and training, the issue becomes more pronounced, as learners are often preparing for specialized roles that require tailored skills and knowledge. Personalization of learning trajectories, adapting the learning content, pace, and methods to suit individual requirements, has become increasingly valuable in bridging this gap and enhancing educational outcomes and learner experience (L. Chen et al., 2020).

Problem statement and research motivation

Context-aware recommender systems (CARS) have emerged as a promising solution to address this need for contextual personalization in education. CARS in education aim to provide learners with personalized recommendations for learning resources based on their specific context, including their learning goals, prior knowledge, preferences, time availability, as well as other environmental factors. By incorporating contextual information into the recommendation process, these systems can potentially offer more relevant and effective learning experiences (Adomavicius & Tuzhilin, 2015).

Recommender systems (RSs) are not new in the scientific literature and have been researched for several decades. They provided learners with customized suggestions for learning resources and activities, whether individual ones or in sequence as learning paths. Traditional recommender systems, however, often operate with limited contextual understanding, focusing instead on general user preferences, historical interaction data, or peer-similarity recommendations (Bobadilla et al., 2013). This, in turn, leads to recommendations that may be irrelevant or inappropriate in certain situations. CARS, on the other hand, take into account various contextual factors such as time, location, social environment, prior knowledge, and user goals, to provide more accurate and personalized recommendations (Hemmler & Ifenthaler, 2022). This contextualization is particularly important in educational settings, where factors such as the learner's current knowledge level, time, and immediate learning objectives can significantly impact the usefulness of recommended resources or activities.

While recommender systems focus on learning personalization, the current advantages of AI in education have shown the importance of the stakeholder's control and agency over those recommendations. Therefore, in addition to personalization and contextualization, there is a growing emphasis on transparency and explainability in educational recommender systems, especially in vocational education and training. The need for transparency arises from the ethical considerations and trust issues associated with AI-driven decision-making systems in education. Learners, educators, and other stakeholders need to understand how and why certain recommendations are made. This is to ensure fairness and accountability of the learning process (Sonboli et al., 2021). The European Union's AI Act, which entered into force on 1 August 2024 and will be fully applicable in 2 years, addresses these concerns by setting requirements for transparent and explainable AI systems in education. According to the AI Act, high-risk AI systems, which include those used in educational or vocational training, must provide clear and understandable information about their purpose, as well as their capabilities and limitations. This includes information about the data used to train the system, the methods employed, as well as the potential risks associated with its use. The Act also emphasizes the importance of human oversight and the ability to contest or appeal decisions made by AI systems in educational contexts.

1.1 Problem statement and research motivation

Personalization in education has long been a focus of pedagogical research, driven by evidence that tailored learning experiences can improve learner's engagement, as well as their motivation, and thus learning outcomes (VanLehn, 2011). In the context of VET, personalization becomes even more critical due to the specialized nature of the skills being developed and the diverse contexts in which learners operate. For example, a construction worker undergoing safety training may require different instructional materials and scenarios than an industrial technician working in a controlled manufacturing environment. Generic training programs often fail to account for such variations, leading to suboptimal learning outcomes and a lack of preparedness for real-world challenges.

Modern educational technologies, including recommender systems, have the potential to address these limitations by leveraging data-driven insights to deliver personalized learning experiences. However, achieving effective personalization requires a nuanced understanding of the factors influencing each learner's journey. This includes not only their immediate learning needs but also their long-term goals, prior knowledge, and the specific challenges they face in their professional or academic contexts. Context-aware recommender systems are uniquely positioned to address these requirements by integrating diverse data sources and applying advanced algorithms to generate tailored recommendations.

Moreover, personalization is not limited to content delivery; it also extends to the design of learning pathways. Adaptive learning systems are designed to dynamically adjust the sequence and structure of instructional activities based on the learner's progress and performance. This level of customization can be particularly beneficial in VET, where learners often need to acquire practical skills through hands-on experiences and context-aware learning content. By aligning learning pathways with the learner's context and objectives, personalization can enhance the relevance and effectiveness of training programs, ultimately improving both learner satisfaction and performance (Pardo et al., 2017).

While the benefits of context-aware and personalized learning systems are well-documented, their implementation still poses several important challenges. One major hurdle is the complexity of modeling and integrating contextual data into recommender systems. Contextual factors are often dynamic and multifaceted, requiring sophisticated algorithms capable of capturing and processing this complexity. For example, a learner's context may include both static attributes, such as their job role or educational background, and dynamic factors, such as their current workload or changes in their professional responsibilities. Developing systems that can effectively account for these variables requires advanced methods for both data representation and algorithm design (Adomavicius & Tuzhilin, 2015).

Problem statement and research motivation

This thesis investigates and develops a novel approach for integrating contextual learning factors structurally and algorithmically in the recommendation algorithm. In the proposed method for Context Awareness and Explainability through Knowledge Graphs (CARExKG), structural integration is designed based on the graphical nature of the data structure itself, namely the knowledge graph. Algorithmic integration of context factors is done by developing a new reward function for a graph-based Markov decision process (MDP), where the design of the reward function allows the context factors to adjust the values of the transitions from one learning state to another. This, in turn, contextually weighs the potential learning paths in the knowledge graph. Top-ranked learning paths are then recommended to the learner.

The success of context-aware systems in education and VET depends not only on their technical sophistication but also on their ability to gain the trust and acceptance of learners who receive the recommendation. Transparency and explainability are critical for ensuring that learners understand and trust the recommendations provided. For that purpose, Explainable AI (XAI) offers mechanisms to make recommendation processes more transparent by providing clear and human-understandable clarifications for decisions made by the AI system (Samek et al., 2017). In the educational context, this transparency reduces resistance to recommendations and encourages a deeper engagement with personalized learning paths. Despite the advancements in CARS and the growing emphasis on transparency, a significant gap remains existing in the literature regarding explainable, context-aware recommendation systems in education. While existing research has explored various aspects of CARS and explainable AI separately, there is a need for integrated approaches that combine context-awareness with explainability in educational recommender systems (Zhong & Negre, 2022). This gap is particularly evident in vocational education and training, where the complex interplay between learning contexts, industry requirements, and individual learner needs calls for sophisticated, yet transparent recommendation systems, which not only engage the learner in following the learning-path recommendation, but most importantly give them the agency over the decision of following that recommendation. Thus, the learner can utilize the explanation of a learning path recommendation to make an informed decision on adopting and following the recommended content. This approach to human-AI collaborative decision-making proved to have higher value in fields like education, where decisions made about learning paths are mostly long-term and time-consuming.

Explainability of AI-supported RSs, in this sense, has pedagogical implications, as it enables learners to understand the rationale behind specific recommendations, which establishes a sense of ownership and agency over their learning journey. For example, a learner using a context-aware recommender system for skill development in nursing might receive a recommendation to complete a module on pediatric care. An explainable system would provide a clear justification, such as highlighting the learner's previous performance in related topics or the relevance of pediatric care to their professional goals. By making the decision-making process transparent, explainable AI can build trust and

reduce learner’s resistance towards learning recommendations, ultimately improving the effectiveness of personalized learning systems.

Despite the importance of transparency and explainability, existing systems often fail to provide explanations that are both accurate and understandable to learners. Effective XAI frameworks in education need to balance technical depth with user-centric design, ensuring that explanations are not only correct but also meaningful to diverse learners in varied educational contexts. Addressing these gaps requires interdisciplinary research that combines the fields of educational technology, artificial intelligence, and pedagogy.

To address this challenge, this thesis completes the design of novel KG-based CARS by developing and implementing an explainability framework. Based on knowledge graphs and large language models (LLMs), the CARExKG method generates pedagogically founded explanations of the recommended learning paths. The explanation content is designed with pedagogy experts to ensure its soundness and usefulness to the learners in their corresponding contexts. This research investigated user-centered requirements for the explanation content, shape, and interaction model. Textual and visual explainability components were then developed with a focus on the goal of the recommendation and the relations among the elements within the recommended learning path.

1.2 Research questions and objectives

In order to break down the challenges in designing and implementing explainable CARSs in education and VET, three core research questions were formulated to cover the gaps identified in the literature in the field of explainable CARSs. The thesis aimed then to develop the methods and systems that answer those questions, as well as the experiments to evaluate those methods and systems. Research questions handed in this research are as follows:

RQ1. How can the integration of a learning context into knowledge graphs be achieved, structurally and algorithmically, to establish a structural foundation for the personalization and adaptability of learning paths?

The first RQ is meant to set the foundation for the development of the educational CARS in the CERExKG method. Despite the existence of several CARS methods, their compatibility with the pedagogical requirements in an educational environment remains limited. A part of that limitation stems from the infrastructure on which the RSs and their contextualization are based.

To that end, it is important to carefully define the nature of a learning context as well as a contextual factor. A learning context is not simply a value of a certain variable that is stored in a database, even though this context is usually represented by specific contextual factors. For example, if we consider the “level of parent education” as a

Research questions and objectives

contextual factor for a primary school student, then the representation of this factor in the database, e.g., using the value “high-school certificate”, is not sufficient, solely, for making a contextual recommendation. In fact, it is the dynamics between this contextual factor and the other learning context factors that influence the learning outcome, and thus affect how the recommendation should be constructed. Take, for instance, the factor “parent engagement” in the learning process. It is clear that not only the separate values of “level of parent education” or “parent engagement” influence the learner, but also the interplay between those factors, i.e., whether the parents’ level of education influences the level of their engagement in their child’s learning process. As learning context reveals this level of complexity, with interconnected contextual components, the data structure that represents this context requires the ability to represent both its individual factors and relationships. That representation of the learning context factors in the database is only then able to provide a solid and clear representation of the potential interplay among them. While relational data structures provide an efficient way of storing individual factor values, graph data structures offer a stronger potential to capture the connections between the different aspects of the learning setting, as well as its context factors.

That being said, the representation of the contextual factors in a data structure is still limited, in terms of affecting the learning recommendations contextually, if that representation is done without considering how the data structure is queried, i.e., how the information about context factors is retrieved. The RS algorithm itself needs to be designed in a way that integrates the effect of each contextual factor, as well as the effect of the dynamics among those factors, while generating a recommendation. Individually, each context factor has a level of importance, which can differ among different learning settings. Collectively, on the other hand, the values and weight of multiple context factors affect the recommendation differently when the learning context includes those factors together. This is why an algorithmic approach to represent and utilize multiple learning factors in one learning setting becomes important for CARS in education.

The first RQ in this thesis aims to investigate the definition of the learning context and contextual factors in the literature, and to develop the structural and algorithmic requirements for integrating those factors in the graph data structure and the recommendation algorithm, respectively. Based on these requirements, the frameworks and methodology for implementing the structural and algorithmic integration are designed and implemented, to set the foundation for the development of the RS of the CARExKG method, whose design is the goal of the second research question.

RQ2. In what way can graph-based learning path recommendations be contextually tailored to improve learning?

Integrating learning context in the data structures sets the foundation for developing the CARS algorithm. This development, however, has to comply with the pedagogical goals of learning personalization, i.e., go beyond general information like learning preferences. Depending on the educational setting, and especially in learning settings like the VET, the learning outcome is influenced by user-centric, yet pedagogy-relevant

aspects, such as the learning goal and learning time. The algorithm of a CARS needs to accommodate those aspects in the personalization approach, in order to generate a pedagogically-sound recommendation. This algorithmic integration complements the structural integration of learning context factors in the graphical database.

Unlike a one-step learning recommendation, constructing a learning path, composed of a sequence of learning elements, poses another challenge for the contextualization of the RS. This stems from several factors, such as the change in the learner's state along the learning progress, the number of possible learning paths that achieve one learning goal, and the coherence between the learning elements in one learning path. This indicates that the recommendation algorithm has to account for these progressive changes in the process of ranking possible learning paths, in order to select one path (or top-n paths) for a specific learner. This should not be separated from the contextualization process of the RS, which is then based on the algorithmic integration of the contextual learning factors developed in the first RQ.

Building on the network nature of the knowledge graph, the discovery of the best learning paths for a learner can be broken down into path-finding and path-ranking problems. Both of those stages are to be designed based on solid requirements from pedagogy, if the final path recommendation is meant to fit the educational setting well. This research question, therefore, focuses on the development of a recommendation algorithm that takes into account all those factors while being able to cover multiple learning contexts in different learning scenarios, such as those in VET.

RQ3. How can graph-based explainable AI be designed to reduce learner's resistance towards accepting correct learning path recommendations?

While explainable AI has a major focus on clarifying how AI systems work, educational stakeholders require a different level of understanding of the predictions of, e.g., a recommendation system. This comes in contrast to the needs of a technical stakeholder, whose main focus is the inner workings of the system itself and how a prediction is achieved, rather than the relation of the predicted output to the problem in hand and why this prediction is useful in a particular solution.

To achieve a deeper level of system explainability in education, the third research question in this thesis addresses the problem and approaches a solution for it by studying the requirements of educational stakeholders from XAI. Then, it prompts a thorough research to develop a framework that utilizes the context integration strategies introduced in the first and second questions, to generate user-centric explanations not only about the inner workings of the CARS, but also context-level explanations that clarify to the learner how their specific learning setting and knowledge state led to recommending a specific learning path. The explanation, in this sense, focuses on clarifying three main aspects of the recommendation, from a pedagogical perspective:

1. The relation of the path recommendation to the learning goal.

Conceptual structure of the proposed solution to thesis RQs

2. The reasoning behind recommending each particular element in the learning path.
3. The relationship and similarities among learning path elements, which led to combining them together in one recommendation.

It is noticeable from this research direction that the educational stakeholder is meant to be at the center of the XAI framework and the CARExKG method presented in this thesis. This is due to the fact that the final aim of implementing XAI in education is not a technical one, but rather pedagogical and social. This means that XAI not only targets enhancing the technical understanding of the learners and educators of the technical system they are using, but most importantly, their ability to make and inform decisions based on the system's prediction. That ability has a strong influence on the way these stakeholders interact, accept, and utilize the intelligent algorithm's output, which is what this thesis uses to measure the effect of the explainability aspect of the CARExKG method on the user's perception of the learning recommendation, represented by their resistance to follow that recommendation, or, in contract, their acceptance of the recommendation.

The underlying goal of addressing this research question is to enhance the level of control and agency that an educational stakeholder, specifically a learner or an educator, has over the complex predictions and recommendations of the CARS. The hypothesis behind this research orientation is that having more agency over the prediction will enhance the learner's acceptance and adoption of the learning-path recommendation, which is measured in the VET context through their resistance to accepting a long, multidimensional path recommendation, which requires a rather long time and high-effort investment to complete since the training takes place alongside the normal professional responsibilities.

1.3 Conceptual structure of the proposed solution to thesis RQs

This thesis approaches the answer to the above-mentioned RQs on two levels:

1. The answer to each individual question, and
2. The answer to the three questions as a whole.

The answer to each question is meant to provide a scientific methodology to achieve the question's specific goal. The answer to the three questions as a whole, on the other hand, is expected to provide a comprehensive and coherent framework in which the requirements of each individual question are carried out and fulfilled by the answer to the other two. In other words, it aims at developing an integrative implementation framework that is capable of approaching a solution to a complex, real-world educational problem as a whole rather than a solution to individual parts of it.

In the center of the proposed solution lies the selection of the data structure, which not only serves the goals of storing data about the learner and the educational content but

is also required to provide the means for the effective discovery of information, which is needed for creating sufficient knowledge that leads to meaningful decision making.

A data structure that enables information discovery and knowledge building has to offer the means of connecting data points in a contextual and meaningful manner. These connections create more complex pieces of information, which, in turn, enable an efficient information retrieval algorithm to make sense of this information within a certain context, thus driving insights and building relevant knowledge from it. That knowledge is then the tool for making informed decisions, which is the overarching goal of all research questions in this thesis.

Therefore, the solution concept in the thesis is designed with the utilization of knowledge graphs as network data structures that allow defining multiple types of graph nodes and connecting those nodes through purposeful graph relations. While at the center of the proposed design, knowledge graphs require a surrounding group of comprehensive algorithmic approaches for handling the three main tasks, which are essential to the efficient use of the KG:

1. Algorithms to construct and complete the KG
2. Algorithms to retrieve data and information from the KG
3. Algorithms that enable reasoning over the KG structure.

Looking back at the RQs, it is noticeable that the first task is directly connected to answering RQ1, while the second and third tasks are both required to answer RQ2 and RQ3.

Figure 1 illustrates the conceptual structure of the proposed solution for the individual and combined RQs. Approaching an answer for RQ1 was done through a systematic study of learning context factors, to enable creating sufficient user and learning object (LO) profiles that are then modeled in the knowledge graphs. To embed those profiles in the KG, an ontological foundation is defined as a schema that allows graph nodes, relations, and data properties to represent user and LO models. Since the user model changes throughout the learning process, e.g., by acquiring new knowledge, the KG was designed with both dynamic and static layers. The dynamic layer represented the changing node types and was supported by update algorithms to track and record that change in the KG. Static KG layer included the node types that modeled fixed or slowly changing node types, such as the content of Los, and the domain-related properties.

KG construction and completion algorithms were developed to utilize embedding models for the numerical representation of textual data in user and LO profiles. The KG completion algorithm had to be designed based on the pedagogical requirements from the knowledge modeling, including the approach of capturing and embedding contextual factors in the KG, so that their retrieval will support the learning recommendation algorithm in creating a contextualized learning path. Moreover, the KG schema also had to provide clear and interpretable information about the Los in a learning path

Conceptual structure of the proposed solution to thesis RQs

recommendation, thus enabling an understandable explanation of the recommended path. For this purpose, KG data properties and relations included explainability-oriented classes, such as similarity reasons, which were calculated alongside semantic relation extraction between the LOs and stored as a semantic-relation property to be called by the explainability algorithm and fill specific information slots in the final textual explanation template.

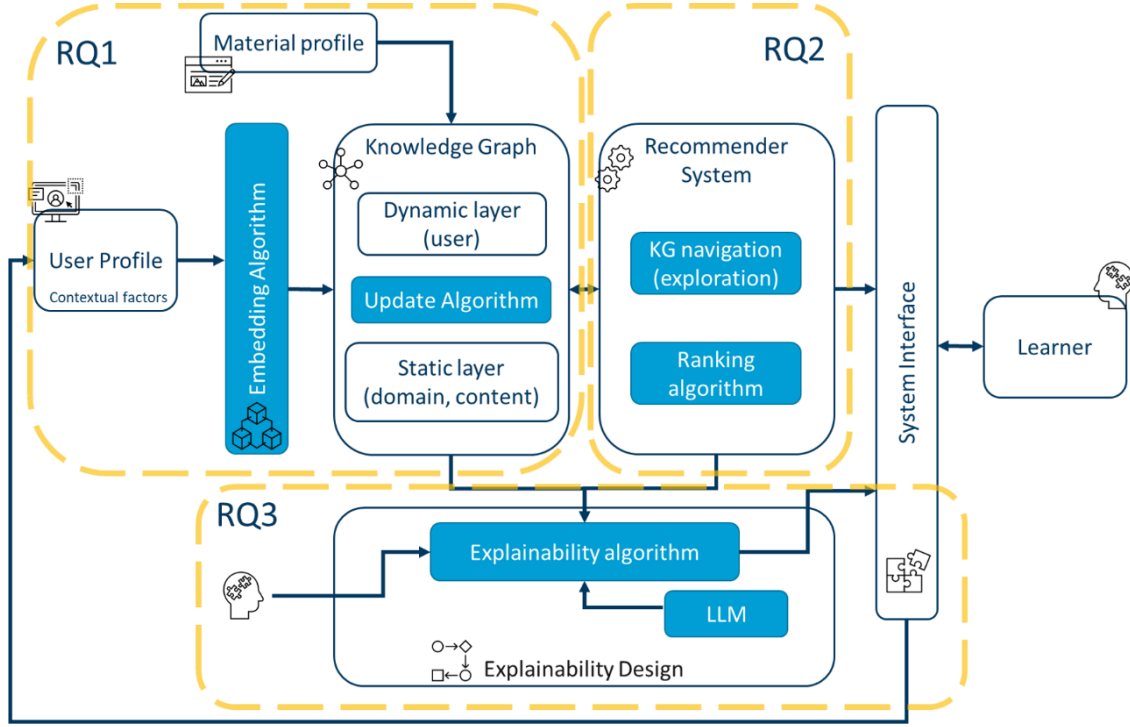


Figure 1. Conceptual design of the solution for the research questions.

The answer to RQ2 was designed as a Markov Decision Process (MDP), which utilized the KG structure as the environment, in which graph path finding and ranking was conducted. MDP is selected to approach the answer of RQ2 for two main reasons:

1. It offered a controlled and structured mathematical formalization to describe the environment where decisions are made sequentially, thus leading to a learning path recommendation, rather than a one-step recommendation.
2. MDP's definition of the environment, in which the reinforced learning can take place, is based on defining states and transitions between those states. This definition aligns strongly with the KG as a data structure, allowing the direct use of the KG as the environment of the MDP, in contrast to creating a separate environment solely for enabling the implementation of MDP.

In the KG, MDP environment states were learning objects that a learner has studied. They represented the current state of knowledge, which was aggregated to all previously visited states in the KG. The transition from one state to another was done by taking an "action", which materialized in the KG as the relation between two learning states. This compliance with MDP allows the KG to reduce the complexity of defining a separate

learning environment and increase the potential of navigation possibilities as a result of the network nature of the KG.

To allow creating contextual learning path recommendations with MDP over the KG, MDP's reward function is used to reflect the role of context on the selection of the next learning state in the path recommendation. In this sense, a virtual agent that is exploring the KG will reward the actions that lead to learning states corresponding to the learner's context, which states that have less relevant to the learner's context will not lead to increasing the overall path value. It is worth mentioning that this mathematical formulation of the path recommendation algorithm problem aligns with the learning theories, in which learning is represented as an accumulation of knowledge and experiences, such as in Kolb's experiential learning (Kolb, 1984), or Constructivism theory (Gogus, 2012). The creation of the path recommendation based on context factors aligns strongly with Student-centered approaches (Dewey, 1916) to learning design, as well as with the Situated learning approach (Lave & Wenger, 1991), which forms one of the main pedagogical foundations of this thesis.

The pedagogical foundation of navigating the KG in search of all possible learning paths and ranking those based on how they fit a learner's contexts, is the core of the reasoning process followed in this thesis to comply with the pedagogical requirements of educational technologies. This pedagogy-founded reasoning over KGs, using the mathematical formalization of MDP, is the method proposed in this thesis to approach solving the level of complexity that educational CARS encounter in the context of VET. Moreover, this contribution aligns with the ever-growing demand for interdisciplinary co-design approaches of education RSs, where pedagogical expertise accompanies the technical one from the very first stages of designing learning path CARS, in contrast to being only the user of CARS.

The answer to RQ3 is not separate from this pedagogical foundation but rather follows the same logic for creating a pedagogy-founded explanation of the inner workings of the proposed MDP and the resulting CARS. Situated learning theory is particularly utilized for the answer that this thesis proposes for RQ3. It emphasizes the role of the situation, in which the learning takes place, in the creation of the path recommendation and then its explanations. This situation is represented by the contextual factors, which include, among others, the goal of the learning, the problem at hand, and the knowledge state of the learner. Utilizing those factors, and the input from pedagogy experts on what the learner needs to know about the recommendation, the proposed explainable IA approach is governed by solid guidelines on "what to explain" and "how to explain".

Under these two questions lies a complex process of designing sound explanations, in terms of their content, shape, representation, interaction method, level of detail, and complexity, as several of those factors have an influence on the stakeholder's acceptance of the explanation, and thus adoption of the path recommendation. Therefore, the proposed design process on XAI in this thesis is strongly connected to domain and pedagogy experts, by whom the guidelines, rules, and safeguarding requirements of the

Structure of the thesis

explanation generation are defined. The technical realization of constructing the explanation is built utilizing the KG structure itself, as a coherent and thorough approach, which aligns with the proposed solution for RQ1 and RQ2, while proposing on its own a novel explainability methodology which reduces the complexity of black-box explanations, and offer a transparent approach to generating contextual explanations, which goes beyond the basic level of knowledge to reach a contextual-knowledge level. In this level, explanations are tailored to the specific problem at hand, and the state of the learner themselves, in contrast to being generic, static explanations for all contexts.

This research approaches creating those contextually oriented, dynamic explanations by designing a comprehensive explanation template, in which specific information slots are defined. The complexity of information slots differs, ranging from basic knowledge to reflective one, going through contextual knowledge. The proposed framework then defined the roles of AI systems and humans in a collaborative process for creating a complex explanation. This is due to the limitations of AI systems in achieving human-level, reflective explanations, which are able to provide contextual information for the explanations effectively.

AI systems used for extracting contextual information for the explanation utilized a set of customized text mining and information retrieval pipelines, in addition to dedicated large language models for the augmented generation of text from the graph data structure.

1.4 Structure of the thesis

Following the three research questions in 1.3, this thesis is structured in three main parts, each focusing on one part of the overall proposed solution, and the proposed technical solution for the respective RQ. The evaluation strategy and the set of experiments conducted to assess the performance of the individual solutions and the overall performance of the framework are then introduced in the fourth part of the thesis. In the following, a breakdown of the thesis chapters within those parts is presented:

1.4.1 Chapter 2: Theoretical Foundations

In this chapter, the foundation of the technical methodologies and learning theories used in this thesis are demonstrated. By analyzing fundamental and recent literature from standard, high-quality research databases, the chapter introduces CARS, along with their classifications and context-capturing approaches. Then, knowledge graphs are discussed as data structures and knowledge sources, pointing out their use in educational technologies, and for context modeling in CARS.

The reasoning over KGs is introduced in chapter 2 through MDP in detail. The chapter presents the mathematical formulation of MDP, its properties, the selection of environment states and actions, as well as the role of the reward function in the sequential

decision-making process. Graph-based MDP is discussed then as a method for defining the MDP environment.

The chapter also sets the foundation for XAI in education, and points out its implementation strategies, theoretical and technical approaches, as well as its relations to pedagogy and the challenges it faces in educational contexts. The regulations of XAI use in education are also presented in this chapter, reflecting on the set of European Acts on AI, data, and data services, which promote and guide the implementation of XAI in moderate- to high-risk fields, such as education.

After the theoretical foundation chapter, the thesis is divided into four parts. The first three parts correspond to the order and main parts of the three research questions, in terms of investigating the definition and meaning of learning context, the development of contextualization methods structurally and algorithmically, and the development of the explainability methods in CARExKG. The fourth part covers the different experiments conducted to evaluate the CARExKG method, and the results achieved from that evaluation.

Part I: Learning Context and Algorithmic Contextualization

In this part, chapter 3 analyzes and demonstrates the selection and modeling approaches of the learning context in modern CARS. It serves as a foundation for the context-factor selection and modelling in the proposed CARExKG method, and aims to provide the first step into answering RQ1.

1.4.2 Chapter 3: Context Indication in Personalized and Adaptive Learning Recommendations. A Pedagogical and Technical Perspective

This chapter presents a pedagogical and technical perspective on context-aware learning recommendations. It includes a systematic literature survey that was conducted in this research, detailing the methodology for selecting and analyzing relevant context factors in education. The analysis identifies different domains and approaches of contextualization, connections to learning theories, and various dimensions of the learning context. The results and discussion section explores the origins of context factors, profile-based context dimensions, and approaches for defining and capturing context from technical and pedagogical perspectives. It also provides a new classification scheme for context factors, driven by the educational stakeholder's needs and requirements, along with a thorough analysis of the observations made from the scientific literature in this field.

Part II: Knowledge Graph Based Learning-Path Recommendations

The second part introduces the graphical data structures, and the algorithms developed in the thesis within the different use cases that were handled. It spans over two chapters, which introduce the KG construction strategies, and the graph-based recommendation algorithm developed in the CARExKG method. The chapters in this part complete the answer to RQ1, which part I begins, and then focus on answering RQ2 by demonstrating how the structural and algorithmic integration of context factors in the KG and the RS are proposed and implemented in this thesis.

1.4.3 Chapter 4: Knowledge Graph Construction for Context-Aware Personalized Learning

Chapter 4 focuses on the approaches followed in this thesis for constructing knowledge graphs for personalized learning applications. It describes methodologies for KG construction and expert integration in three distinct use cases. The three use cases covered in this thesis are: 1) lessons-learned recommendations for engineer training in chip design domain, 2) personalized learning in nursing vocational education and training domain, and 3) learning pathway recommendation in higher education. Each use case outlines KG construction methods, taxonomies, topic modeling, and relation extraction. The integration of user models and context factors in KGs is also examined, and their role in creating a transparent and interpretable system for explainability is described.

1.4.4 Chapter 5: Markov Decision Process (MDP) over Knowledge Graphs for Learning Path Recommendation System

In this chapter, the proposed MDP-based learning path recommendation system is introduced in detail. It describes the graph-based recommender system, navigation strategies for learning path finding, and the design of the MDP on the basis of the KG as an environment, both mathematically and algorithmically. Furthermore, the design and optimization of the proposed reward function in this thesis are presented and discussed, in accordance with its technical and pedagogical meaningfulness for the educational setting.

Part III: Explainable AI for Learner Agency over Learning Recommendation

This part covers the explainable AI methods and framework proposed and implemented in this thesis, to generate graph-based textual and visual explanations for learning path recommendations. Thus, part III focuses on answering RQ3 by investigating and implementing the infrastructures and user interfaces needed to provide user-centric XAI for supporting learner agency over the learning path recommendation.

1.4.5 Chapter 6: A Framework for Explainable Recommendation Systems over Knowledge Graphs

This chapter introduces TExKG, a framework for explainable information retrieval (IR) and recommendation systems based on knowledge graphs. The framework is designed as a part of the CARExKG method to offer a transferable, graph-based, and domain-agnostic foundation for creating textual and visual explanations. The chapter describes the proposed approach for multi-source knowledge graph construction in the framework, as well as the integration of domain experts and domain rules in the process of creating the KG and the explanations. Additionally, the chapter presents the integrative concept developed for designing and creating the final explanation from human and automated algorithms collaboratively. Chapter 6 demonstrates how the proposed TExKG framework is flexibly adapted and implemented in different use cases.

1.4.6 Chapter 7: Explainability for Learning: Supporting Learner’s Agency with Interactive, Conversational, and KG-based Explainability

This chapter explores the role of explainability, particularly through knowledge graphs and conversational AI, in supporting the learner’s ability to make informed decisions about the learning recommendation. It examines methodologies for utilizing large language models (LLMs) and strategies for contextualizing prompt engineering to enhance explainability in learning environments. Furthermore, the chapter describes the research on integrating human and AI systems in a collaborative process of creating the final explanation, addressing the limitations of automatic explanations in cognition-critical explainability tasks.

Part IV: Evaluation, Results, and Discussion

1.4.7 Chapter 8: Evaluation Strategy, Experiments and Results

This chapter presents the evaluation methodologies for the proposed CARExGK method, focusing on its three main components. It highlights the series of experiments that took place within this research, which were designed to evaluate the three main pillars of this thesis: the knowledge graph, the recommender system, and the explainability approach. The chapter presents the structural evaluation of knowledge graphs quantitatively and qualitatively, through centrality measures and graph quality metrics that are interpreted in educational contexts. Additionally, it describes the evaluation of recommendation systems through user studies in VET and higher education, with a focus on performance, explainability, and context alignment. Lastly, the chapter details the experimental validation of the proposed explainability methods, including the TExKG framework and the conversational explainability method. Experiments evaluating

Structure of the thesis

user acceptance of conversational explainability platforms in nursing VET and higher education settings are presented, along with their results.

1.4.8 Chapter 9: Conclusion

This chapter concludes the thesis by highlighting its theoretical and practical contributions, and reflecting on the challenges and research questions addressed in this work. The chapter also discusses the limitations of the research and outlines potential directions for future work.

2 Theoretical Foundation

This chapter explores the theoretical scientific foundations of the interrelated fields addressed in this thesis. It aims to provide a thorough understanding of the state of the art in these areas and to highlight their potential in shaping intelligent systems and their role in the overall proposed solution in this thesis.

The chapter dives into the core concepts, methodologies, and applications of personalized learning, context awareness in learning personalization, knowledge graph structures, and explainable AI for learning path recommendations. It begins with an examination of the principles of learning itself and continues to introduce the models underlying personalized and adaptive learning, followed by an exploration of knowledge graphs as a structural and computational enabler for learning personalization and explainability. Subsequently, the chapter addresses the theoretical and algorithmic aspects of recommender systems, focusing on context-aware recommendation approaches. Finally, the chapter examines Explainable AI, its importance in fostering trust and user agency, and the integration requirements in the educational domain and learning processes.

2.1 What is learning?

Learning represents a fundamental cognitive and social process through which individuals acquire, modify, or reinforce knowledge, behaviors, skills, or values. This is accomplished through various cognitive processes, such as instruction, experience, study, observation, or practice (Schunk, 2020). It is clear that learning is a multifaceted and complex phenomenon. The complexity of learning is also evident through the fact that human learning is examined across several disciplines, including, but not limited to, psychology, education, cognitive science, and neuroscience. Our understanding and conceptualization of learning has evolved considerably over the past century, with a spectrum of theoretical frameworks attempting to explain how learning occurs. For example, (Ertmer & Newby, 2013) identified major theoretical frameworks that illuminate different aspects of learning: behaviorism, cognitivism, constructivism, and

What is learning?

connectivism. These theories offer perspectives on how knowledge is formed, retained, and applied, as well as how educational strategies can optimize learning outcomes.

From a cognitivist perspective, learning involves the acquisition and reorganization of cognitive structures through which humans process and store information (Anderson, 2000). Constructivist approaches, as pioneered by Piaget and later expanded by others, imply that knowledge is actively constructed by learners through their experiences and interactions with their environment (Glaserfeld, 1995). Socio-cultural theories, stemming from Vygotsky's work, emphasize that learning is fundamentally a social process that occurs within specific cultural and historical contexts (Wertsch, 1985).

Modern educational research has increasingly embraced integrative approaches to learning, recognizing that effective learning involves cognitive, affective, motivational, and social dimensions (Sawyer, 2014). These multifaceted conceptions of learning acknowledge that individuals learn differently based on prior knowledge, learning preferences, motivational states, and contextual factors. Knud Illeris (Illeris, 2017) proposes a comprehensive model that describes learning as involving three dimensions: content (what is learned), incentive (motivation and emotions), and interaction (with the social and material environment). This multidimensional understanding of learning has significant implications for educational design, particularly in developing systems that can adapt to individual differences.

Moreover, educational research distinguishes between several types of learning, namely formal, non-formal, and informal, which differ in their structure, purpose, and context (Weber, 2016). Formal learning takes place within organized educational institutions with explicit learning objectives, structured curricula, and recognized certification processes (McGivney, 1999). Non-formal learning, while still somewhat structured, occurs outside traditional educational institutions and typically does not lead to formal qualifications. Informal learning happens spontaneously through daily activities and experiences, without explicit educational intent or structure (Eraut, 2004). These distinctions are not only taxonomical but also reflect fundamentally different approaches to knowledge acquisition and skill development that influence how educational technologies should be designed. Furthermore, this organizational perspective of learning is linked to the theory of human capital and its three components defined by (Blundell et al., 1999): the early ability (acquired or innate), qualifications and knowledge acquired through formal education, and skills, competencies, and expertise acquired through training on the job. Weber argues that learning becomes an investment in the education of the individual, in which a part of the current income is invested in favor of acquiring a higher income in the future, as well as a higher productivity of the individuals working in an organization (Weber, 2016).

Learning can also be categorized based on cognitive engagement levels, from surface learning (focusing on memorization and reproduction) to deep learning (emphasizing understanding and sense-making, and not to be confused with the algorithmic deep learning) (Marton & Saljo, 1976). This distinction has important

implications for educational design and education software development, as technologies that promote deep learning require different features than those supporting memorization or procedural learning. For example, educational software that supports students in retention tasks by presenting a pre-defined set of digital learning cards does not reach the depth of learning that contextualized software offers when considering the student's previous knowledge to recommend learning cards from a certain domain, or with a certain level of difficulty. Similarly, the distinction between self-regulated learning, where learners actively control their learning process (B. J. Zimmerman, 2001), and externally regulated learning has implications for the design of systems that appropriately scaffold learner autonomy, which may require, in some cases, personalized feedback on the student's answers, or explainability of the algorithmic recommendations.

2.1.1 Learning in formal education

Formal education typically occurs within institutionalized settings such as schools, colleges, and universities. It features structured curricula, systematic assessment mechanisms, and leads to recognized qualifications. In these contexts, learning is often organized into disciplines or subject areas, with progression through increasingly complex content guided by established pedagogical frameworks. Teachers typically serve as content experts and facilitators of learning, designing instructional activities and assessments to promote knowledge acquisition and skills development (Biggs & Tang, 2011).

Assessment in formal education tends to be summative and criterion-referenced, measuring achievement against standardized expectations (Black & Wiliam, 2009). These assessment practices are integral to formal education systems, serving both accountability functions and providing feedback to learners and instructors. Educational technologies developed for formal learning contexts must typically integrate with these established assessment frameworks while potentially offering new approaches to formative and continuous assessment.

Formal education has traditionally followed relatively standardized approaches, with limited accommodation for individual differences in learning needs, preferences, or pace. However, recent decades have seen increased attention to differentiated instruction and personalized learning pathways within formal settings (Tomlinson, 1999). The integration of digital technologies has expanded possibilities for personalization within formal education, enabling adaptive content delivery, individualized feedback, and flexible learning pathways that accommodate diverse learner characteristics.

2.1.2 Learning in vocational education and training

Vocational Education and Training (VET) differs from general education in its focus on developing specific occupational skills and competencies for the workplace. VET serves as a critical bridge between educational systems and labor markets,

Learning personalization

emphasizing practical application and industry relevance (European Centre for the Development of Vocational Training., 2017). While formal education often prioritizes theoretical knowledge, VET typically balances theoretical understanding with practical skills development through approaches that integrate classroom learning with workplace experiences.

VET systems vary considerably across countries, ranging from dual apprenticeship models that combine workplace training with school-based education (as in Germany and Switzerland) to more school-centered approaches (Pilz, 2016). Despite these structural differences, VET is consistently distinguished by its alignment with industry needs, competency-based assessment approach, and integration of work-based learning experiences. This distinctive orientation presents both opportunities and challenges for educational technology development, as systems must accommodate diverse learning contexts and integrate theoretical knowledge with practical application.

Learning in VET contexts often follows principles of situated learning, where knowledge acquisition occurs within authentic contexts that mirror real workplace environments (Lave & Wenger, 1991). This approach recognizes that professional competence involves not just technical skills but also tacit knowledge, professional attitudes, as well as socialization into communities of practice (Eraut, 2000). Educational technologies for VET must therefore support learning across multiple contexts, facilitate knowledge transfer between classroom and workplace, and promote the development of both explicit and tacit professional knowledge.

Recent developments in VET have emphasized the importance of developing not only occupation-specific competencies but also transversal skills like problem-solving, communication, and digital literacy that support career adaptability in changing labor markets (Calero López & Rodríguez-López, 2020). This evolution reflects the changing nature of work in knowledge-based economies, where routine tasks are increasingly automated while complex problem-solving and interpersonal skills gain importance. Accordingly, educational technologies for VET must support the development of both specialized occupational competencies and broader transversal skills that enhance employability across evolving workplace contexts.

2.2 Learning personalization

Learning personalization represents a paradigm shift from traditional "one-size-fits-all" educational approaches toward tailored learning experiences that accommodate individual differences in dimensions like prior knowledge, learning goals, and learning contexts. This shift is grounded in the recognition that learners vary in how they acquire knowledge, process it, and then apply it in their respective contexts. It is also evident that addressing these differences can significantly enhance learning outcomes (Sancenon et al., 2022). As educational systems increasingly adopt digital technologies, personalization

has become one of the central topics in educational research and practice, with important implications for vocational education and training.

The primary goal of learning personalization is to optimize educational experiences to match individual learner characteristics, thereby improving learning effectiveness, efficiency, and engagement (US Department of Education, 2017). By adapting content, pace, instructional approaches, and learning environments to individual needs, personalized learning aims to address the limitations of standardized approaches that may either overwhelm struggling learners or insufficiently challenge advanced ones (Prain et al., 2013). This optimization goal aligns with fundamental principles of instructional design and the educational theories that emphasize the importance of matching instructional conditions to learner characteristics, such as Vygotsky's Zone of Proximal Development (ZPD) (Vygotsky, 1978).

Beyond cognitive outcomes, personalization seeks to foster learner agency and self-regulation. By providing learners with appropriate choices and control over their learning process, personalized approaches aim to develop metacognitive skills and learning autonomy that support lifelong learning (B. J. Zimmerman, 2002). This emphasis on learner agency acknowledges that effective personalization involves not just system-directed adaptations but also developing learners' capacity to make informed educational choices based on their goals, preferences, and self-assessment.

Personalization also addresses motivational aspects of learning by connecting educational experiences to personal interests and goals. Evidence can be found in educational psychology that relevance and personal connection enhance motivation and engagement (Cordova & Lepper, 1996), potentially reducing dropout rates and increasing persistence. In VET specifically, personalization can better align training with both individual career aspirations and labor market demands, creating more efficient pathways to employment.

2.3 Learning personalization through recommendation systems

Recommender systems (RS) are defined as software tools and techniques that providing suggestions for items to be of use to a user (Ricci et al., 2022). They serve as information filtering systems that predict users' preferences and needs. RSs have become well-rounded tools, used across various domains, to help users navigate large amounts of information by suggesting relevant items based on certain user-related factors, such as their preferences, previous interaction patterns, as contexts.

Educational recommender systems (ERSs) represent a specialized application of RS techniques, focused specifically on supporting learning processes. In this sense, ERSs are recommender systems that have been designed for an educational setting, with a primary goal of supporting learners in finding appropriate learning resources or activities (Drachsler et al., 2015). Unlike general recommender systems, ERSs need to account for

Context-aware recommender systems (CARS)

pedagogical objectives, as well as align with learning theories and educational contexts, to ensure that the recommendations contribute meaningfully to the knowledge acquisition and skill development processes. The educational domain introduces unique requirements including consideration of learner knowledge states, learning goals, educational standards, and long-term learning trajectories.

In other words, ERSs consider not only what learners want but what they need from an educational perspective. These aspects are meant to ensure immediate engagement and long-term educational benefits (Askarbekuly & Luković, 2024). However, such balance requires sophisticated modeling of learning objectives beyond simple preference matching. This contextual challenge is accompanied with the fact that learning context are not static but rather change with time. The learner's level of knowledge differs along the learning path, and so are their needs and sometimes learning goals. A static use model is, therefore, unable to provide the ERS with sufficient information for creating long-term, progressive path recommendations. Considering static and dynamic contextual factors of the learning and their learning environment has significant impact to the relevance of the recommendation, especially in an educational setting, where the usefulness of a resource depends heavily on circumstantial factors (Adomavicius et al., 2022).

2.4 Context-aware recommender systems (CARS)

While recommender systems traditionally match users with items based on past behavior or explicit ratings. Context-aware recommender systems extend this by incorporating contextual information, i.e., any information that can characterize a user's situation during recommendation (Adomavicius et al., 2022; Dey, 2001) into the recommendation process. In other words, CARS adapt recommendations to factors beyond just the user and the recommended item, such as the time and place of interaction, the user's current goal or mood, or the social environment around the user. The motivation is that the relevance of recommended items to a user can vary across different contexts (Zheng, 2017; Zheng & Jose, 2019). For example, recommending audible learning materials to a student may differ if the student is learning in a quiet environment or in the outdoors. Therefore, CARS utilize information about this context to improve the relevance and timeliness of recommendations (Haruna et al., 2017; Unger et al., 2020).

2.4.1 Definition of “Context”

According to Merriam-Webster dictionary, the term “context” is linguistically defined as *“the interrelated conditions in which something exists or occurs”* (Definition of CONTEXT, 2023). Originally, the terms described the part of discourse, which surrounds a word (Definition of CONTEXT, 2023). This definition was later extended to include events in addition to words. In this sense, the context refers to the environment or setting in which something exists. Following this notion, a “contextualized” entity is one that is *“placed in an appropriate setting, one in which it may be properly considered”* (Definition of CONTEXT, 2023).

The literal definition of context is then interpreted and implemented differently, depending on the domain and the nature of the entity itself. For example, in pedagogy, a learner's context can describe the situation of that learner, in terms of their social connections (Vygotsky & Cole, 1978) motivation (Ames, 1992) or current level of knowledge (Kuger & Klieme, 2016), while a technical consideration of the same learner's context can be described with the sensory data that a mobile device collects from the learner's interaction with the learning material on the device (Alnuaim et al., 2016). In the literature, multiple definitions of the learning context can be identified, which are mainly influenced by their dependency on pedagogical foundations or technical ones. The overlap and differences between those definitions are not only on the wording level of the definition itself, but also on the implementation of context-capturing approaches in a real-world scenario. In an interdisciplinary domain that addresses both pedagogical and technical requirements of the learning process, such as technology enhanced learning (TEL), a comprehensive and unified approach to defining and utilizing learning context is essential, yet it is still rarely investigated in the literature (Mayeku & Hogrefe, 2017). When it comes to technical implementations of the context in a learning process, the meaning of contextual concepts like "time", "space" or "place" has been under-theorized (Pimmer et al., 2013). This, in turn, reflects on the approaches used for capturing these concepts, which will also not be theory- or pedagogy-driven, but rather technology-driven. In that case, the real pedagogical value of the resulting contextualization methods cannot be guaranteed.

In situated learning theory, the definitions of learning context extend from the learner themselves to include their "interactions" with the environment (Bredo, 1994; Hung & Chen, 2001). In this sense, knowledge is not an exclusive element in the learner's mind, but it is rather situated as a part of the activity or culture, where it has been developed, or where it is used (Pimmer et al., 2013).

A "meaningful context" in the field of vocational education and training is not separable from its socialization aspect, which refers to the cooperative character of work processes and their influence on the division of labor, e.g., to lead to cost reduction or digitization (Buchmann, 2022). This aspect of the context shows that other dimensions, such as the legal and psychological ones, are to be considered in addition to the pedagogical ones in defining the context. This definition has its origins in social constructivism and connectivism, where the role of social interaction and the technologies that enable it is key in creating the knowledge in the mind of the learner.

In contrast to the philosophical nature of context definitions in pedagogy, and its focus on knowledge construction in the learner's mind, context awareness in technology has been defined by some scholars as the ability of software and hardware to predict and propose services to a user based on analyzing their position, time, living patterns, surrounding space, bodily signals and vital signs (Ku, 2014). Learning contextualization in this sense refers to using sensors, e.g., of a mobile device, to collect information about

Context-aware recommender systems (CARS)

the learner to personalize and adjust the learning offers they receive to their current context.

While the two previous definitions build on the same merits of the general context definition, being the situation in which learning happens, it is clear that the focus of each definition is different. However, it is also important to point out here that contextual definitions are not opposites between pedagogy and technology, but rather spread over a spectrum of slightly different interpretations of “what the learning situation is” and “how it can be captured”. Between the two fields, personalized and adaptive learning (PAL) and TEL research has provided context definitions that address both technical and pedagogical requirements for contextual learning and personalization. A common context definition in this field is proposed by Dey (Dey, 2001) which addresses “any information that can be used to characterize the situation of an entity”, where the entity can then be a person, place, or object. The information in this definition refers to any element or piece of data that enables describing the condition or state of the entity (Gómez et al., 2014). In TEL particularly, Luckin defines the context as “the current situation of a person related to a learning activity” (Luckin, 2010).

In this thesis, the definition adopted from situated and subject-oriented learning theories (Buchmann, 2022), where context is defined as the situation in which learning happens. This includes the learner’s situation (location, previous knowledge, etc.) and the LO’s situation (type, length, implementation scenario, etc.). Context is described by factors that can relate two or more LOs to each other if they serve the same learning (Hemmler & Ifenthaler, 2022; Verbert et al., 2012).

2.4.2 Context capturing in CARS

Early work in CARS established three primary strategies for integrating context into recommendations: pre-filtering, post-filtering, and contextual modeling (Adomavicius et al., 2020; Haruna et al., 2017), see Figure 2.

- In pre-filtering, the system first restricts or adjusts the input data based on the current context (e.g., consider only similar learners or resources from the same context) and then applies a standard recommendation algorithm.
- In post-filtering, an initial recommendation list, which is obtained without considering the context, is re-ranked according to contextual relevance.
- Contextual modeling (also called integrated modeling) incorporates context directly into the model, e.g., by adding context factors as additional dimensions in matrix factorization or features in a learning-to-rank model.

Each approach of the previous ones has strengths and weaknesses. Pre- and post-filtering are relatively simple to implement with existing recommendation algorithms. However, they may discard useful data or apply ad-hoc adjustments. Integrated modeling can, in principle, exploit context more fully by learning complex interactions between

context, users, and recommended items, but it often requires more data and complex models to avoid sparsity (Adomavicius et al., 2020; Haruna et al., 2017). It is worth noticing that despite the advantages that contextual modelling methods offer in comparison to pre- and post-filtering, those advantages are themselves not consistent across all scenarios. The effectiveness of each approach can depend on the domain and data characteristics (Panniello et al., 2014). Thus, hybrid approaches are also present in the literature, which utilize adaptive methods that choose a strategy based on context type or quality of available information (Kulkarni & Rodd, 2020; Unger et al., 2020).

In addition to the integration strategy, CARS can be classified by how context is captured and used. Contextual information can be explicit, e.g. provided by users or sensed directly, like time or location, or implicit, e.g., inferred from user behavior or other data (Dey, 2001; Haruna et al., 2017). Early CARS focused on a few explicit context variables, such as location, time, or weather conditions, to adapt recommendations, which had the advantage of interpretability (Adomavicius et al., 2022). However, context can also include latent factors that are not directly observed. Modern approaches often employ machine learning to infer context or its effects from data (Haruna et al., 2017). For instance, sequential recommendation techniques treat a user's recent interaction history as a form of contextual sequence, predicting the next item by modeling the sequence of past items (Zheng & Jose, 2019). Similarly, AI models can learn complex context–interaction patterns, e.g. a model might learn that a student's “preference” for certain content is different when they study late at night versus in the morning, without being explicitly told about time-of-day as a context.

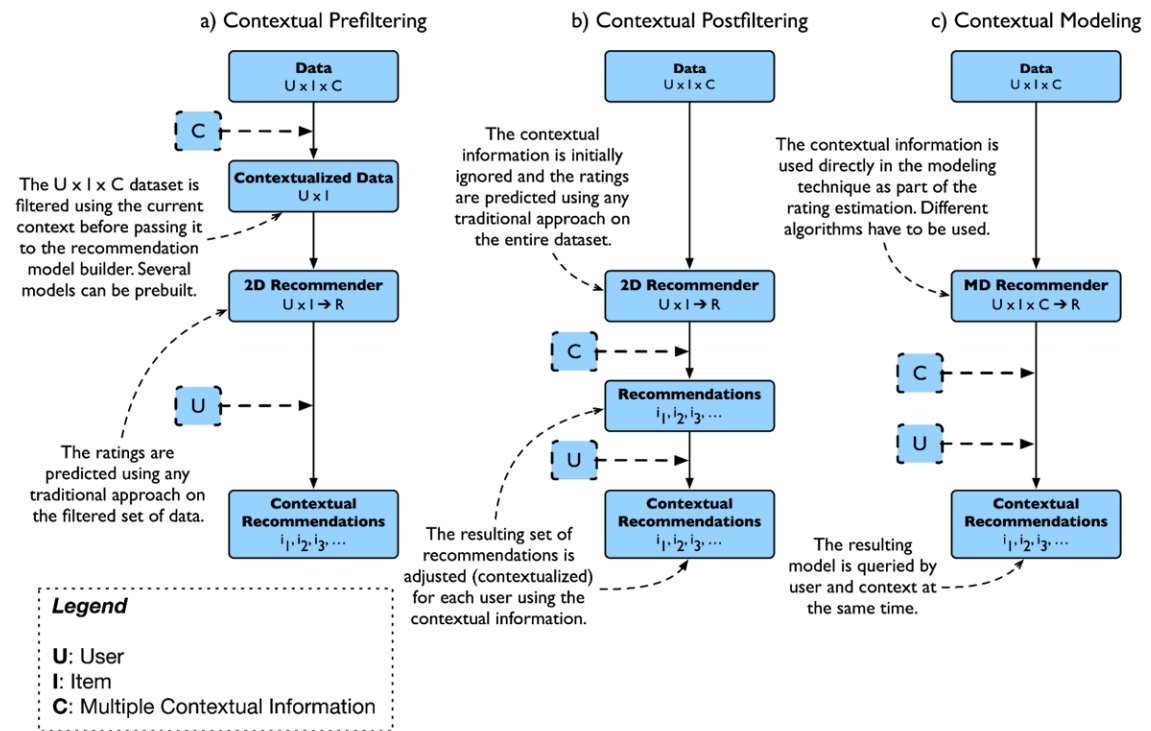


Figure 2. strategies for integrating context into recommendations (Adomavicius et al., 2022).

Context-aware recommender systems (CARS)

In general, context can be categorized into static and dynamic (Kulkarni & Rodd, 2020). Static context includes relatively stable user attributes, such as age, or some of the environment settings, such as course type, which do not change during a session. Dynamic context refers to factors that can differ with each interaction or between learning sessions. These include the temporal context, such as time, recency of last activity, spatial context, such as physical location, social context, such as the presence of peers in group learning settings, or the state of the user, such as their current goal or emotional state. Zheng and Jose capture this aspect of context in their definition, which describes context in CARS as “dynamic variables which may change frequently when a user is performing an activity,” including the dynamic attributes of a user or their environment (Zheng & Jose, 2019).

With the rise of deep learning and advanced feature modeling, recent CARS have moved beyond simple context filters. One of the modern directions is combining factorization methods with context. For example, context-aware matrix factorization and Factorization Machines (FM) treat context variables as additional factors in the factorization model, see Figure 3, enabling the model to learn interactions (user $\times$ item $\times$ context) triples as shown in Eq. 1.

$$R: User \times Item \times Context \rightarrow Rating \quad (1)$$

Wu et al. (2019) argue that representing user–item interactions in a context-dependent manner can better capture users’ varying preferences across items. They propose the model (CARL) that derives a joint representation for each user–item pair instead of independent user and item vectors, and then uses an FM to model higher-order interactions with contextual features. This yields a pair-based relevant information, i.e., context-aware information, that improves rating predictions and can be interpreted through attention mechanisms (Wu et al., 2019). Similarly, deep learning frameworks have been applied to learn latent representations of context. Unger et al. (2020) present a neural framework that incorporates context into neural collaborative filtering by representing contextual data in multiple ways (explicit features, unstructured signals, and structured latent factors). They demonstrate that using a structured latent representation of contexts, such as treating contextual variables collectively and learning an embedding for context state, outperforms both traditional context-aware models and other deep models, across several datasets. This indicates that capturing interactions among contextual factors, rather than using each context variable separately, can yield better recommendations. However, the interpretability and explainability of context modeling approaches that utilize machine learning and deep learning remain a challenge. Due to the black-box nature of these models, the reasoning behind considering a certain pattern, such as the time of learning, as a contextual factor, and interpreting how it affects the recommendation, is not information that can be easily acquired from the model. This, in turn, can result in serious limitations in an education use case, where understanding the role of each individual context factor may be essential for the stakeholder’s acceptance and engagement with the recommendation.

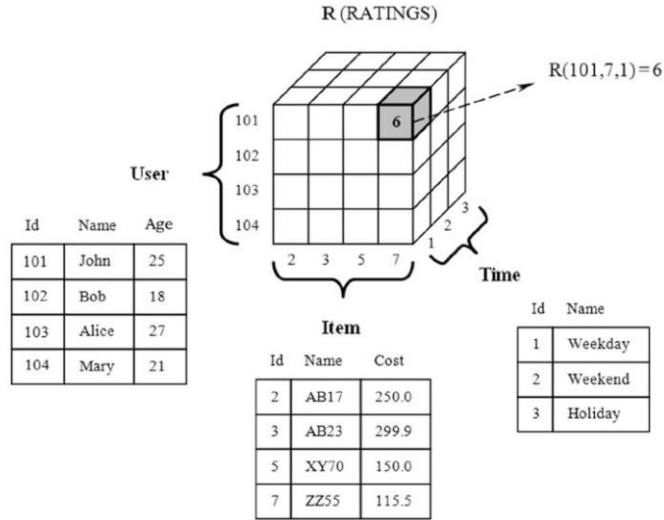


Figure 3. Multidimensional model for the User × Item × Time, where “Time” represents the context factor (Adomavicius et al., 2022).

One of the main approaches to tackle the explainability issue without compromising the aspect of modeling and connecting context factors is found in the recent development of using graph-based representations within CARS. Liu et al. (2022) note that many context-aware methods focus on improving the feature interaction layer of a model but overlook the embedding layer, which can suffer from sparsity of contextual signals. To address this, they construct a user–event graph connecting users, items, and “intent” nodes that represent combinations of contextual features, and apply graph convolution to propagate information in this graph, see Figure 4. This approach effectively learned richer embeddings for sparse context features by utilizing the graph structure itself, leading to improved accuracy in context-aware recommendations (D. Liu et al., 2022).

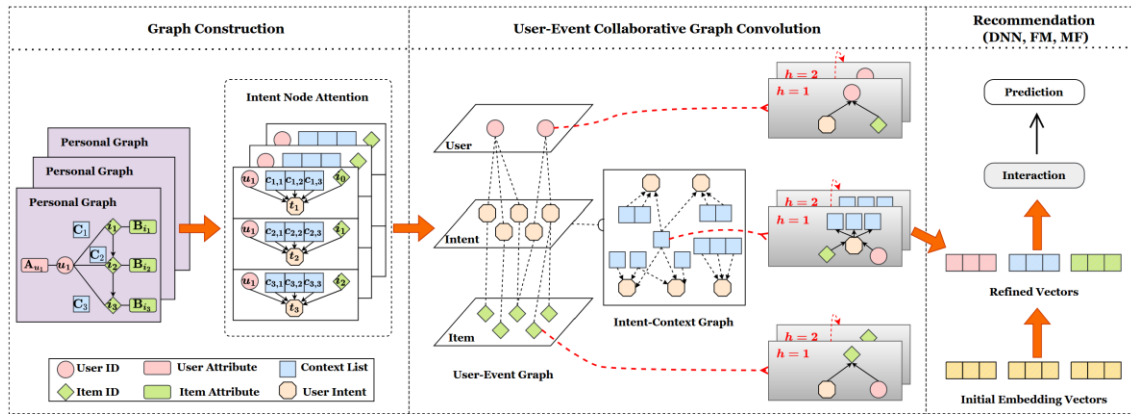


Figure 4. The architecture of the user-event graph embedding learning (UEG-EL) framework as proposed by (D. Liu et al., 2022).

2.5 Graph-based CARS

In educational recommender systems, knowledge graphs have a special role by representing formal educational knowledge and context in a structured way. Unlike

commercial domains where KGs often capture general item properties, in education, the KG typically encodes domain-specific pedagogical knowledge, including:

- relationships between concepts, such as prerequisite relationships or ontology of topics,
- relationships between learning resources and the concepts they teach,
- as well as relationships involving learners, e.g., which concepts a student has mastered, or which courses they have taken.

2.5.1 Knowledge graphs

Knowledge graphs represent powerful data structures for representing complex knowledge domains, particularly in contexts such as education, where context modelling requires defining and drawing relationships between educational concepts, learning materials, and learner characteristics.

A Knowledge Graph (KG), see Figure 5, is generally defined as a directed, labeled graph that encodes facts as relationships between entities. Formally, a KG can be represented as a set of triples, as in Eq. 2:

$$G = \{(h, r, t) | h, t \in E, r \in R\} \quad (2)$$

where E is the set of entities (nodes) and R is the set of relation types (edge labels). Each triple (h, r, t) is often denoted as subject–predicate–object. This triple represents a factual statement that *entity* h is related to *entity* t by relation r . In other words, nodes in the graph correspond to real-world objects or abstract concepts, and edges denote semantic or logical connections between these entities (On the Evolution of Knowledge Graphs: A Survey and Perspective). For example, a triple might be (“Course A”, *prerequisiteOf*, “Course B”), indicating a relationship between two course entities. This triple-based representation is common in the literature and underpins Semantic Web formalisms (e.g. RDF triples). It provides a machine-readable structure where each edge (relation) asserts a link between a “head” entity and a “tail” entity under a specific relation type. The collection of all such triples forms the knowledge graph, which can also be seen as a knowledge base organized in graph form (nodes as entities and edges as relations). Beyond the triple notation, KGs are often accompanied by a schema or ontology that defines the types of entities and relations. In formal terms, a KG typically comprises a *vocabulary* of classes (types of entities) and predicates (relation types), and an *instance set* of triples that instantiate those classes and predicates.

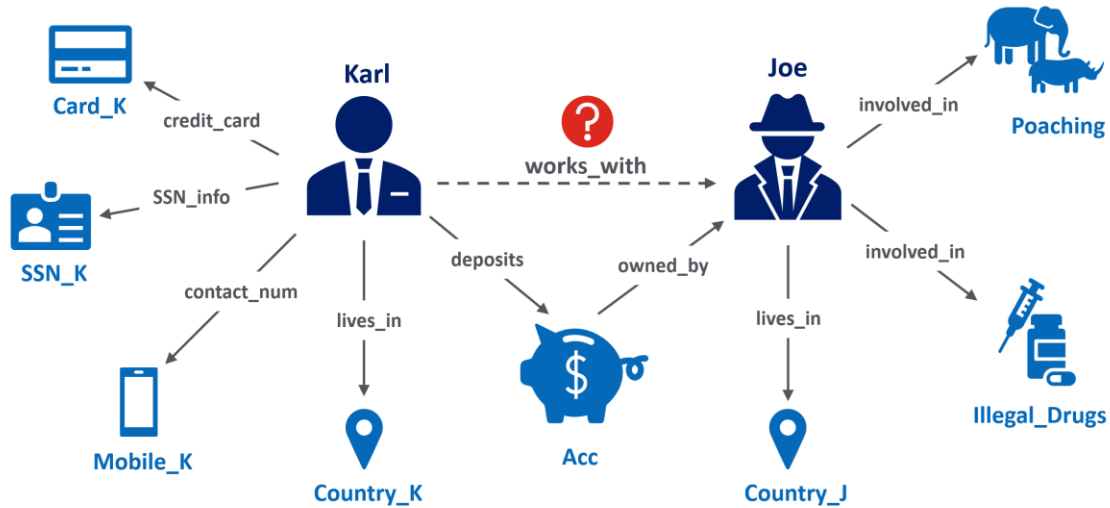


Figure 5. An example Knowledge Graph structure (Bhardwaj, 2022).

Construction of knowledge graphs

Constructing a knowledge graph involves both the design of a schema (ontology) and the population of that graph with factual data. A domain ontology provides the conceptual backbone of a KG – it defines the *classes* of entities, the *relationships* among those classes, and constraints or hierarchies (taxonomy) that organize domain knowledge. In practice, building a KG often follows one of two approaches (or a combination): a top-down ontology-driven approach, and a bottom-up data-driven approach:

- **Ontology Design (Top-Down):** Domain experts or *ontologists* first develop a formal ontology that captures the key concepts and relationships in the domain. This ontology serves as the schema or blueprint for the KG. For example, in an educational domain, the ontology might define classes like *Student*, *Course*, *Topic*, and relations like *enrolledIn*, *prerequisiteOf*, *teaches*. Ontologists carefully structure these definitions to reflect the domain’s logic, ensuring that entities and relations are well-typed and hierarchically organized (e.g., a *Course* might be a subclass of *LearningObject*). This step may involve using standards like OWL or RDFS to formally specify class hierarchies and property characteristics. The role of ontologists is crucial – by crafting the schema, they ensure the KG’s structure is semantically sound and unambiguous. A well-designed ontology prevents inconsistencies and provides a shared vocabulary for integrating data. However, if the ontology is incomplete or inaccurate, it can limit the KG’s quality (Abu-Salih & Alotaibi, 2024).
- **Knowledge Extraction and Graph Population (Bottom-Up):** With a schema in place, the next step is to populate the KG with instances (facts). This can be done by extracting entities and relationships from various data sources. In educational contexts, sources may include structured databases (student information systems, curriculum databases), semi-structured content (XML/JSON course descriptions),

or unstructured text (lecture notes, textbooks). Knowledge extraction techniques from NLP are often employed: for example, Named Entity Recognition to find key concepts, and Relation Extraction to identify relations in text (D’Souza & Auer, 2022; Jiang et al., 2025). Automatic construction pipelines can ingest documents and output triples that align with the ontology. For instance, an algorithm might read course descriptions and identify “Course B *requires* Topic X” as a relation to add to the graph. In some cases, existing knowledge bases (like Wikidata or DBpedia) are mined and mapped into the domain ontology. The process may be partially automated and *semi-supervised*: one study combined automated web data acquisition with manual annotation by experts to build a course knowledge graph, ensuring the resulting ontology and KG were rich and accurate (Abu-Salih & Alotaibi, 2024). This human-in-the-loop approach illustrates that while automation can gather large amounts of data, expert validation is needed to maintain correctness and relevance of the knowledge.

- **Integration and Refinement:** Often, KGs are constructed by merging multiple data sources and aligning them to the ontology. This might involve ontology mapping (aligning or merging ontologies from different sources) and entity resolution (ensuring that the same real-world entity is not duplicated as separate nodes). For example, if multiple universities provide their course data, an ontologist might map all their schemas to a unified ontology and then merge the instances. Tools and frameworks (graph databases, RDF triplestores) are used to store and manage the triples. Once the KG is populated, consistency checking and refinement are performed – e.g., applying logical reasoners to infer new relations or detect contradictions based on the ontology’s axioms. Throughout this construction process, ontologists and knowledge engineers play a supervisory role, curating the ontology and validating samples of the extracted triples. Their expertise helps address errors from automated extraction and adjusts the ontology when new concepts or relations emerge. In summary, building a KG is an iterative engineering process: define the schema (ontology), extract and import data as triples, and refine the graph with expert feedback. When done well, the result is a structured, interconnected representation of domain knowledge that can be used by algorithms for search and inference.

It’s worth noting that knowledge graph construction can be schema-based or hybrid. A schema-based approach uses the predefined ontology as a scaffold (a *top-down* method, as in many educational KGs that start from curriculum standards), whereas a hybrid approach may start with data (like text mining to discover candidate relations) and then have experts incorporate those into the ontology (a mix of bottom-up and top-down). In educational domains, many KGs are indeed *ontology-driven*. For example, EducOnto/EduKG was an ontology and knowledge graph developed for academic curriculum planning, modeling entities such as *student*, *curriculum*, *major*, etc., and their interrelations (Hubert et al., 2022). Other efforts use a bottom-up strategy: one study built a KG for a computer networking course by mining textbooks and Q&A forums, using text

analytics (LDA, TF-IDF) to extract concepts and relations, then storing them in a graph database (Abu-Salih & Alotaibi, 2024). Regardless of approach, the key components of any KG are the entities (with their types, as defined by the ontology), the relations between entities (structured by the ontology schema), and optionally attributes (data properties giving literal values for entity features). Together, these components form a rich network of knowledge. Crucially, the ontologist’s role ties it all together: by designing and overseeing the KG’s schema and content, they ensure the knowledge graph is a faithful and useful model of the educational domain.

User and learning object modeling in KGs

In educational recommender systems, knowledge graphs are used not only to represent domain knowledge (such as academic concepts and course content), but also to model users (learners) and learning objects (educational resources like courses, lessons, or problems). By embedding both learner profiles and learning materials into a unified KG, the system can capture the contextual relationships that are essential for personalized recommendations.

Learner modeling in KGs: A learner (user) can be represented as an entity node in the knowledge graph, with edges connecting that node to various pieces of information about the learner. This *learner profile subgraph* might include the learner’s demographic info, skills mastered, courses completed, preferences, and performance data. For example, a student node could be linked to concept nodes for topics the student has demonstrated proficiency in (*hasMastery*→ *Algebra*), to course nodes they have taken (*enrolledIn*→ *Calculus I*), and even to goal nodes representing learning objectives (*goal*→ *Data Science Certificate*). By structuring the learner’s data in the KG, we obtain a knowledge-rich profile that goes beyond a flat list of interests – the profile is interwoven with the domain ontology. This enables a finer-grained understanding of the learner’s state. For instance, if a student node is connected to a concept node “Photosynthesis” with a relation “*difficulty*” → *high* (indicating they struggled with that concept), the graph can propagate this information when recommending biology learning materials. KGs thus support modeling of learning trajectories: as a student progresses, new edges (e.g., *completed*→ *Lesson 5*) and possibly new nodes (new concepts encountered) are added to their subgraph. The sequence of learning events can be viewed as a path in the graph. In fact, a student’s entire learning history can be represented as a traversable path through content and concept nodes – effectively a personalized learning path within the KG.

Importantly, representing users in a KG allows capturing relationships *between* users and content in a semantic way. In a standard recommender, one might have a user–item interaction matrix; in a KG-based recommender, those interactions are explicit edges in the graph (e.g., *Alice* –[*completed*]→ *Course X*, *Alice* –[*interestedIn*]→ *Topic Y*). This opens the door to graph-based analytics on the learner profile. For example, one can apply graph algorithms or embeddings to compute similarity between learners based on the network of connections (two students with many shared concept nodes would have

similar knowledge profiles). Moreover, the explainability of recommendations is improved: if a course is recommended to Alice, the system might find a path in the KG explaining that recommendation (e.g., *Alice* $\rightarrow$ (*hasInterest*) $\rightarrow$ *Machine Learning* $\rightarrow$ (*isCoveredBy*) $\rightarrow$ *Course on Machine Learning*). Prior research has leveraged KGs to explicitly construct user profiles enriched with knowledge entities for better recommendations. For instance, knowledge-aware models have been proposed that “*explicitly construct user profiles leveraging knowledge and interaction graphs*” to improve recommendation accuracy and transparency. By incorporating domain knowledge (entities like skills or topics) into the user’s profile, the recommender can align the recommended content with the learner’s needs more precisely. In summary, KGs enable a holistic learner model: one that includes not only what the learner has done, but what they know, what they need, and how these relate to the domain’s conceptual structure.

Modeling learning objects: Educational content and learning resources (often termed *learning objects*) are also modeled within the knowledge graph. Each learning object (e.g., a course, lesson, tutorial, or quiz) is an entity in the KG, typically typed by the ontology (e.g., *Course* is a class of entities). These objects are connected to other entities to capture their context and structure. Key relationships include:

- Domain knowledge links:** A learning object may be linked to the knowledge topics or concepts it covers. For example, a course node “Calculus I” might have edges to concept nodes like *Derivative* and *Integral* (e.g., *teaches* $\rightarrow$ *Derivative*). This creates a mapping between content and the conceptual domain. Through these links, a KG can represent *which concepts a learning resource addresses*, enabling pre- and post-requisite reasoning. If “Calculus II” teaches *Integral* and requires *Derivative*, the KG can encode *Derivative* as a prerequisite concept node that connects “Calculus I” and “Calculus II” (perhaps “Calculus I” *prerequisiteOf* $\rightarrow$ “Calculus II”). Indeed, many educational KGs emphasize prerequisite relations among concepts or courses to model learning dependency. An *Educational Knowledge Graph (EKG)* often includes these prerequisite paths that students are expected to follow in their learning process (Aytekin & Saygin, 2024). For instance, an EKG of mathematics might show that the concept “*Algebra*” is a prerequisite for “*Calculus*”, which in turn is a prerequisite for “*Differential Equations*”. Such a graph provides a structured roadmap of the domain’s concepts and how learning objects connect to them.
- Learning object structuring:** KGs can represent the internal structure or classification of learning objects. Courses can be broken down into modules or lessons as part-whole relationships (e.g., *Course X hasPart* $\rightarrow$ *Lesson 1*, *Lesson 2*, etc.). Learning objects can also be categorized into types or subjects. A multidimensional KG framework may organize learning objects into several classes or layers – for example, separating resources into *Video*, *Text*, *Assessment*, and linking them through semantic relations (D. Shi et al., 2020). One study

designed a multidimensional KG for e-learning where *learning objects are stored in multiple classes and connected via semantic relationships*, to enable comprehensive learning path recommendations (D. Shi et al., 2020). This means the KG might have parallel subgraphs: one for knowledge concepts, one for course materials, one for assessments, etc., all interlinked. By doing so, the graph can answer complex questions like “find an assessment that tests a concept taught in this course” by traversing these links. Furthermore, converting traditional hierarchical structures (like a curriculum outline) into a graph can reveal additional context. In an earlier work, the author of this thesis demonstrate how transforming a hierarchical curriculum model into a knowledge graph of learning objects allows mining extra semantic relations between content items (Abu-Rasheed, Dornhöfer, et al., 2023). The resulting KG enriched the connections among learning objects (increasing graph connectivity and meaningful clusters) and provided a better foundation for recommending personalized learning paths.

- **Learner-object interactions:** The KG can also include edges representing interactions between users and learning objects (as mentioned above). For modeling purposes, these can be considered part of the user’s profile, but they also effectively link the “user subgraph” with the “content subgraph.” A learner’s *learning trajectory* can thus be traced as a path through the KG that alternates between user and object nodes: e.g., *Student Alice* –completed→ *Lesson1* –next→ *Lesson2* –enrolledIn→ *Course X*, etc. Some frameworks introduce the notion of a learning graph or learning path graph as a subgraph highlighting the sequence of content a student undergoes (Barlovits et al., 2022). For example, a *learning path* could be a directed path through the course prerequisite graph tailored to the student (including optional supplementary materials as side branches). In one approach, a *learning graph* was defined as a directed acyclic graph of learning tasks, where each node is a task and edges represent the recommended order; such graphs can accommodate alternative tasks (supporting or advanced) branching from the main sequence to adjust to learner needs (Barlovits et al., 2022). This concept, when instantiated for each learner, yields a personalized subgraph in the KG that represents both the curriculum structure and the learner’s journey through it.

2.5.2 Knowledge graphs in recommender systems

Knowledge Graphs (KGs) have emerged as a valuable resource for enhancing recommender systems in recent years. A knowledge graph is a structured network of real-world entities (e.g. people, items, concepts) and their interrelations, often described in a graph format (nodes and edges) with semantic labels (Hogan et al., 2021). In recommender systems, KGs are typically used as side information to enrich the understanding of users or items beyond what is available in the user-item interaction data (Guo et al., 2020). By connecting items to a broader context of related entities and attributes, a KG can help recommenders in multiple ways: improving accuracy via

additional signals, enabling explainability through semantic paths (“This book is recommended because it is about *Machine Learning*, which is a topic you showed interest in”), and alleviating cold-start by providing content-based features for new items or users.

There are broadly two categories of KG integration in recommender algorithms: path-based reasoning and embedding-based methods (Y. Li et al., 2024). In path-based (or connection-based) methods, the algorithm explicitly explores the graph structure to find connections between users and items. For example, given a user who liked item A, and knowing from the KG that A is connected to another item B through a sequence of relations (say, A is of genre Science Fiction, and Science Fiction is a subgenre of B’s genre), the system can infer that B might be a relevant recommendation. Path-based recommenders often utilize algorithms to search for meaningful relation paths or use measures like meta-path similarity in heterogeneous graphs (C. Shi et al., 2019). These approaches can naturally provide human-readable explanations (the path itself is an explanation) but may suffer if the graph is sparse or noisy. In contrast, embedding-based methods learn a latent representation for nodes (and sometimes edges) in the knowledge graph, which can then be integrated into a recommender model (Wang, He, et al., 2019; F. Zhang et al., 2016). For instance, each entity (item or concept) in the KG is embedded as a vector in a continuous space such that related entities have similar embeddings. The recommender can then use these embeddings as features in prediction (e.g. concatenating an item’s KG embedding with its ID embedding in a neural model) or to compute similarity scores. Many embedding-based approaches use Knowledge Graph Embedding (KGE) techniques like TransE, DistMult, or more recent graph neural networks to encode the graph structure. Some hybrid approaches combine both: for example, KGAT: Knowledge Graph Attention Network (Wang, He, et al., 2019) performs embedding propagation over the graph structure, effectively learning from multi-hop connections with attention mechanisms – blending path reasoning with learned embeddings. Overall, incorporating a KG allows the recommender to “comprehend the relationships between items” in a richer way than using interaction data alone (Troussas & Krouska, 2023).

The inclusion of KGs in recommenders has shown to boost performance in various domains such as movie, music, and e-commerce recommendations. By adding a semantic layer, KGs help to address cold-start issues, i.e., a new item can be recommended if the KG reveals it shares traits with items the user likes, and to improve coverage/diversity of recommendations, i.e., the system can find less obvious items connected via the graph. As Troussas and Krouska (2023) note: “knowledge graphs have been introduced in recommender systems as auxiliary data that can track changes in user interests and add insight to recommendations”. In dynamic domains, KGs can even track concept drift – for example, if a user’s recent interactions indicate a shift in interest, the KG can highlight new related topics to explore. Another advantage is explanation: many KG-based recommenders can easily generate explanation sentences by tracing the connections used for recommendation. This transparency is beneficial for user trust. However, using KGs also introduces complexity: constructing or obtaining a high-quality KG is non-trivial (it may require expert curation or information extraction from text), and large graphs pose

scalability challenges for real-time recommendation, hence the interest in efficient embedding techniques and graph neural networks. Nonetheless, the state of the art in general RS increasingly leverages knowledge graphs, as evidenced by surveys that categorize numerous recent systems using KG-based techniques for both rating prediction and top-N recommendation (Chicaiza & Valdiviezo-Diaz, 2021; Guo et al., 2020). Common benchmarks, such as the Netflix movie data enriched with DBpedia knowledge, show significant gains when knowledge graph information is used (Y. Li et al., 2024), e.g., when a system knows a movie’s actors, directors, genres via the KG, it can better match user preferences even with few user ratings.

2.5.3 Educational knowledge graphs

An educational knowledge graph can represent, through graph nodes, knowledge points, such as skills or concepts, and educational resources, such as lessons, problems, or courses. Through graph edges, KGs capture various relations among the knowledge topics and concepts in the graph nodes. For instance, a graph relation can define that “*concept A is prerequisite of concept B*”, “*resource X teaches concept A*”, or “*student Y completed resource X with score Z*”, etc.

By structuring this information as a graph, educational CARS not only can model and represent learning context structurally, but also utilize powerful graph-based reasoning to personalize learning path recommendations. For example, if the system knows via the KG that *Concept2* builds on *Concept1*, and a particular learner has not yet mastered *Concept1*, it can avoid recommending resources for *Concept2*, thus respecting pedagogical sequencing. This is an explicit use of context, here the learner’s knowledge state on prerequisite concepts, represented in the knowledge graph as nodes and relations.

For learning path recommendations, Shi et al. present a model that uses a multidimensional knowledge graph framework for e-learning to recommend learning paths. In their approach, the KG encapsulates multiple dimensions of context: the domain knowledge structure, learners’ performance data, and learning resource metadata. The recommender then searches this graph to find an optimal path of content tailored to the learner’s current knowledge and target learning outcomes. Empirical results showed improved learning efficiency when following KG-informed paths (D. Shi et al., 2020). This aligns with pedagogical strategies of structured learning progressions. Similarly, Troussas and Krouska propose a path-based recommender for learning activities that uses a knowledge graph of students and educational activities. Their model uses connection similarity of users and/or items in a student-activity graph to improve recommendations. Essentially, by analyzing the graph paths (e.g., student $\rightarrow$ concept $\rightarrow$ activity), the system finds activities that are connected to the student’s profile through learned concepts, mimicking how a teacher might choose the next exercise based on what a student has done so far. Notably, they include context like learning activity metadata and learners’ action logs in the graph, so the recommender can reason over rich contextual relations (Troussas & Krouska, 2023).

Knowledge graphs also facilitate concept prerequisite modeling, which is crucial in education. Khan Academy, for example, utilizes an internal knowledge graph of mathematics skills, where graph nodes are skills and edges indicate prerequisite relationships, to drive its mastery-based recommendation system (Y. Li et al., 2024). The KG enables the system to recommend practice tasks that logically follow from what the learner knows: if the learner is struggling on concept X , the system might recommend reviewing the prerequisite Y (if the graph shows $Y \rightarrow X$). This KG-driven adaptation ensures pedagogical soundness; the context of what the student should learn next is determined by the knowledge structure.

From the context-awareness perspective, KGs can explicitly model context as part of the graph. In a simple form, context variables, e.g. learning environment, or difficulty level, can be nodes or attributes in the graph, connecting to relevant entities. For example, a “difficulty” node might connect to all problems of that difficulty; a “classroom” node might connect all students in a class. This allows graph-based recommenders to naturally incorporate these contexts in their reasoning, e.g., when finding what “*students in this class*” are studying via traversing the classroom node. A more direct approach is the user-event graph as per Liu et al. (2022), where an event, such as a user interacting with an item under certain context, is a first-class element of the graph. Although Liu’s work was not specific to education, the concept can be applied through 1) modeling each study session or learning event as a graph node linking a student, a resource, and context descriptors (time, mood, etc.), 2) applying graph algorithms to derive embeddings that capture context-informed relationships (D. Liu et al., 2022). This can mitigate the sparsity of contextual data by sharing structure through the graph.

Another use of KGs in educational CARS is to incorporate learning domain content. For instance, an educational KG may contain detailed information about course materials (textbook chapters, videos) and their topic annotations. By linking learning resources through common topics or skills in the KG, a recommender can find alternate resources teaching the same concept, supporting diversification and remediation (if one resource was not effective, another covering the same concept can be recommended via the KG link). Li et al. (2024) describe CourseKG, a knowledge graph of course information (including concepts, objectives, and materials) aimed at “precision teaching”. While CourseKG is focused on organizing course knowledge, such a graph can feed a recommender that suggests to instructors or learners the most relevant resources for a given topic or the next course to take in a curriculum progression.

2.5.4 Reasoning over knowledge graphs

Reasoning over knowledge graphs refers to the process of deriving new insights, recommendations, or decisions by exploiting the graph’s structure and semantics. This broad area encompasses various computational approaches – from graph algorithms and query answering to machine learning techniques – that *use the KG to infer relevant information*. Below, several reasoning paradigms relevant to knowledge graphs are

outlined, and then a focus is put on how they apply to educational recommender systems to enhance personalized learning.

- Knowledge Graph-Based Recommendation Algorithms:** In recommender systems, KGs are used to improve the matching between users and items (or learning resources) by utilizing the connections in the graph. One common approach is graph-based similarity and path reasoning. For example, algorithms may perform random walks or path searches on the KG to find how strongly a candidate item is connected to the user's profile. If a student node is closely connected to a particular course node via a series of relevant relations (e.g., *Student* $\rightarrow$ *knows* *Concept A* $\rightarrow$ *is taught in* $\rightarrow$ *Course X*), the system can infer that Course X might be a good recommendation. Some algorithms materialize these connections using embedding techniques – they learn vector representations of entities and relations such that the graph's structure is preserved in a latent space. Then recommendations are made by computing proximity in that space (e.g., recommending items whose embedding is close to the user's embedding). Another strategy is rule-based or logic-based reasoning on the KG: using the ontology and logical rules to infer new links. For instance, if the KG knows that *Course X covers Concept Y* and *Concept Y is a required skill for Job Z*, a rule-based reasoner could infer a recommendation like “Course X is useful for achieving Job Z” for a career-guidance system. In education, recommendation algorithms often combine content-based filtering with KG reasoning: a system might recommend learning materials that are *prerequisite-connected* to what the learner has already mastered, effectively following the graph of domain concepts (Manouselis et al., 2011). The KG thus enables context-aware recommendations – beyond raw usage data – by providing semantic paths that connect user needs with content. Research shows that using knowledge graphs in recommendations can significantly enhance accuracy and also provide explanations; for example, a path “*Alice* $\rightarrow$ *Linear Algebra* (interest) $\rightarrow$ *isRequiredFor* $\rightarrow$ *Machine Learning Course*” can explain why Alice might need that course.
- Information Retrieval and Querying:** Knowledge graphs support powerful information retrieval through structured queries. Instead of keyword search, one can perform semantic queries (e.g., SPARQL queries on an RDF graph) to retrieve information that satisfies certain conditions in the graph. For example, a query might ask: “Find all courses that *teach* Concept X and are *recommended for* learners with skill level Y.” The KG can be queried to return precisely those courses, because it encodes relations like *teaches* and possibly a relation or attribute indicating required skill level. This is essentially *reasoning by graph query*: the system deduces an answer by matching patterns in the graph. In educational applications, this underpins semantic search and question answering systems. A QA system might use the KG to answer a question like “What should I learn before taking Course X?” by traversing *prerequisiteOf* edges to list the prerequisites. Likewise, an intelligent tutor might retrieve “related examples” by

finding content that shares common concept nodes with the current lesson. The KG’s ontology ensures that searches can respect the hierarchy (e.g., find resources teaching a broader concept if specific ones are not available). Compared to unstructured information retrieval, KG-based search can yield more accurate and context-aware results because it uses the relationships explicitly. For instance, KGs have been used to enable fragmented knowledge integration – linking small pieces of knowledge so that both students and teachers can navigate through connected concepts (Barlovits et al., 2022). In practice, this kind of reasoning improves learners’ ability to find relevant materials and answers, since the system can infer connections (like “Concept A is related to Concept B, so materials on A might help with B”) rather than rely on exact keyword matches.

- Reinforcement Learning (RL) and Markov Decision Processes:** Reinforcement learning provides a framework for sequential decision making, which is highly pertinent when the goal is to recommend not just one item, but a *sequence* of learning activities tailored to a student. In a reinforcement learning approach, the problem is modeled as a Markov Decision Process (MDP): a tuple (S, A, P, R) where S are states, A are actions, P are transition probabilities, and R is a reward function. In an educational context, a state $s \in S$ could represent the learner’s current knowledge state (the set of concepts mastered so far, for example), an action $a \in A$ could be recommending a particular learning object or exercise, and the reward $R(s, a)$ might correspond to the learner’s performance or knowledge gain after that recommendation (e.g., a positive reward if the student learned the concept, or based on quiz results). The KG comes into play by defining the state space and transition dynamics: since the KG encodes relationships like prerequisites, it can inform which next actions are valid or pedagogically sensible (you wouldn’t recommend a course if the prerequisite concept is not yet mastered by the student, which is a notion directly encoded in the KG). The reinforcement learning agent’s goal is to learn a *policy* $\pi: S \rightarrow A$ that maximizes cumulative reward – in other words, an optimal personalized learning strategy for each student. This is essentially automated tutoring: the system “reasons” about what to recommend next by simulating outcomes (with the help of the KG structure as a map of the domain) and learning from feedback. Several works have applied RL on knowledge graphs for recommendations. For example, in non-educational settings, the policy-guided path reasoning approach uses RL to find paths in a KG that lead from a user to a target item, ensuring recommendations have an interpretable path explanation (Verma et al., 2025). In the educational domain, researchers have begun to use KGs as the environment for an RL agent to navigate. One study proposes using a MOOC knowledge graph as the environment for an RL agent to generate explainable course recommendation paths (i.e. sequences of courses) for learners. The KG here serves as a state graph where each node or region could represent a set of mastered concepts, and the agent learns to move the student through the graph towards a goal state (mastery of a target set of skills). Empirically, such

approaches can yield better long-term outcomes: by planning a sequence with awareness of knowledge dependencies, the agent can avoid “dead-end” recommendations (too easy or too advanced content) and maximize the student’s progress.

- Markov Decision Processes & Educational Planning:** Even outside of learning a policy via RL, just modeling the educational recommendation problem as an MDP provides a powerful reasoning lens. For example, educational planners have used MDP formulations to compute the optimal next problem or lesson by lookahead search on a graph of knowledge states. The knowledge graph can supply the state transition graph: state defined by what a student knows, and edges (transitions) defined by learning a new concept or passing a course. By assigning rewards to states (say +1 for each concept mastered), one can use dynamic programming or simulation to recommend actions that lead to high-reward states (i.e., mastering many concepts). Early work by Durand et al. (2011) treated learning object recommendation as an MDP, tailoring sequences of activities for e-learning courses; more recently, Zhang et al. (2019) employed a form of hierarchical RL (which can be seen as operating on an MDP) to refine students’ course-taking sequences. The key advantage of the MDP/RL approach in the KG context is that it naturally handles the sequential and dependency-aware nature of learning – it’s a form of planning over the knowledge graph. The KG contributes the transition structure (what leads to what) and a rich representation of state (detailed profile of what the learner knows), while RL contributes an optimization technique to choose the best actions. When we zoom in on education applications, we observe that these reasoning techniques, empowered by KGs, directly enhance personalized learning and recommendations:
- Personalized and Adaptive Learning Paths:** KGs enable systems to recommend learning paths that are customized to the individual’s knowledge and goals. Because the KG encodes prerequisite relationships and topic connections, a reasoning algorithm can assemble a sequence of learning objects that respect these constraints and target the learner’s knowledge gaps. For instance, an algorithm might reason that “to learn concept C, the student should first solidify concept A and B” – a conclusion drawn by traversing the graph of concept prerequisites. Studies have shown that students following concept prerequisite pathways drawn from an educational KG have better learning outcomes than those who don’t (Aytekin & Saygın, 2024). By tailoring the path (which course or lesson to take next) to the student’s current state (as represented in the KG), the system implements adaptive learning. In essence, the KG reasoning acts like an expert teacher, ensuring the recommended material is neither redundant (already known) nor too advanced. One example is a personalized learning path recommendation system that uses a “cognitive graph” – a knowledge graph combined with cognitive state modeling – to suggest next steps; this system visualizes a student’s current knowledge structure and recommends appropriate subsequent concepts or

courses (Abu-Salih & Alotaibi, 2024). The result is a guided trajectory through the curriculum that is unique to each learner, continually updated as the learner progresses.

- Improved Recommendation Accuracy and Relevance:** By leveraging the rich context in KGs, recommender systems can move beyond superficial matching to *knowledge-level matching*. A KG-based recommender in education might know, for instance, that *Course X* covers topics that align with a learner’s interest or fill a gap in their skill set. This knowledge comes from the KG’s connections (e.g., learner’s known skills vs. course’s required skills). Therefore, the system can recommend *Course X* not simply because similar users took it, but because semantically it fits the learner’s profile. This often improves accuracy. Empirical research in various domains has found that incorporating KG side-information alleviates problems like cold-start (since even a new course can be recommended if its content concepts match the user’s needs) and sparsity (the graph provides additional paths connecting users to items). In education, where new learning materials or new learners with sparse data are common, this is valuable. Moreover, the KG can be used to diversify and explain recommendations: it might suggest a resource that complements the learner’s current activities (e.g., a practical project to apply a theory they learned), which a pure collaborative filter might miss. Knowledge-driven recommendation algorithms can thus reason about pedagogical utility (not just popularity).
- Explainability and Trust:** Reasoning on KGs often naturally yields explanations, which is crucial in educational systems. Both learners and educators are more likely to trust and adopt recommendations that come with understandable justifications. Because a KG encodes meaningful relationships, the system can explain a recommendation like “*We recommend Course Y because it covers Concept Z which is a gap in your profile*”. This explanation corresponds to a path in the KG: *Student* $\rightarrow$ (*lacks*) *Concept Z*; *Course Y* $\rightarrow$ (*teaches*) *Concept Z*. Many modern approaches explicitly focus on explainable recommendations by mining such paths. For instance, reinforcement learning methods that traverse the KG can output the path taken as an explanation (Verma et al., 2025). In an educational scenario, an RL agent might find an optimal path *Student* $\rightarrow$ *Module A* $\rightarrow$ *Module B* $\rightarrow$ *Certificate* and recommend *Module B* next; the sequence itself explains the pedagogical reasoning. The literature stresses that explainability is not only an added feature but a need to enhance learning. When students understand *why* a resource is recommended, it can increase their motivation and trust in the system (Q. Liu et al., 2024). Thus, KG reasoning contributes to a more transparent and intelligible personalized learning experience.
- Adaptive Assessment and Feedback:** Although not explicitly asked, it’s worth noting that reasoning over KGs can also improve adaptive assessment – deciding what question to ask or what feedback to give. For example, if a student makes an

error on a problem, the system can consult the KG to identify which prerequisite concept might be weak and then recommend a remedial resource targeting that concept. This kind of diagnostic reasoning is enabled by the concept graph linking problems to underlying knowledge components (skills). In intelligent tutoring systems, this is akin to *model-tracing* and *knowledge tracing*; a KG provides a structured knowledge model for these algorithms to reason over. Some systems use a form of MDP where states include the probability of mastery of each concept, and actions include giving hints or easier sub-problems targeting a concept – thus reasoning about interventions in a principled way to maximize the student’s mastery across the knowledge graph.

2.6 Foundations of Markov decision processes

Formally, an MDP comprises the tuple $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \gamma)$ where:

- $\mathcal{S}$: State space representing current graph position and context
- $\mathcal{A}$: Action space of permissible relation traversals
- $\mathcal{T}(s'|s, a)$: Transition probability to state s' from s via action a
- $\mathcal{R}(s, a, s')$: Reward function encoding path utility
- $\gamma \in [1]$: Discount factor for future rewards

In KG environments, states correspond to entity-relationship tuples (e_c, e_s, o) where:

- e_c : Current entity/node
- e_s : Source entity initiating the query
- o : Target objective (entity or relation)

Transition dynamics derive from the KG's connectivity patterns, with $\mathcal{T}(s'|s, a)$ representing the likelihood of reaching s' by following relation a from s . For educational KGs containing curriculum dependencies, transitions might encode prerequisite probabilities between course concepts.

The value function $V^\pi(s)$, representing expected long-term reward from state s under policy π , satisfies the Bellman equation:

$$V^\pi(s) = \mathbb{E}_\pi \left[\sum_{t=0}^{\infty} \gamma^t \mathcal{R}(s_t, a_t, s_{t+1}) | s_0 = s \right] \quad (2)$$

Optimal policies π^* maximize expected cumulative reward through recursive value updates:

$$V^*(s) = \max_a \left[\mathcal{R}(s, a) + \gamma \sum_{s'} \mathcal{T}(s'|s, a) V^*(s') \right] \quad (3)$$

where $\mathcal{R}(s, a) = \mathbb{E}[\mathcal{R}(s, a, s')]$. This formulation enables gradient-based optimization of reasoning paths through backward induction.

The defining characteristic of an MDP is the Markov property, which stipulates that the future state depends only on the current state and action, not on the sequence of states and actions that preceded it. Mathematically, this can be expressed as:

$$P(s_{\{t+1\}} | s_t, a_t, s_{\{t-1\}}, a_{\{t-1\}}, \dots, s_0, a_0) = P(s_{\{t+1\}} | s_t, a_t) \quad (4)$$

Value functions and optimal policies

$$Q^*(s, a) = \sum_{s'} P(s'|s, a) \times (R(s, a, s') + \gamma \left(\max_{a'} Q^*(s', a') \right)) \quad (5)$$

Once the optimal value functions are determined, an optimal policy can be derived as:

$$\pi^*(s) = \operatorname{argmax}_a Q^*(s, a) \quad (6)$$

2.6.1 Knowledge graphs as MDP environments

When a knowledge graph serves as the environment for an MDP, the components of the MDP are mapped to elements of the knowledge graph as follows:

State representation

In a knowledge graph-based MDP, states typically correspond to nodes in the graph, representing entities of interest. The state can be formally defined as a tuple $s = (u, e, h)$, where u represents a user entity, e represents the current entity node, and h represents the history of the path traversed so far. For educational applications, states might represent:

- Knowledge concepts or skills
- Learning resources (e.g., videos, articles, exercises)
- Learner profiles or knowledge states
- Educational activities or assessments

Action representation

Actions in a knowledge graph MDP generally involve traversing edges in the graph, moving from one entity to another. The set of available actions at a given state is typically defined as the outgoing edges from the corresponding node. Formally, the action space A_t at state s_t can be defined as:

$$A_t = \{(r, e') \mid (e_t, r, e') \in G \text{ or } (e', r, e_t) \in G, e' \notin h\} \quad (7)$$

where G represents the knowledge graph, e_t is the current entity, r is a relation, and h is the history of previously visited entities.

Transition dynamics

In a deterministic knowledge graph, transitions are governed by the graph structure itself. Given a state $s_t = (u, e_t, h_t)$ and an action $a_t = (r_t + 1, e_t + 1)$, the transition to the next state $s_t + 1$ is:

$$P(s_t + 1 = (u, e_t + 1, h_t + 1) \mid s_t, a_t) = 1 \quad (8)$$

Where $h_t + 1$ incorporates e_t and possibly $r_t + 1$ into the history h_t (Xian et al., 2019). In stochastic formulations, uncertainty may be introduced to model various factors such as learning variability or environmental noise.

Advantages of knowledge graph-based MDPs

Using knowledge graphs as environments for MDPs offers several advantages for educational recommender systems:

1. **Rich semantic representation:** Knowledge graphs capture nuanced relationships between educational concepts, enabling recommendations that respect the inherent structure of the knowledge domain.
2. **Heterogeneous information integration:** They seamlessly integrate diverse types of educational entities (e.g., concepts, resources, assessments, learners) and their interrelationships.
3. **Context awareness:** Knowledge graphs naturally represent contextual information such as learning environment, device characteristics, temporal factors, and social context, enabling more relevant recommendations.
4. **Cold-start mitigation:** Even when direct user-item interaction data is sparse, knowledge graphs provide rich semantic connections that can inform recommendations for new users or items.
5. **Multi-criteria optimization:** Knowledge graphs support balancing multiple educational objectives simultaneously, such as knowledge acquisition, engagement, and efficient learning pathways.

2.6.2 Reinforcement learning for MDP-based knowledge reasoning

Policy gradient methods have demonstrated particular efficacy in KG reasoning tasks due to their ability to handle large action spaces and sparse rewards. Consider an actor-critic architecture where:

Actor network $\pi_\theta(a|s)$:

Outputs action probabilities using graph embeddings of the current state:

$$\pi_\theta(a|s) = \text{softmax}(\mathbf{W}_2 \cdot \sigma(\mathbf{W}_1[\mathbf{e}_s \oplus \mathbf{r}_a])) \quad (9)$$

Where $\mathbf{e}_s$ and $\mathbf{r}_a$ denote entity and relation embeddings, respectively.

Critic network $V_\phi(s)$:

Estimates state value through temporal difference learning:

$$V_\phi(s_t) \leftarrow \mathbb{E}_\pi[r_t + \gamma V_\phi(s_{t+1})] \quad (10)$$

The advantage function $A(s, a) = Q(s, a) - V(s)$ drives policy updates via gradient ascent:

$$\nabla_\theta J(\theta) = \mathbb{E}_\pi[\nabla_\theta \log \pi_\theta(a|s) A(s, a)] \quad (11)$$

Sample efficiency improves through two key innovations from recent literature:

1. **Knowledge injection:** Pre-trained KG embeddings initialize network weights, encoding ontological relationships
2. **Hierarchical abstraction:** Multi-level policies handle curriculum dependencies (high-level) and concept transitions (low-level)

Experimental results on educational KG benchmarks demonstrate 38% faster convergence compared to vanilla RL approaches when using curriculum-guided exploration^{[1][6]}. The hierarchical architecture reduces effective action space dimensionality by factorizing decisions into:

1. **Macro-actions:** Select curricular units based on long-term objectives.
2. **Micro-actions:** Choose specific concept transitions within units.

This dual-level approach mirrors human pedagogical strategies while maintaining mathematical tractability through decomposed state spaces.

RL approaches for knowledge graph navigation

Reinforcement learning provides a framework for solving MDPs when the model (transition probabilities and rewards) is not known in advance. Several RL approaches can be applied to navigate knowledge graphs:

Value-based methods

Value-based methods learn a value function that estimates the expected cumulative reward for each state-action pair. Deep Q-Networks (DQN) extend traditional Q-learning by using neural networks to approximate the Q-function, enabling application to large

state spaces typical in knowledge graphs (Xian et al., 2019). The Q-function is updated according to:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_t + \gamma \cdot \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right] \quad (12)$$

Where α is the learning rate and γ is the discount factor.

Policy-Based Methods

Policy gradient methods directly optimize policy parameters θ by performing gradient ascent on $E[R_t]$, the expected return:

$$\theta \leftarrow \theta + \alpha \nabla_{\theta} E[R_t] \quad (13)$$

These methods can better handle the continuous state representations often used in knowledge graph embeddings and are particularly effective when the action space is large or continuous.

2.7 Educational CARS and pedagogical alignment

In educational settings, incorporating learning context in the learning recommendation is particularly crucial, as learning is a contextual and sequential process. Educational CARS considers not only who the learner is and what resources are available, but also the learner’s current knowledge state, learning history, objectives, cognitive/emotional state, and/or the situational factors of the learning environment. Context-awareness in this domain aligns naturally with pedagogical principles of personalization, adaptivity, and learner-centered instruction, which are key tenets in modern education (Brusilovsky, 2001; Chau et al., 2018). By tailoring recommendations to the learner’s context, an educational RS can ensure that the suggested activity is at an appropriate difficulty, relevant to the current learning goal, and delivered at the right moment to maximize impact on learning.

Several types of contextual information have been studied in learning recommenders:

1. Learner state and traits, e.g., the student’s prior knowledge, skill mastery level, learning style, or even affective state
2. Task or content context, e.g., the difficulty level of content, topic of the current lesson, or prerequisites needed
3. Temporal context, i.e., the timing of learning (study session length, time of day, recency of last activity)
4. Social context, whether the learner is working alone or in a group, interactions with peers or instructors
5. Environmental context, such as the device or learning setting (e.g., in a distance learning scenario, classroom vs. home).

Educational CARS and pedagogical alignment

Many educational CARS focus on learner-centric contexts, as understanding the learner's internal state is critical for effective personalization (Shute & Zapata-Rivera, 2012). For instance, learner ability or knowledge level is often treated as context so that recommended materials match the learner's proficiency. Ortegon-Romero and Krug-Wives (2021) observe a surge in systems that integrate Item Response Theory (IRT) models, a psychometric approach to estimating learner ability, into educational recommenders. In their systematic review, they classify such systems into three themes: identifying learning trajectories, predicting learning performance, and recommending educational resources. All three tasks benefit from context-awareness, where knowing a student's skill mastery (context) helps identify an optimal next step in a learning trajectory, and IRT provides a principled way to model this context. By combining IRT with recommendation strategies, systems can recommend exercises of appropriate difficulty, aligning with the pedagogical principle of adaptive sequencing (Ortegon Romero & Krug Wives, 2021). This approach is essentially an adaptation of adaptive testing for recommender purposes, ensuring that each learner's path is neither too easy nor too hard, consistent with Vygotsky's notion of the zone of proximal development.

Beyond cognitive ability, other contextual factors, such as personality, have been explored as context dimensions. For example, Gianotti et al. (2019) incorporate personality traits into an educational RS for recommending courses, aiming to align recommendations with individual differences in behavior. The rationale is that acknowledging stable learner traits (visual vs. verbal learning preference, introverted vs. extroverted tendencies, etc.) as part of the context can lead to recommendations that fit the learner's preferred way of learning, which is hypothesized to increase engagement.

Affective context is another example, where, e.g., Chanaa and El Faddouli analyze learners' affective and cognitive traits in a CARS using feature interaction techniques (factorization machines), essentially treating factors like motivation or attentiveness as contextual features (Chanaa & El Faddouli, 2022). Such affect-aware recommenders align with theories that emotional state can significantly impact learning; by recognizing when a learner is frustrated or bored, the system might adjust what it recommends, e.g., by offering a simpler task or a motivational prompt. However, accurately sensing affect is challenging, and these systems often rely on proxies (e.g. time taken on tasks, or self-reports).

Context of the learning task and environment also plays a role. For instance, Intayoad et al. (2017) use social context in an online learning recommender, utilizing information about a learner's social network and past collaboration to personalize course recommendations. The idea is consistent with social constructivist pedagogy: if we know a student's community or group context, we can recommend resources that have helped similar peers or even suggest peer helpers. Another example is context-aware mobile learning: Rodriguez-Hernández and Ilarri (2021) discuss mobile context-aware recommenders that adjust content delivery based on context like location and connectivity, ensuring recommendations are viable in the user's context, e.g., suggesting

a short quiz when the user has a 10-minute break. Temporal context in education can be aligned with spacing effects, where a system might recommend review materials when it detects that a learner hasn't practiced a skill recently, thus timing recommendations for better retention. Tarus et al. (2018) combine context awareness with sequential pattern mining to suggest e-learning activities. Their hybrid recommender learns common sequences of learning actions and uses context factors, such as current position in a sequence, to predict next useful resources. This sequential context approach aligns with the pedagogical strategy of scaffolding, where the next task builds upon what was previously done.

The benefit of context-aware educational recommenders is evident in improved personalization and potentially better learning outcomes. Many studies report that incorporating contextual factors (skill level, learning preferences, etc.) yields recommendations that are more relevant and useful to learners than one-size-fits-all suggestions (Haruna et al., 2017). For example, an adaptive learning path recommender by Carbone et al. (2021) uses context to dynamically build learning paths; it was shown to adapt to different student profiles and improve course completion rates. Moreover, context-awareness can support pedagogical goals like maintaining appropriate challenge and motivation. Zhou et al. (2021) demonstrate this in the domain of cognitive diagnosis: they model how “contexts and cultures have a direct impact on student learning by affecting the student’s implicit cognitive states”. By capturing educational context (such as country-level curricular differences and school environments) with a hierarchical attention model, their system more precisely diagnoses knowledge gaps and predicts performance, which in turn allows it to recommend targeted remedial content (Zhou et al., 2021). Interestingly, their analysis revealed differences in context impacts across countries, reflecting how pedagogical context (like national curricula or cultural attitudes) modulates learning. This underscores that context in education operates on multiple levels, from immediate situational factors to broader cultural settings, and effective CARS should consider these to truly align with the learner’s needs.

2.7.1 Challenges for CARS in education

Despite these advances, there are challenges and weaknesses in current educational CARS. One challenge is obtaining reliable context data. Unlike simple contexts (e.g. timestamp), educational contexts like “student’s current motivation level” or “prior knowledge of topic X” can be hard to measure objectively. Systems that rely on self-reported context (asking students how confident they feel, for example) may burden the user or get inaccurate data. Implicit inference (e.g. inferring knowledge from quiz results) can help but may introduce uncertainty. If context is misidentified, recommendations might backfire pedagogically (for example, suggesting an overly difficult problem because the system overestimated the student’s skill). Another issue is data sparsity and scalability: incorporating many context factors might slice the data into very fine segments (the “curse of dimensionality”), requiring large datasets for model training (Kulkarni & Rodd, 2020). This is problematic in education, where each student’s learning

trajectory is unique and datasets are not as large as commercial domains. Some studies caution that including too many contextual variables can hurt recommendation quality due to overfitting or increased complexity (Haruna et al., 2017). Therefore, a balance must be struck in selecting the most impactful context factors. From a pedagogical perspective, there is also a need to ensure that recommendations align with instructional design. A context-aware tutor might be very effective at keeping a student engaged with personalized suggestions, but educators must oversee that these suggestions still follow the curriculum and learning objectives. Involving teachers in the loop or using context that includes teacher preferences can be a way to reconcile algorithmic recommendations with pedagogy. Recent systematic reviews highlight the importance of integrating educational theories into the design of adaptive learning recommenders, so that the systems do not operate as black boxes but as tools guided by didactic strategies (Urdaneta-Ponte et al., 2021). Context awareness is thus a powerful approach in educational recommender systems, enabling personalization that is learner-centric and pedagogically informed, but it requires careful design to capture meaningful context and maintain trust and alignment with educators' goals.

Adomavicius et al. (2022) classify CARS challenges in education into four categories:

- **Algorithms:** Developing recommendation algorithms that effectively integrate contextual information into recommender systems to improve the quality of recommendations. This area has received significant attention, but continuous improvement is still a challenge.
- **Evaluation:** Conducting thorough performance evaluations of various context-aware recommendation techniques, including understanding their benefits and limitations under diverse conditions (such as accuracy vs. diversity trade-offs and domain-relevant metrics like economic or behavioral outcomes).
- **Engineering:** Designing architectures, frameworks, and systematic approaches to facilitate the implementation, deployment, and usability of context-aware recommender systems. Historically, this area has been less explored, emphasizing the need for new data structures, efficient storage methods, and modular system architectures.
- **Fundamentals:** Achieving a deeper understanding of the notion of context itself and effectively modeling context within recommender systems. This includes recognizing which contextual factors are relevant, meaningful, and influential enough across different users and scenarios

2.8 Explainable AI in Education

Explainable Artificial Intelligence (XAI) has emerged as a critical paradigm in educational technology, responding to growing concerns about algorithmic opacity and the need for transparent learning systems. This section examines the theoretical foundations and current state-of-the-art in XAI within educational contexts, providing a

comprehensive analysis of its definition, implementation methodologies, knowledge-based approaches, and ethical considerations.

Explainable AI refers to the capacity of artificial intelligence systems to elucidate their decision-making processes, enabling stakeholders to understand how and why particular outputs are generated (Rachha & Seyam, 2023). In educational contexts, XAI addresses a fundamental tension: as AI-powered systems become increasingly sophisticated in analyzing learning data, they often become more opaque in their functioning. This opacity creates what Khosravi et al. (2022) identify as a "black box" problem, where critical decisions affecting educational outcomes are made through processes that remain inscrutable to the very stakeholders they impact—students, educators, administrators, and policymakers.

The necessity for XAI in education extends beyond mere technical transparency; it addresses fundamental pedagogical, ethical, and regulatory imperatives. Educational AI systems—from intelligent tutoring systems (ITS) to learning management systems (LMS) and educational data mining (EDM) tools, increasingly make high-stakes decisions that directly impact learning trajectories and outcomes. When these systems recommend learning materials, predict academic performance, assess student work, or identify students at risk of failure, they profoundly shape educational experiences and opportunities (Khosravi et al., 2022). The opacity of such systems creates several interrelated problems:

- First, educational data contains inherent noise and complexity that complicates algorithmic processing. As Khosravi et al. (2022) emphasize that "learning data has many sources of noise, and reasoning about such data is noisy at many levels". This noise increases the risk of algorithmic misinterpretation and therefore requires transparent mechanisms to verify the soundness of AI-driven inferences. For instance, predictive models that categorize students as "at-risk" based on demographic variables may inadvertently perpetuate existing inequalities if these variables correlate with protected attributes like socioeconomic status (Lee & Kizilcec, 2020).
- Second, education inherently values metacognition—the ability of learners to understand and regulate their own learning processes. XAI systems align with this value by promoting what Bull and Kay (2016) term "open learner models" (OLMs), which enable students to visualize and understand their assessed knowledge states, fostering self-regulation and agency. These models transform AI from an opaque authority into a transparent partner in the learning process.

The development of XAI in education is further accelerated by evolving regulatory frameworks. The European Union's General Data Protection Regulation (GDPR) explicitly codifies a "right to explanation" for algorithmic decisions affecting individuals (Rachha & Seyam, 2023). Similar protections are emerging in educational policy globally, with frameworks like the UNESCO Beijing Consensus articulating ethical

principles for AI in education, including transparency and accountability requirements that necessitate explainable approaches.

Beyond regulatory compliance, XAI addresses what Khosravi et al. (2022) identify as the distinctive needs of education as a domain where human development and agency are central values. Educational systems must not only be accurate but also pedagogically sound, explaining not just what to learn but supporting understanding of how and why learning occurs. This aligns with established educational theories that emphasize the importance of metacognitive awareness in effective learning (Bull, 1997).

Empirical research demonstrates that explainable educational AI systems produce measurable benefits. Studies by Duffy and Azevedo (2015) reveal that students interacting with explainable intelligent tutoring systems show increased motivation and higher achievement in self-regulated learning. Similarly, Long and Aleven (2017) found that open learner models enhanced learning outcomes by supporting metacognitive monitoring and self-reflection.

The bidirectional relationship between explainability and educational outcomes is particularly significant. As Ogata et al. (2024) argue, educational XAI should encompass two essential components: "providing explanations to raise students' awareness of their learning progress and explaining recommendations to establish trust and motivate students to continue using the system". This dual focus recognizes that explanation in education serves both cognitive purposes (enhancing understanding) and affective purposes (building trust and engagement).

2.8.1 Adding explainability in educational recommender systems

Educational XAI methods span from inherently interpretable approaches to post-hoc explanation techniques. These methods reflect varying trade-offs between model complexity, predictive accuracy, and interpretability. Several comprehensive taxonomies have been proposed to categorize these approaches based on their temporal relationship to the modeling process (pre-model, in-model, and post-model) and their scope of explanation (global model understanding versus local instance explanation), see Figure 6.

Inherently interpretable models

Inherently interpretable models represent a "transparent by design" approach where the model's architecture facilitates direct human understanding of its decision-making processes. In educational contexts, several interpretable models have demonstrated efficacy:

- **Linear and logistic regression models:** operate by assigning weights to input features, creating a transparent relationship between inputs and predictions. In education, these models have been effectively deployed for predicting student performance in specific courses (Lakkaraju et al., 2015) and identifying factors

influencing degree completion (Aguiar et al., 2014). The simplicity of these models allows educators to directly observe which factors (e.g., attendance patterns, assignment completion rates, or prior knowledge assessments) most strongly influence predicted outcomes. This transparency facilitates data-informed interventions by identifying specific actionable factors.

- **Decision trees and rule-based systems:** represent knowledge through hierarchical structures that map decision pathways. These models excel in educational contexts where decision logic must be communicated to non-technical stakeholders. For example, Pardo et al. (2017) implemented decision trees to identify students at risk of academic failure, generating rule-based explanations such as "Students who submitted fewer than three assignments AND scored below 65% on the midterm have an 83% probability of failing." Such explicit rule structures facilitate both understanding and actionability. Similarly, rule-based expert systems have been deployed for educational tasks ranging from course recommendations to scholarship eligibility determinations (Engin et al., 2014), where transparent decision criteria enhance trust and system adoption.
- **Generalized additive models (GAMs):** extend linear models by incorporating non-linear transformations while maintaining interpretability. In educational applications, GAMs have enabled researchers to identify complex relationships between psychological traits and learning preferences. For instance, Dikaya et al. (2021) utilized GAMs to analyze correlations between students' interpersonal communication skills and their attitudes toward remote learning during the COVID-19 pandemic. The interpretable nature of GAMs allowed researchers to visualize how specific student traits (e.g., shyness or manipulativeness) related to remote learning adaptability, which are critical insights for designing personalized learning interventions.

Post-hoc explainability techniques

When educational systems employ more complex models (e.g., deep neural networks or ensemble methods), post-hoc explainability techniques provide mechanisms to interpret otherwise opaque models:

- **LIME** (Local Interpretable Model-Agnostic Explanations) generates simplified local approximations of complex models to explain individual predictions. In educational contexts, LIME has been applied to explain predictions about student performance by identifying the most influential factors for specific cases. Alwarthan et al. (2022) applied LIME to explain predictions of at-risk students in higher education, enabling advisors to understand which specific behavioral patterns or performance indicators triggered the system's concern. This local explainability facilitates targeted intervention strategies tailored to individual student needs.
- **SHAP** (SHapley Additive exPlanations) quantifies feature contributions to predictions using game theory principles. SHAP values have been utilized in

educational fairness audits to identify whether specific demographic or behavioral factors disproportionately influence system predictions.

- **Feature importance and saliency mapping** techniques highlight which input features most significantly influence model predictions. In educational recommender systems, these approaches visualize which aspects of student profiles or learning histories drive specific content recommendations. For instance, Aguiar et al. (2014) utilized feature importance metrics to explain why certain learning resources were recommended to students, highlighting whether recommendations were primarily driven by content similarity, performance patterns, or learning style assessment.

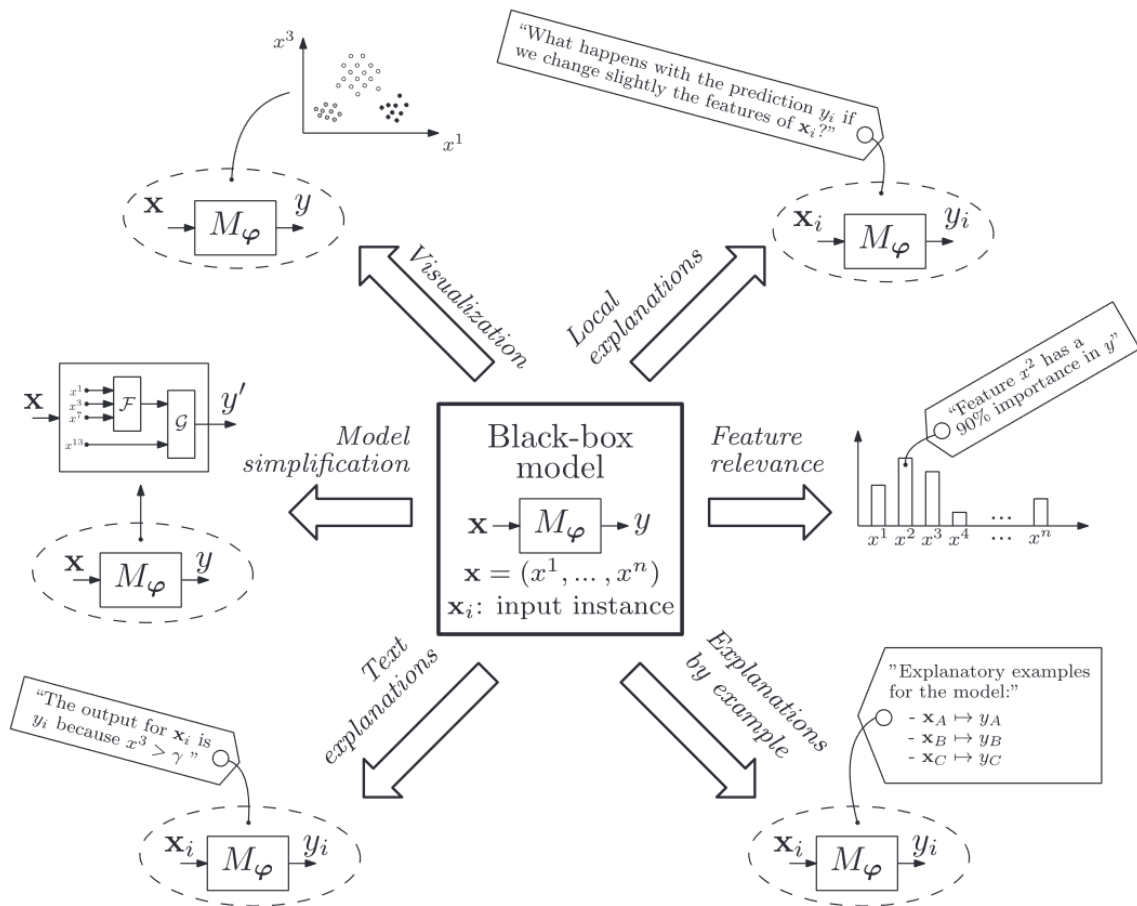


Figure 6. Explanation approaches for a black-ox model (Barredo Arrieta et al., 2020).

2.8.2 XAI application in educational recommender systems

Educational recommender systems present unique explainability challenges due to their impact on learning trajectories and resource allocation. Several approaches have emerged specifically for enhancing explainability in these systems:

Knowledge Component-Based Explanation leverages domain knowledge structures to explain recommendations in terms of pedagogical concepts. Dai et al. (2022) developed a model-driven quiz recommendation system that provided explanations based

on knowledge components, explaining recommendations in terms like: "This quiz is recommended because it covers concept X, which our system identifies as an area where you need more practice based on your previous performance on related questions". This approach connects recommendations to learning objectives, enhancing both transparency and pedagogical alignment.

Model Parameter-Based Explanation extracts interpretable information from model parameters to generate explanations. Takami et al. (2021) proposed methods to generate explanations from learner knowledge tracing models, explaining quiz recommendations based on specific parameters that represent mastery levels of different concepts. This approach bridges the gap between statistical modeling and pedagogical understanding by translating abstract parameters into meaningful learning insights.

Example-Based Explanation provides concrete examples that contextualize system decisions. Conati et al. (2021) incorporated clustering techniques to generate interpretable and personalized hints in exploratory learning environments, supporting explanation through exemplars rather than abstract rules. This approach aligns with cognitive science principles suggesting that humans often understand concepts through examples rather than formal definitions.

Comparison-Based Explanation contextualizes recommendations by comparing them with alternative options or reference points. In plagiarism detection systems, Foltýnek et al. (2020) implemented comparison-based explanations that highlighted similarities between submitted essays, providing visual evidence for potential academic integrity concerns. This comparative approach helps stakeholders understand system decisions relative to alternatives.

Research indicates that the effectiveness of explanation methods varies based on stakeholder roles, expertise, and the specific educational context. For instance, the study by Conati et al. (2021) revealed that the impact of explanations on student learning and perception varies based on individual characteristics, suggesting the need for personalized explanation approaches. Similarly, Abdi et al. (2021) demonstrated that complementing recommender systems with open learner models enhanced student engagement and perceived system effectiveness, indicating that explanation methods should be evaluated not only for transparency but also for their impact on learning outcomes and user experience. This user-centric approach for explainability and its evaluation is highlighted by Afchar et al. (2022), as shown in Figure 7.

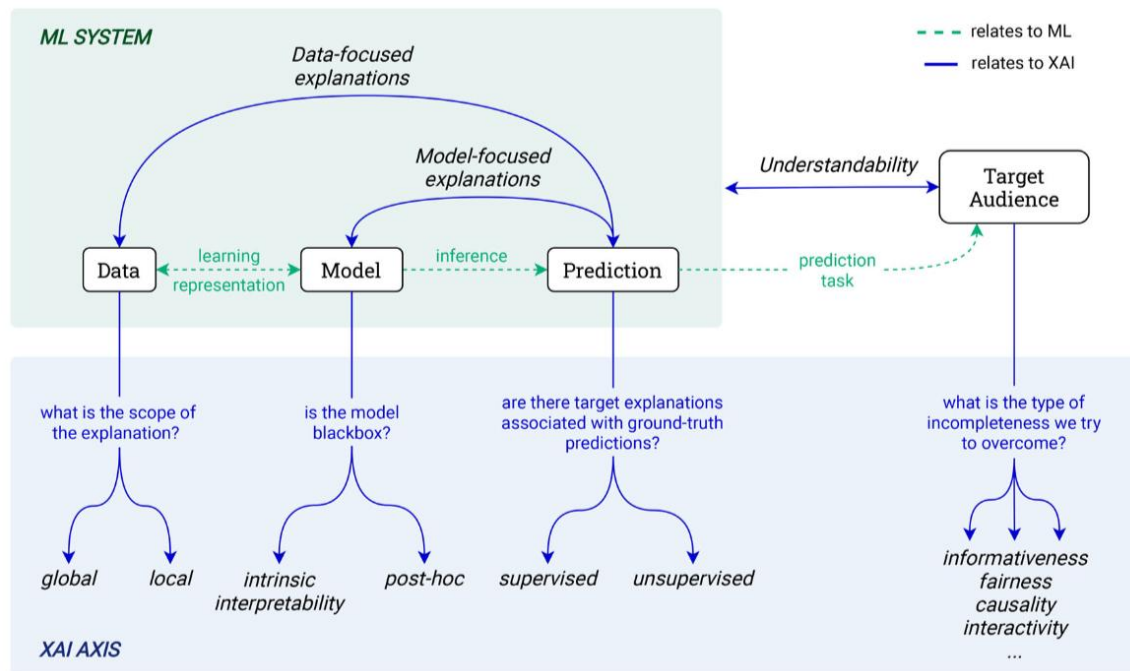


Figure 7. Algorithmic and user-centered explainability (Afchar et al., 2022).

2.8.3 Knowledge graph-based explainability

Knowledge graphs represent a potent framework for enhancing explainability in educational AI by structuring domain knowledge as interconnected entities and relationships. Unlike statistical models that function primarily as predictive mechanisms, knowledge graphs serve dual roles as both representational frameworks and reasoning systems that mirror human conceptual organization. In educational contexts, knowledge graphs formalize curriculum structures, learning objectives, and pedagogical relationships in ways that facilitate transparent AI reasoning and human comprehension.

The fundamental architecture of educational knowledge graphs typically encompasses three primary components: (1) concept nodes representing knowledge elements or skills; (2) relationship edges encoding pedagogical connections such as prerequisites, generalizations, or applications; and (3) attributes that qualify nodes or edges with additional information such as difficulty levels or learning outcomes. This structure creates what Khosravi et al. (2022) describe as a "semantically rich representation" that enables AI systems to reason about learning in ways that more closely align with human pedagogical thinking.

Several implementation strategies have emerged for leveraging knowledge graphs in educational XAI:

- Conceptual explanation generation utilizes knowledge graph traversal to construct natural language explanations of AI decisions. When a system recommends a learning resource or identifies a knowledge gap, graph-based reasoning can trace the conceptual pathways underlying this decision. For example, a system might

explain: "This resource on quadratic equations is recommended because your knowledge graph profile indicates mastery of linear equations (prerequisite) but difficulty with polynomial factorization (related concept)". This explanatory approach aligns with cognitive theories suggesting that human understanding is enhanced when new information connects to existing knowledge structures.

- Visual knowledge representation transforms abstract knowledge graphs into interactive visualizations that communicate AI reasoning to educational stakeholders. Abdi et al. (2021) implemented this approach by visualizing student knowledge states as graph structures where node colors indicated mastery levels and edge thicknesses represented relationship strengths. Such visualizations enable students to "see" their knowledge state from the system's perspective, facilitating metacognitive awareness and agency.
- Path-based reasoning explanation explicates decision processes by highlighting critical paths through knowledge graphs. Such path-based explanations can clarify recommendation processes by showing "why the questions were recommended to improve academic performance". This approach reveals not only what was recommended but the inferential chain leading to that recommendation, enhancing both transparency and pedagogical alignment.

The integration of knowledge graphs with other XAI methodologies creates particularly powerful explanatory systems:

- KG-enhanced open learner models combine knowledge graph representations with student-facing interfaces that promote self-regulation. Such integrated approaches enable students to "explore their knowledge gaps through navigable KGs". This interactive exploration transforms the learning process from passive consumption to active investigation of one's knowledge state.
- KG-augmented recommender systems enhance content recommendations with conceptual explanations derived from knowledge graph relationships. Barria-Pineda and Brusilovsky (2019) adopted this approach by complementing a knowledge-tracing model with concept-based explanations, helping students understand recommendations in terms of knowledge components rather than opaque algorithms. This integration enhances both the accuracy of recommendations and their explanatory power.
- Multi-modal knowledge representation combines knowledge graphs with other representational formats to provide richer explanations. For instance, the EXAIT system developed by (Ogata et al., 2023) integrates knowledge graphs with time-series analysis of student problem-solving processes, enabling explanations that connect conceptual understanding with procedural knowledge application. This multi-modal approach addresses the complexity of learning processes that span declarative, procedural, and metacognitive domains.

2.8.4 Limitations and research challenges

Despite their explanatory potential, knowledge graph approaches face several challenges in educational contexts (Khosravi et al., 2022):

The knowledge acquisition bottleneck presents a significant obstacle, as constructing comprehensive educational knowledge graphs requires substantial domain expertise and curation effort. Recent research has explored semi-automated approaches using natural language processing and machine learning to extract knowledge structures from educational materials, but human validation remains essential for ensuring pedagogical validity.

Personalization of knowledge graph explanations presents another challenge, as different stakeholders require varying levels of detail and perspective. Students may benefit from simplified graph visualizations focused on their immediate learning path, while educators might need more comprehensive views that situate individual student progress within broader curricular frameworks.

Dynamic knowledge evolution poses particular challenges for knowledge graph maintenance, as educational domains continually evolve with new research and pedagogical approaches. Systems must balance stability with adaptability to remain both trustworthy and current.

2.8.5 Ethical implications of explainability in education

Explainable Artificial Intelligence (XAI) in education offers significant advantages for enhancing learner agency and fostering trust in AI-driven recommendations. XAI serves as a crucial tool for addressing algorithmic bias and promoting fairness in educational technology. By providing transparency in decision-making processes, XAI enables stakeholders to detect and mitigate biases that may disadvantage certain student groups, particularly those from underrepresented backgrounds.

The implementation of XAI in educational contexts facilitates what Páez (2019) terms "functional understanding," allowing users to interpret AI decisions and comprehend the relationship between inputs and outputs. This understanding is essential for calibrating appropriate levels of trust in AI systems, preventing both uncritical acceptance (algorithmic deference) and categorical rejection (algorithmic aversion) of AI recommendations.

Furthermore, XAI approaches support the preservation of student and teacher agency in educational processes. By revealing not only system recommendations but also underlying parameters and assumptions, XAI empowers stakeholders to actively engage with and challenge algorithmic determinations. This aligns with Selwyn's (2019) concept of "exerting agency through AI," fostering critical engagement rather than mere compliance with AI-driven interventions.

However, the implementation of XAI in education also presents challenges, particularly in balancing transparency with privacy protection and system security. Ethical frameworks and mitigation strategies, such as human-centered design methodologies and differential explanation approaches, have emerged to address these concerns and ensure responsible deployment of XAI in educational settings.

Part I

Learning Context and Algorithmic Contextualization

3 Context Indication in Learning Recommendations: Perspectives from Technology and Pedagogy

3.1 Introduction

Learners pursuing the same educational objectives usually possess varying backgrounds, learning practices, preferences, and perhaps career orientations. While evidence indicates that individualizing educational experiences for specific learners or learner groups yields enhanced learning outcomes (Taylor et al., 2021), greater engagement and satisfaction levels (Rajabalee & Santally, 2021), and stronger commitment to lifelong learning pursuits (Ramírez Luelmo et al., 2021), the question remains on how to represent the learner's context in the personalization process.

To that end, human educators have always utilized their expertise and professional experience to understand their students' needs and backgrounds, and customize this learning experiences, as much as possible in the learning setting in hand. Digital educational platforms, on the other hand, and TEL approaches generally employ filtering and recommendation algorithms to accomplish personalized content identification or learning pathway construction for individual learners (Manouselis et al., 2013). Recommendation systems within TEL utilize available information about the learner and the content, to generate personalized recommendations. Given that learning always occurs within specific contexts (Brown et al., 1989), RSs approach incorporating contextual information automatically into the personalized path recommendation through a selection of a set of measurable and available context factors, which then are embedded in the user model, and affect the recommendation algorithm. Since the conceptualization of "context" exhibits considerable variation among researchers in scholarly literature, as examined by (Verbert et al., 2012), the identification of dimensions and factors considered by recommender systems to "capture" the learning context is dependent on these definitional differences. For instance, a context definition from a technical perspective will usually include factor such as the type of student devise, levels of lighting, or level of surrounding noise, while a pedagogical definition of the learning context is likely to include factors such as the level of parent's education, level of teacher

involvement, or the type of school. In other words, a context definition from a pedagogical viewpoint (Koubek et al., 2009), characterizes context as the dimensional aspects of situations experienced by students, encompassing not merely technological tools or demographic information but also ethical considerations, legal frameworks, and economic conditions.

Despite the overlap between these definitions and perspectives, the differences reflect domain-specific emphasis on factors, which characterize a student's learning context. In digital and e-learning platforms, pedagogical foundations derived from learning theories should inform recommendation system development to achieve educational objectives (Abu-Rasheed, Weber, Dornhöfer, et al., 2023). This necessitates consistent context definitions and factor indications that align with pedagogical requirements and contextual learning considerations (Isaias et al., 2022).

This chapter aims to provide a comprehensive examination of existing literature regarding contextual factors utilized within TEL domains, in the light of the pedagogical foundations from learning theories. It examines various learning context definitions adopted in literature during the eleven-year period from 2012-2022 within TEL and context-aware recommendation system domains. The pedagogical and technical foundations of these definitions are investigated by examining context conceptualizations and operationalization across different learning theories. Additionally, the chapter surveys contextual factors employed by researchers from both pedagogical and technical perspectives, highlighting their similarities, differences, and the implications of these factors.

To define the scope and objectives of this study concretely, research questions were formulated through preliminary review of existing literature and similar surveys, identifying knowledge gaps and areas requiring further investigation. The research questions that this study focuses on are as follow:

1. Which contextual factors appear in TEL and context-aware recommendation literature?
2. What categories do these factors belong to?
3. Which factors are derived from pedagogical foundations in learning theories?
4. Which learning theories are addressed in the surveyed literature?
5. What contextual information sources are considered (e.g., user, learning material, or learning environment context)?

In the following sections, the systematic literature review process the research methodology, the analysis of surveyed literature, and the synthesizes of the review results, patterns, and observations, are introduced. The chapter also addresses similar reviews in this domain.

3.2 Review methodology

To achieve the objectives of the review and answer the above-mentioned research questions, the systematic literature review methodology was employed following the framework and guidelines proposed by Vom Brocke et al. (2009) and Webster and Watson (2002). This framework and guidelines aim to ensure the rigor of the study, and the objective perspective during the analysis and synthesis of the review results. The review is conducted in this thesis as the foundation for adopting a comprehensive and consistent definition of the learning context, as well as a solid foundation for the selection of the contextual factors that are to be embedded in the knowledge graph and the recommendation algorithm, which are presented in the following chapters.

3.2.1 Literature search strategy and data collection

Search strategy was developed to ensure, to the best of our knowledge, a comprehensive identification and collection of all relevant literature to the review objectives. Comprehensive search was conducted using various databases, namely Google Scholar, ACM Digital Library, Springer Link, EbscoHost, and LearnTechLib. Search was limited to 2012-2022 period. Databases and search engines were selected based on topic relevance and comprehensiveness regarding high-quality literature coverage on technical and pedagogical domains as literature sources on learning context and factors.

Search terms were identified based on research questions and developed using keyword and controlled vocabulary combinations. Terms were selected to ensure retrieval of all relevant literature while minimizing risk of missing important studies. Selected search terms include:

- *((learner OR learning) AND context) AND (TEL OR e-learning OR (technology AND enhanced AND learning))*
- *((learner OR learning) AND context) AND (learning AND theories)*
- *((learner OR learning) AND context) AND (indicators OR factors OR dimensions)*
- *((learner OR learning) AND context) AND (modeling OR model)*

Each selected database possesses unique search engine. To accommodate terms in each database's search engine, search strategy was adjusted following technical requirements of that engine, ensuring results generated from different database searches follow same search-term-combination logic.

Following literature review taxonomy introduced by Cooper (1988), who presented six characteristics of the systematic literature review, this review categorizes as follows:

1. Focus: focusing on theories underlying definition, selection, and utilization of context factors.

2. Goal: aiming for integration of contextualization concepts between pedagogy and technology.
3. Organization: organizing review based on concepts and methods introduced in literature rather than historical order.
4. Perspective: seeking objective literature observation and neutral analysis of domain differences and overlaps.
5. Audience: targeting specialized scholars in pedagogy, TEL, and personalized and adaptive learning environments (PALE) domains.
6. Coverage: following exhaustive and selective approach, through comprehensive literature search and selection criteria application to limit survey scope within objective boundaries.

After retrieving search results from all selected databases, most frequent title and abstract terms are analyzed to control for alignment with research objectives. Figure 8 shows word cloud visualization of most common terminology in retrieved results.



Figure 8. Most common terms appearing in titles and abstracts of retrieved search results.

3.2.2 Selection Criteria

Articles were screened based on research question relevance, inclusion and exclusion criteria, and quality assessment.

Inclusion criteria require selected papers to:

- represent original work or surveys
- be accessible to study authors

- be written in English language

Exclusion criteria eliminate papers appearing in search results but not serving review goals:

- Paper considers non-"learning context" contexts
- Paper represents workshop description, proceedings preface, or similar
- Paper was already found in another database (duplicate)

In addition to database-retrieved papers, eight papers were added to study. These papers did not appear in search results since search terms do not appear in titles or abstracts. However, they were identified as relevant due to content implicitly describing contextual factors. Additional papers were found either in references of included review papers or based on authors' prior knowledge of their content.

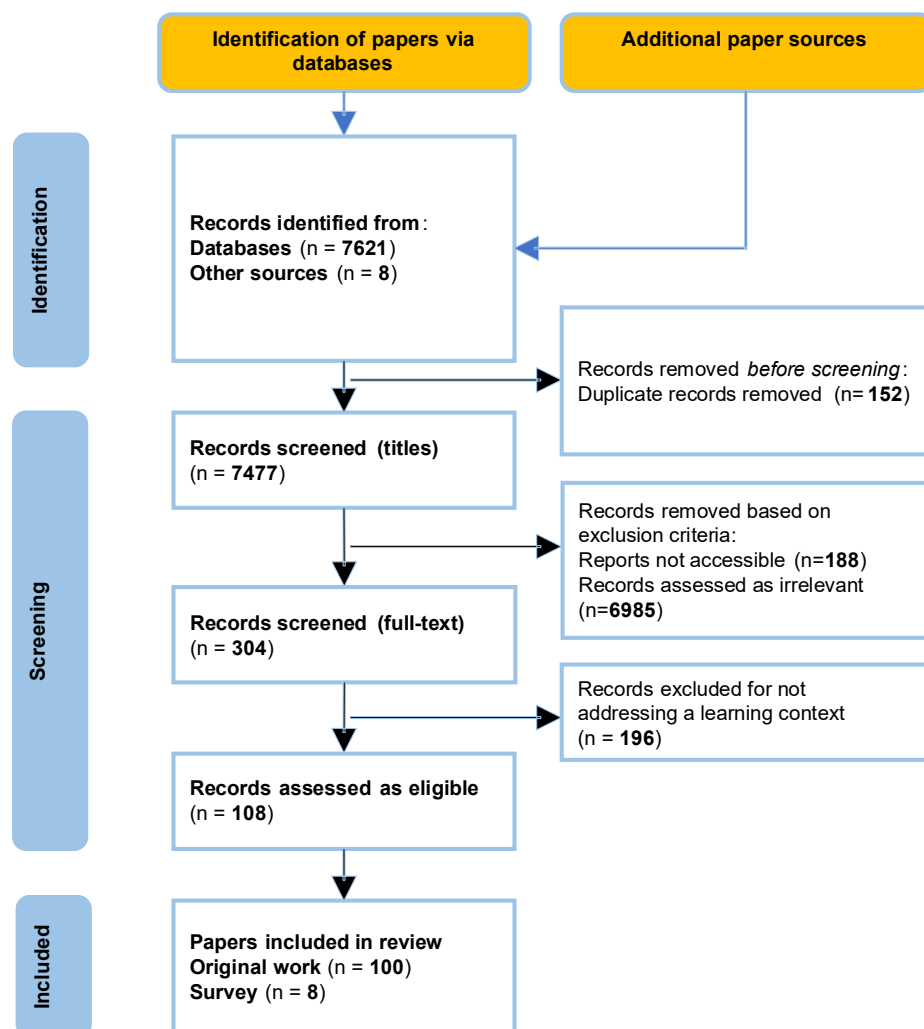


Figure 9. Search, screening, and selection process of the papers included in the review (Abu-Rasheed, Weber, & Fathi, 2023).

3.2.3 Literature Screening Process

Reviewers screened papers based on titles and abstracts. When doubt existed regarding paper inclusion or exclusion, full-text screening identified inclusion and exclusion criteria fulfillment. Final papers considered in review were analyzed to extract findings relevant to research questions. Figure 9 illustrates screening and selection process of papers included in this survey.

3.3 Analysis of the Surveyed Literature

This section discusses the analysis of the surveyed literature and highlight the findings from the concepts identified. To perform this analysis, data collected in the previous phase has been arranged in a database, which represented each paper with the concepts it included. Zotero literature management tool is used to create a report of each paper, reflecting the concept it includes regarding the research questions of this survey. We, therefore, investigate and document what contextual factors are addressed in the paper, the categories and sources of those factors, the application domain(s) the paper covers, the foundation of the factor selection and the learning theories addressed, if any.

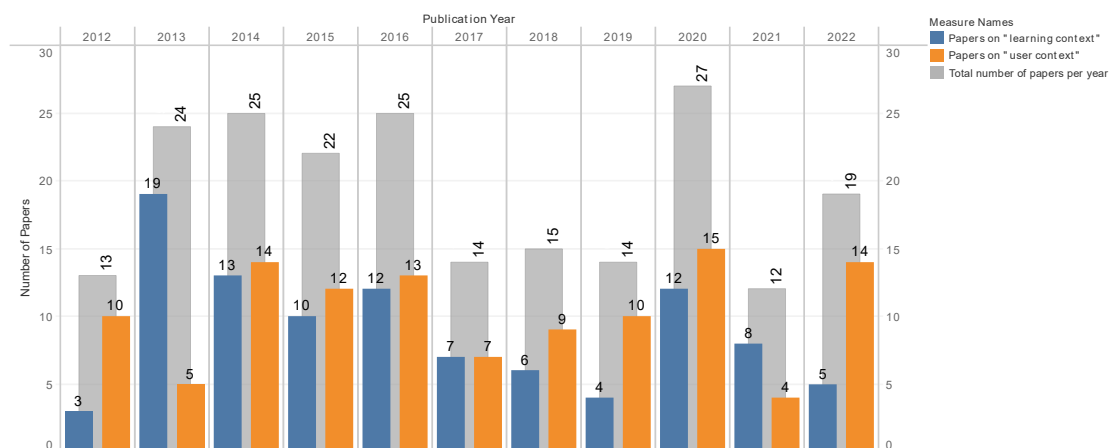


Figure 10. Number of papers covering "user context" (orange) and "learning context" (blue) as parts of the total paper sum per year (grey) (Abu-Rasheed, Weber, & Fathi, 2023).

The overall number of papers reviewed in this research is spread over the period between 2012-2022. In the selection process, papers that considered the “context” were nominated for the final screening phase. From those papers, it was found that not all authors explicitly addressed the “learning context”, which is the focus in this survey. Rather than an explicit learning context, these papers described context-based algorithms or approaches which could be implemented in a learning-related application. This portion of papers still addressed contextual factors and dimensions, but from a general point of view that does not necessarily apply to a learner. Those papers are pointed out to highlight the tendency and patterns in the literature for considering a “learner context” in the learning process” against a “user context” in a general application. Figure 10 shows the number of papers in both categories during the period between 2012 and 2022. While

papers on the learner context and those on user context are relatively comparable in numbers before 2017, the trend for papers on a learner context slightly decreases afterwards, while the number of those discussing a user context relatively increases in comparison. This fact does not necessarily mean that learning contexts became less interesting for researchers.

3.3.1 Contextualization Domains of Application

Beyond general-purpose contextualization, tendency to use technical foundations for selecting context factors remains visible in literature focusing on learner contexts. To analyze this aspect, application fields covered by surveyed papers are investigated. Figure 11 shows different fields authors focused on in proposed context-aware approaches. Application domain distribution shows two dominant fields: recommender systems and mobile learning (M-learning). Both domains form over 31% of surveyed papers. Following these domains are ubiquitous learning (U-learning) and adaptive e-learning. Remaining implementation fields spread across 38% of papers, including authoring tools, game-based learning, blended learning (B-learning), and adaptive e-learning, among others. Papers not addressing specific application areas were included in "General" category, covering 11.7% of papers. The general category isn't visualized in Figure 4 to enhance readability.

In recommender systems domain, Niemann and Wolpers (2013) develop content recommendation system in TEL, utilizing web portal learning objects to find semantic similarity. Shcherbachenko and Nowakowski (2018) utilize user context to enhance learning recommendation system accuracy. Authors propose architecture for context-aware e-learning system based on analysis of means for creating m-learning recommendation systems.

In m-learning domain, increased student mobile device use enabled utilizing integrated sensors to capture additional user contextual data. Mobile devices have been recognized as promising technological means facilitating teaching and learning for individual learners (Gómez et al., 2014; Sotsenko et al., 2016). This indicates increased potential for personalizing content delivered through mobile devices based on learner characteristics and situation (Gómez et al., 2014). This argument receives support from Alnuaim et al. (2016) and Ryu and Parsons (2008), who argue mobile learning can enhance student learning process and fulfill personal needs due to information provided about learners. Moebert et al. (2016) propose generalized approach integrating context in mobile learning applications, providing contextualization framework implementable across different domains and learning scenarios.

Contextualization support extends beyond learners to content creators using educational authoring tools, as in Gallego et al. (2013), who propose context-aware recommender system suggesting educational resources for content creators adding new learning objects on authoring tools. Authors note that captured and utilized context

includes not only learner context but learning material context. This educational context can include material information such as language and target age, enabling linking material to certain user contexts.

Some surveyed literature publications include original work focusing on context solely from data-driven perspective, relying on data from, e.g., mobile devices, to represent learner context. Other publications show pedagogically informed approaches for handling user context, considering pedagogical requirements and learning theory rules to define and use learner context. To differentiate both contextualization approaches, context dependence on learning theories and dimensions is investigated in following section, then introduce categorization approaches to formulate differences and similarities.

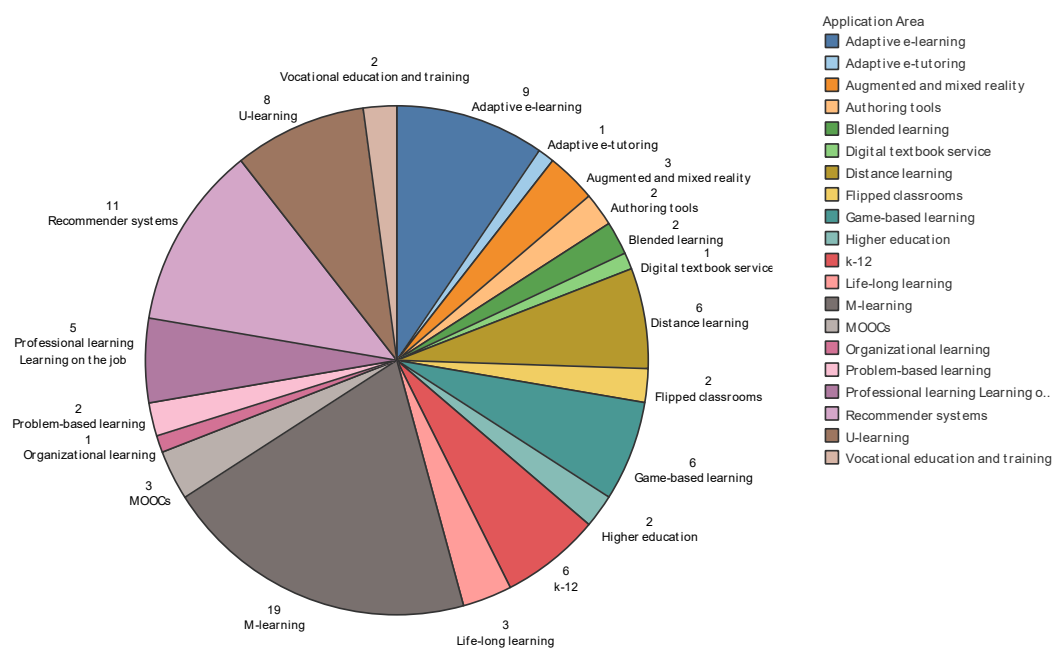


Figure 11. Domains of application for the context-aware methods and systems in the surveyed literature (Abu-Rasheed, Weber, & Fathi, 2023).

3.3.2 Learning Theories in Context-aware Approaches

Considering differences and overlaps between context definitions in surveyed literature (see section 2.4.1), most surveyed literature falls into two categories: (1) explicitly pointing out need for building learning-personalization solutions on pedagogical foundations, as in Peña de Carrillo and Choquet (2013), Soualah-Alila et al. (2013), Gómez et al. (2014), or (2) implicitly addressing pedagogical learning aspects by focusing on contextual factors corresponding to one or more learning theories, as in Musumba and Wario (2019), who propose adaptive e-learning architecture building user and material profiles considering learner decision-making process and social learning aspects, implicitly reflecting problem-based learning, student-centered learning, and social constructivism theory. Figure 12 shows main learning theories addressed in surveyed literature.

Bougsiaa (2016) and Zimmerman et al. (2016) utilize augmented reality (AR) as ubiquitous learning tool providing contextual learning experience to students. Bougsiaa argues, based on Dunleavy et al. (2009), that AR aligns well with both situated learning and constructivism theories because learners are situated in real-world physical environments and potential social settings (depending on AR application), offering learners guidance and facilitating participation in meta-cognitive learning processes (Bougsiaa, 2016). Zimmermann et al. argue AR offers learning setting supporting disciplinary thinking by providing just-in-time information from surrounding environment and incorporating learning process scaffolding (Zimmerman et al., 2016).

Gamified learning has been proposed as means for realizing self-determination theory, as shown in works by Shi et al. (2014) and Hwang and Chang (2020). Realization occurs through supporting learner autonomy in game decision-making and competence toward controlling learning outcomes and learning content mastery. This approach aims to increase learner motivation in social e-learning environments (L. Shi et al., 2014). This approach further reflects goal orientation theory by utilizing ubiquitous computing environment to enhance learning efficiency through increased learning process motivation (Chiou & Tseng, 2012). Chiou et al. built on goal orientation theoretical concepts to develop ubiquitous learning environment support system offering navigation guidance for learners among existing learning content (Chiou & Tseng, 2012).

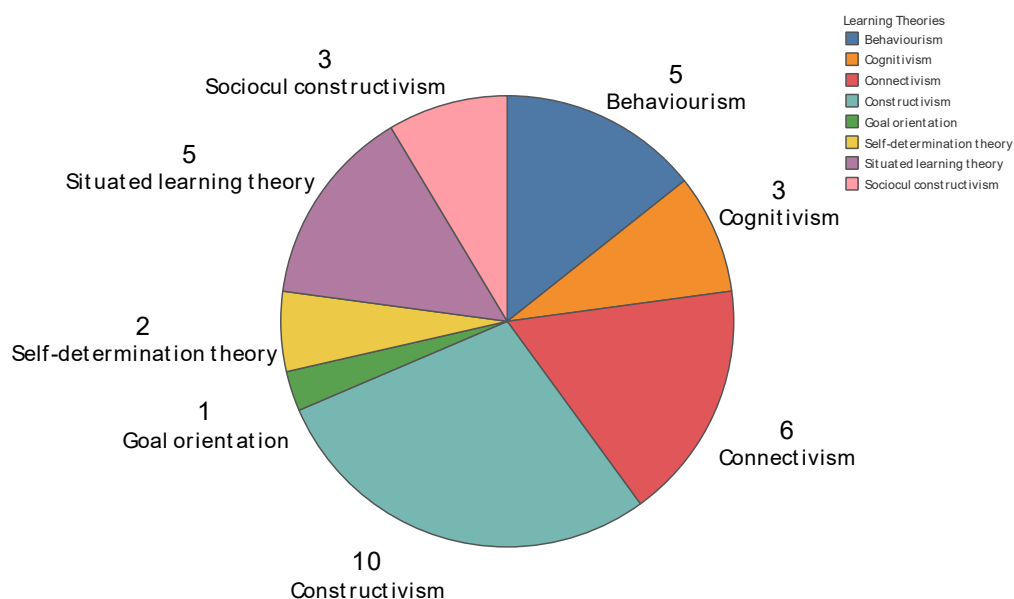


Figure 12. Learning theories addressed by the proposed solution in the surveyed literature (Abu-Rasheed, Weber, & Fathi, 2023).

3.3.3 Context Dimensions

Learning context definition indicates that situations where learning occurs describe not only learner situation but also environment and interaction between them. This

definition establishes first foundation for categorizing learning context into different dimensions: user, environment, and interaction means.

Throughout learning contextualization research, scholars have extended these dimensions to include finely granulated ones representing smaller or more specific context factor groups. Bases for selecting certain dimensions or dimension groups varied between researchers. While some literature specifies contextual factor dimensions corresponding to different sensory information that, e.g., mobile devices can collect, other literature defines dimensions detailing learning pedagogical setting, thus separating, e.g., teacher interaction from school administration staff interaction.

An early thorough investigation of contextual dimensions was introduced in 2012 by Verbert et al. (2012), identifying eight learning context dimensions in TEL:

1. Computing, including software and hardware specification, network characteristics, etc.
2. Location, including quantitative GPS readings or qualitative values such as home or school.
3. Physical conditions, describing user surroundings, such as lighting and sound conditions.
4. Time, referring to learning material access point or learning duration.
5. Activity, referring to actions or tasks users conduct during learning process, usually captured as events on, e.g., learning platforms during learning sessions.
6. Resource, describing learning resources regarding technical aspects (video, audio, text, etc.) or associated metadata, such as resource annotations.
7. User, describing basic user information, plus knowledge, interests, and goals.
8. Social relations, including interactions with peers, colleagues, or educators in learning settings.

Cerinšek et al. (2013) explore contextual enrichment of competence model. They preview TEL from business- and organization-oriented perspective, leading authors to recognize four context dimensions: individual, organizational, knowledge, and environmental. While Verbert et al. address technical and pedagogical perspectives of contextual factors, Cerinšek et al. focus on qualitative factors within each previous dimension. For example, individual dimension includes learner information such as multi-tasking ability or knowledge acquisition ability. Organizational factors include organization ethics, social responsibility, and multidisciplinary aspects. Knowledge dimension describes legislation factors and business models, resource management, and planning. Environment factors include available funds, community stakeholders, market trends, and climate changes, among others.

Sudhana et al. (2013) categorize context factors into two dimensions: static and dynamic. Static dimension includes user personal details and environment factors, while dynamic dimension includes user preferences. This categorization is based on contextual information acquisition. Here, environment dimension includes location and time factors

acquired from user profile and system interaction. Authors also categorize learning context from learner perspective into three dimensions:

1. learner situation, including learning environment such as device, location, and time.
2. learning domain, representing subject or learning area details.
3. learner activity, including learning approach and interaction events.

Yoo et al. (2013) consider contextualization process from recommendation system algorithm perspective. Authors classify context factors under three dimensions:

1. Source, representing recommendation source, i.e., recommender system itself.
2. Message, representing recommendation communicated to learner.
3. Receiver, representing learner receiving learning recommendation.

Mohammad et al. (2015) analyze learning context in massive open online courses (MOOCs), classifying context dimensions as: course context, class context, and student context. Bicans (2016) addresses context in classrooms and intelligent tutoring systems, recognizing five context modeling dimensions related to learning object, pedagogy, student, learning session, and learning environment. Li (2019) addresses contextual factors in MOOCs within dimensions: external context, learner context, social interaction, instructional design, and delivery platform. Author arranges dimensions morphogenetically, i.e., sequentially regarding time, to describe different contextual levels learners experience throughout learning process. Hemmler and Ifenthaler (2022) also classify contextual factors as external and internal, former referring to learning environment and setting, latter to factors related to learners themselves. Authors further extend context dimensions within these categories covering fine-grained aspects of learner, environment, and resource.

Moebert et al. (2016) classify context dimensions as:

1. Physical context describing surrounding environment
2. Mobile context including location and movement factors
3. Situational context representing learner physical situation, e.g., facial expressions or body gestures
4. Scenario, representing learning progress, tasks, and time required for tasks
5. Personal context representing previous knowledge, motivation, preferences, social interactions, and expectations
6. Technical context, representing infrastructure used for learning, such as available device types and specifications.

Huang et al. (2017) classify context into: time, location, task, device, social, and environmental, defining each dimension similarly to Verbert et al. (2012). Authors further propose learner-characteristics framework including learner basic information, cognitive level, learning style, and interest preferences.

Macchia and Brézillon (2021) proposed three overarching dimensions, or frames, to map learning objects and pedagogical methods in pedagogical training. These frames further break down into contextual dimensions similar to previous classifications. Proposed dimensions are:

1. Learner, including contextual factors within sub-classes of physical characteristics and state, socio-cultural characteristics, emotional, intellectual, motivational, learning objectives, and learning profile
2. Training, representing domain and including sub-classes of training objectives, operational objectives, global pedagogical objectives, and session context
3. Learning activities, representing learner interaction and manipulation of learning objects.

Lallemand and Koenig (2020) categorize context factors in user experience (UX) domain within six dimensions:

1. Physical, describing physical conditions where learning application is deployed, describing user interaction with other users
2. Internal, referring to user status such as motivation and expectations
3. Technical, referring to technologies used and user experience with them
4. Task, referring not to task description but user task perception, such as task focus, control, distractions, and potential multitasking
5. Temporal, describing time-related factors such as system interaction time, interaction duration, frequency, or interaction time pressure.

3.4 Results and Discussion

This section discusses results, patterns, and lessons-learned from literature. literature categories were identified, which were not explicitly addressed in previous works by introducing two context factor categorization schemes. First scheme indicates pedagogical and technical origins of context in literature, especially when origins overlap or differ. Second scheme indicates different profiles factors belong to when used in digital learning platforms.

Findings are supported with deeper discussion on observations from 2012-2022 literature landscape. Four observations are formulated, highlighting critical issues identified from literature analysis and provide overview of potential future research directions to bridge identified gaps and promote greater alignment between pedagogy, TEL, and PALE.

3.4.1 Context-factor Origins

Building on the overlap and differences between learning theories and technological solutions in defining learning context and developing contextual learning approaches, learning contextualization categorization is identified, which points out multiple proposed

concepts and solution origins. Table 1 shows the distribution of surveyed papers over the context-origin categories. Three categories were concluded in surveyed literature:

- **Pure pedagogical contextualization:** Concepts from learning theories are considered for defining learning context and indicators. Papers in this category usually present deep philosophical argumentation on context meaning and knowledge construction implications in learning process. These papers discuss context at high level, rarely addressing how context is captured or measured in real-world scenarios. Examples include (Buchmann, 2022; Danes & Brewton, 2014; Fancsali & Ritter, 2014; Negovan & Bogdan, 2013; Opel & Brinda, 2013; Pimmer et al., 2013; L. Shi et al., 2014; Tempelaar et al., 2013).
- **Pure technical contextualization:** This category represents opposite side of contextualization approaches, where context definition is solely based on devices and technological solutions used to capture learner data. Papers in this category focus on measuring contextual data through sensors and technological observations, defining learning context directly from that angle, as seen in (Baccari & Neji, 2016; Chorfi et al., 2012; Clarkes-Nias et al., 2015; Gallego et al., 2013; Hwang, 2013; Ku, 2014; Madhu Sudhana, 2015; Niemann & Wolpers, 2013; Qiuyan et al., 2013; Supic, 2016).
- **Pedagogically informed technical contextualization:** This category includes technical implementations of contextualized learning based on solid, explicit or implicit foundations in learning theories and pedagogical requirements. Papers in this category argue relationships between learning technologies (e.g., AR, recommendation systems, mobile devices, learning games) and specific learning theories realized through these technologies. Examples include (Alnuaim et al., 2016; Borgnakke, 2015; Brézillon, 2014; Ciordas-Hertel, 2020, 2020; Ciordas-Hertel et al., 2022; Colman et al., 2014; di Mascio et al., 2013; Edmonds, 2014; Evans, 2015; Forissier et al., 2014; Hampel & de los Arcos, 2013; Herrero-Martín et al., 2022; Knobelsdorf & Tenenbergh, 2013; Moebert et al., 2016; Mohammad et al., 2015; Mowafi et al., 2014; Parker & Hollister, 2014; Peña de Carrillo & Choquet, 2013; Ramakrishnan et al., 2014; Reynolds & Chiu, 2013). In the past ten years, papers in this category formed 58% of the surveyed literature, in comparison to 23% purely pedagogical approaches, and 19% purely technical ones.

Table 1. Frequent contextual factors, their context-origin, profiling categories, and examples of papers addressing them in the surveyed literature (Abu-Rasheed, Weber, & Fathi, 2023)

Factor	Contextualization foundation			Context profiling category			Addressed in
	Pedagogical	Technical	Pedagogical+ Technical	User	Resource	Environment	
Location			●	●		●	(Alnuaim et al., 2016; Baccari & Neji, 2016; Bougsiaa, 2016; Ciordas-Hertel, 2020; Musumba & Wario, 2019; Siraj et al., 2018)
Time			●	●		●	(Bidarra & Rusman, 2017; Bouihi & Bahaj, 2017; Brik & Touahria, 2020; Sevkli & Abdulkarem, 2015; Starr, 2020; Supic, 2016; Wongchiranuwat et al., 2020)
Previous knowledge	●			●			(Bicans, 2016; Bicans & Grundspenkis, 2017; Hilkenmeier et al., 2021; H. Li et al., 2017; Nijenhuis-Voogt et al., 2018; Tarus et al., 2018)
Peers / colleagues			●			●	(Daoudi et al., 2020; Dennen et al., 2018; Junianto & Wutsqa, 2019; Lallemand & Koenig, 2020; Shcherbachenko & Nowakowski, 2018; Starr, 2020)
Activity			●	●			(Bouihi & Bahaj, 2017; Junianto & Wutsqa, 2019; H. Li et al., 2017; Nijenhuis-Voogt et al., 2018)
Interests	●			●			(Brik & Touahria, 2020; Mayeku & Hogrefe, 2017; Sotsenko et al., 2016; Thüs et al., 2015; H. T. Zimmerman et al., 2016)

Instructional strategies	●				●	●	(Buchmann, 2022; Caponera & Losito, 2016; Kuger et al., 2016; Zhou et al., 2021)
Motivation	●			●			(Lallemant & Koenig, 2020; Z. Li, 2016; Mutambik et al., 2020; Starr, 2020)
Noise level		●				●	(Ciordas-Hertel, 2020; Ciordas-Hertel et al., 2022; Lallemant & Koenig, 2020; Zheng et al., 2019)
Psychological state	●			●			(Buchmann, 2022; Ciordas-Hertel et al., 2022; Daoudi et al., 2020; Ilkou et al., 2021; Shcherbachenko & Nowakowski, 2018)
Culture	●					●	(Aldowah et al., 2019; Bidarra & Rusman, 2017; Kuger et al., 2016; Mutambik et al., 2020; Nijenhuis-Voogt et al., 2018)
Ambient light		●				●	(Ciordas-Hertel, 2020; Ciordas-Hertel et al., 2022; Huang et al., 2017; Lallemant & Koenig, 2020; Mausz & Tavares, 2017)
Learning goal			●	●	●		(Bicans, 2016; Daoudi et al., 2020; Lallemant & Koenig, 2020; Mayeku & Hogrefe, 2017; Tarus et al., 2018)
Network bandwidth		●				●	(Baccari & Neji, 2016; Huang et al., 2017; Sotsenko et al., 2016)
Socio-economic status	●			●		●	(Buchmann, 2022; Caponera & Losito, 2016; Starr, 2020; Zhou et al., 2021)
Language			●	●	●		(Aldowah et al., 2019; Baccari & Neji, 2016; Daoudi et al., 2020; Sotsenko et al., 2016; Starr, 2020)
Device interaction type		●				●	(Alnuaim et al., 2016; Dennen et al., 2018; Lallemant & Koenig, 2020; Shcherbachenko & Nowakowski, 2018)
Organization	●			●		●	(Aldowah et al., 2019; Bidarra & Rusman, 2017; Caponera & Losito, 2016;

							Hilkenmeier et al., 2021; Mutambik et al., 2020)
Learning history			●	●			(Bicans & Grundspenkis, 2017; Chung, 2020; Huang et al., 2017; Ilkou et al., 2021; Musumba & Wario, 2019)
Weather conditions			●			●	(Lallemant & Koenig, 2020; Sotsenko et al., 2016; Zheng et al., 2019)
Physical state			●	●			(Daoudi et al., 2020; Lallemant & Koenig, 2020; Macchia & Brézillon, 2021; Sotsenko et al., 2016)
Task	●			●		●	(Huang et al., 2017; Hwang & Chang, 2020; Siraj et al., 2018)
Legal / Laws	●					●	(Aldowah et al., 2019; Buchmann, 2022; Starr, 2020)
Teaching materials			●		●		(Caponera & Losito, 2016; Dennen et al., 2018; Musumba & Wario, 2019; Nuankaew & Nuankaew, 2019)
Ambient temperature			●			●	(Ciordas-Hertel et al., 2022; Huang et al., 2017, 2017)
Wireless communication		●				●	(Aldowah et al., 2019; Daoudi et al., 2020; Huang et al., 2017; Siraj et al., 2018)
Preferences			●	●			(Amasha et al., 2020; Canbaloğlu & Treur, 2022; Ilkou et al., 2021; Musumba & Wario, 2019)
System interactions			●			●	(Chung, 2020; Shcherbachenko & Nowakowski, 2018; Tarus et al., 2018; Zarrad & Zaguia, 2015)
Device ID		●				●	(Baccari & Neji, 2016; Bouihi & Bahaj, 2017)
Age			●	●			(Bouihi & Bahaj, 2017; Daoudi et al., 2020)

Classroom condition	●					●	(Caponera & Losito, 2016; Zheng et al., 2019)
Battery charge		●				●	(Siraj et al., 2018; Sotsenko et al., 2016)
Screen size		●				●	(Alnuaim et al., 2016; Siraj et al., 2018)
Ethics	●			●		●	(Kuger et al., 2016; Starr, 2020)
Parental involvement	●					●	(Caponera & Losito, 2016; Zhou et al., 2021)
Movement		●				●	(Ciordas-Hertel et al., 2022; Shcherbachenko & Nowakowski, 2018)
Gender	●			●			(Mutambik et al., 2020)
Teacher involvement	●					●	(Aldowah et al., 2019; Herrero-Martín et al., 2022)
Device type		●				●	(Lallemand & Koenig, 2020)

3.4.2 Profile-based Context Dimensions

Surveyed approaches have proposed various dimensions intersecting and differing on multiple levels. Application domain (Zbick et al., 2016), learning tools and infrastructure (Hwang, 2013), and reliance on technical or pedagogical learning foundations (Macchia and Brézillon, 2021) all influence adoption of certain dimensions over others.

On learning platforms, information about learners, learning resources, and learning environment typically appears as profiles. Learners have unique profiles including information about age, interests, learning goals, etc. Learning resources similarly have profiles including resource type information, length, etc. Environment information, such as lighting and noise levels, is typically measured and included in temporary "learning session" profile. While previous classification schemes address context factors within different learning settings, one missing approach is addressing how learning context factors are practically handled on learning platforms. For example, learner profile covers multiple dimensions about learner/user from above schemes, including "Physical," "Task," "Internal," "Situational," and "Personal" classes. To address this issue from practical perspective considering different profiles that context factors belong to on

learning platforms, new scheme is proposed for categorizing contextual factors within three dimensions:

1. Learner,
2. Educational resource, and
3. Environment.

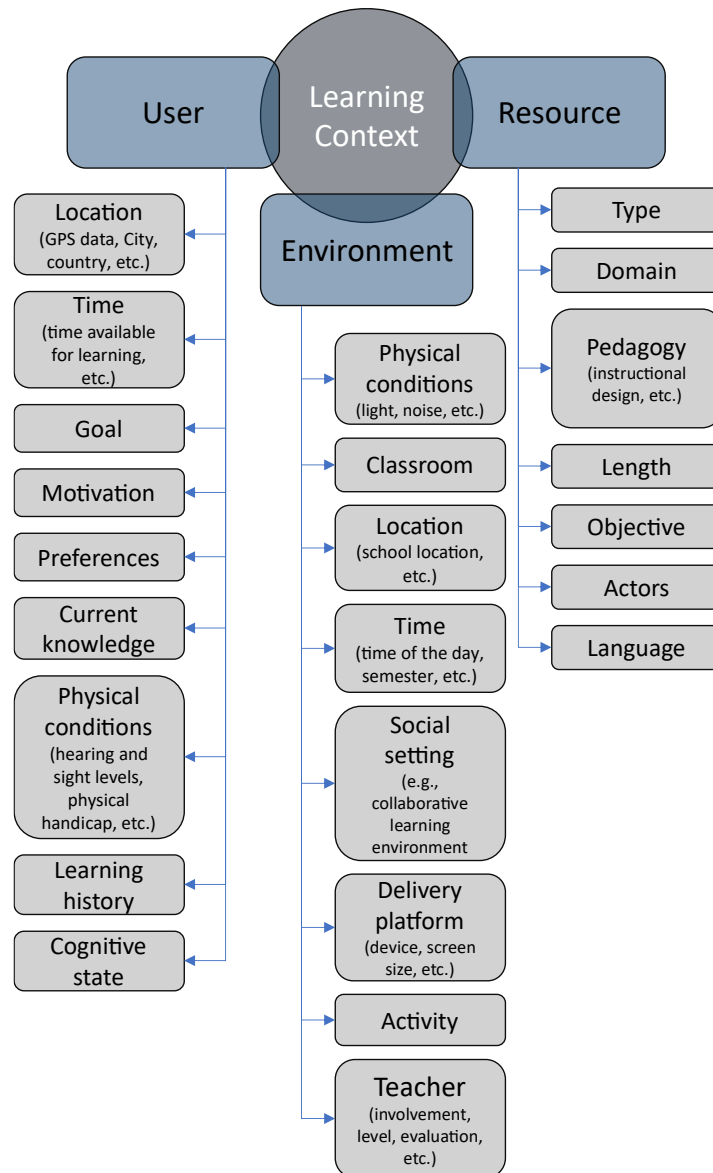


Figure 13. Proposed classification scheme of context factors based on their representation in learning platforms (Abu-Rasheed, Weber, & Fathi, 2023).

The latter dimension breaks down into sub-classes describing technical, dynamically changing factors, such as noise and light levels, and pedagogical factors such as organizational structure, teacher assessment, or parent and teacher involvement. Figure 13 illustrates proposed dimensions and sub-classes. This classification aims to accommodate majority of frequently used contextual factors from pedagogical and technical viewpoints. Dimensional structure is defined on multiple granularity levels, adaptable to domain-specific requirements and technical implementation requirements,

while aligning with common learner and resource profiling practices on learning recommendation platforms.

3.4.3 Context Capturing and Contextual Factors

Within Technology Enhanced Learning (TEL), the contextualization of learning processes demands accurate representation and capture of the learning context within technological implementations. The varying definitions of learning context lead to different factors being utilized for its representation. These factors serve to provide quantitative manifestations of what would otherwise be qualitative contextual aspect definitions. For example, the social context of a learner, which is a contextual aspect originating from constructivist theory, can only be implemented in learning recommender systems if quantifiable values exist that recommendation algorithms can utilize for calculating top-n suggestions. Such quantitative values might include identifiers of user profiles connected to the current learner's profile, or quantities of shared courses attended by both the current learner and another member of their social group. Without these quantitative representations, recommender systems cannot evaluate or leverage social factor indicators when generating learning suggestions.

Reflecting on the learning context definitions across the three categories identified in section 2.4.1, it becomes evident that technical algorithms' inherently quantitative nature sometimes encounters challenges when addressing pedagogically-defined contexts. For instance, legal or psychological contextual aspects (Buchmann, 2022) do not readily translate into simple data structures or parameter values easily measured through sensor technology. These factors also present complexity for human-based evaluation methods such as surveys, as they require specialized knowledge and expertise for proper assessment. By contrast, contextual factors like location, which can be selected from predefined options or automatically determined via mobile GPS, illustrate the substantial difference in difficulty when capturing certain contextual factor values compared to elements like the legal learning context or a learner's socioeconomic status.

Such challenges and similar limitations have prompted scholars like Taylor et al. (2021) to argue that scaling personalized pedagogical approaches to individual levels necessitates technologies not yet fully available. This assertion, however, does not suggest that personalization based on individual learner contexts is impossible. Rather, as this research interprets it, it indicates two imperatives: (1) contextualization technologies and algorithms should address pedagogical aspects such as legal factors to the best of their capabilities, acknowledging limitations in capturing or representing these elements rather than avoiding them entirely as observed in most technical solutions; and (2) if contextualization approaches seek pedagogical compatibility to scale effectively, qualitative contextual factors require further research to develop suitable quantitative representations for algorithmic implementation. This endeavor necessitates equal effort from both pedagogical and technical experts, as effective solutions remain unattainable without close collaboration between these domains. Within the surveyed literature, this

concept appears only partially in Moebert et al. (2016), who classified different contextual factors' relevance to predefined educational settings (formalized, physical, collaborative, immersive, teaching, and learning). The authors also categorized measurement accuracy for these factors within each setting and proposed a framework for detecting, collecting, and utilizing contextual factors in adaptive mobile learning applications. Their framework emphasizes adaptability to diverse learning settings, recognizing that factors vary in both relevance and measurement accuracy across different contexts.

The capture of learning context factors depends on their definition, the underlying algorithms utilizing them, and the tools employed in learning activities. To illustrate this point, consider "location" as a contextual factor. Location may refer to a learner's physical position expressed through geographical coordinates, or alternatively, to their virtual presence represented as a node within a connected device network. Each definition necessitates a different capture approach—using GPS readings for the former definition and device IP addresses for the latter. Algorithms and tools similarly influence context capture methodologies. Mobile learning personalization utilizes device sensors to acquire contextual learner information, while recommendation algorithms implemented on web-based learning platforms typically access learner profiles for contextual information, often manually provided during platform registration (e.g., location city).

Another approach for capturing more qualitative, pedagogically-oriented contextual factors involves surveys analyzed and interpreted by domain experts. This methodology appears in Caponera and Losito (2016) for assessing learners' socioeconomic status. Similarly, the Programme for International Student Assessment (PISA) (Kuger et al., 2016) captures diverse contextual factors influencing student learning through tests and surveys that collect data on performance and other potentially influential aspects, including learner context.

In this literature survey, various contextual factors and their implementation frequency in proposed solutions and approaches were examined. Factors identified in this research and other valuable studies, including Caponera and Losito (2016), Kuger et al. (2016), Lallemand and Koenig (2020), Hilkenmeier et al. (2021), and Hemmler and Ifenthaler (2022), receive varying attention across the literature. Several considerations influence a factor's prevalence: (1) its origin—pedagogical, technical, or hybrid; (2) ease of capture; and (3) impact on learning personalization. Figure 14 illustrates frequency distribution of factors between pedagogical and technical domains. Contextual factors are positioned between these domains based on their treatment in literature from both areas. Consequently, factors appearing near the diagonal line represent the third category introduced (Pedagogically informed technical contextualization). Factor size indicates frequency of literature discussion. A contextual factor is considered to be addressed when authors explicitly identify it as representing learning contexts, even if not implemented in their proposed solution, since factor selection in each solution reflects additional influences including application domain, learning tools, and algorithm characteristics.

Figure 14 includes only factors with frequency ≥ 3 due to the extensive number of less common contextual factors. Table 1 provides a comprehensive listing of these factors with their classifications and relevant literature references.

Figure 14 demonstrates that location and time rank among the most frequently addressed factors in the literature. This prominence can be attributed to both their robust pedagogical foundation regarding learning impact and their feasibility of capture in most technical learning personalization algorithms. Other factors, such as socioeconomic indicators, while common in pedagogical approaches due to their demonstrated influence on learner performance (Caponera & Losito, 2016), appear considerably less frequently than other factors, potentially reflecting measurement challenges in technical implementations.

Similar to location's varying interpretations and definitions, other contextual factors identified in the literature exhibit definitional diversity. Time as a contextual factor may indicate the specific moment when learning occurs, the season, or extend to learning duration or learning tool interaction periods. This conceptual breadth explains why some literature elevates location and time to dimensional status rather than treating them as individual factors. In other words, they may represent entire groups of spatio-temporal factors instead of single elements, reflecting their definitional variety. Contextual factor categorization within different dimensions appears throughout the literature using numerous dimensional frameworks. The following section explores these diverse dimensional approaches to contextual factor classification.

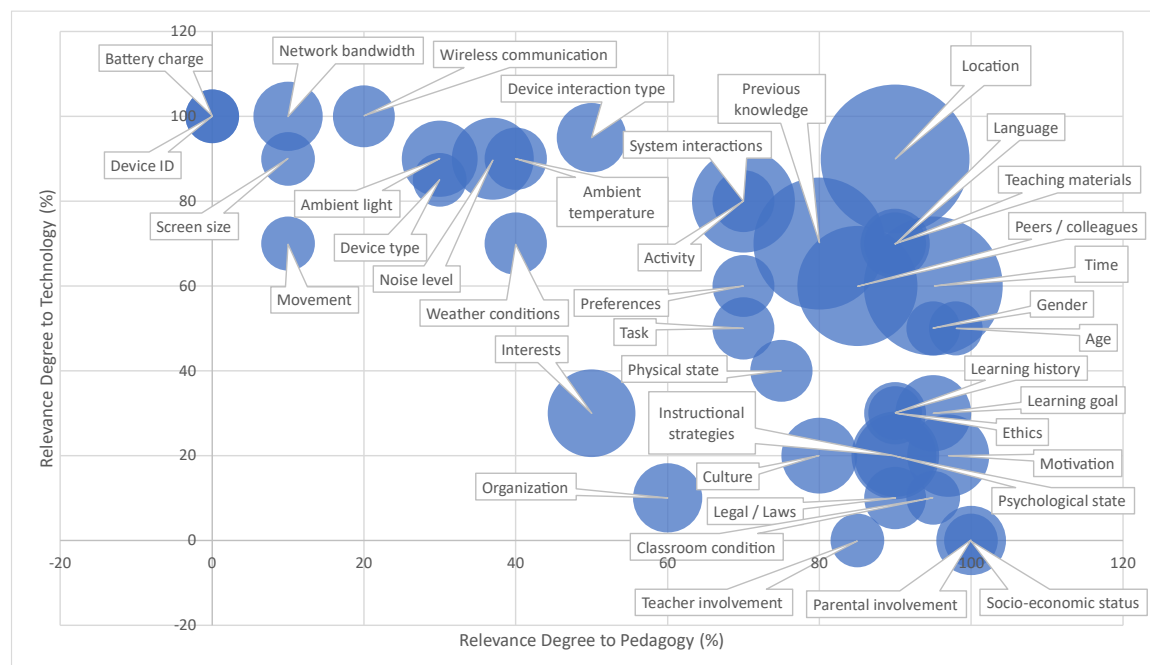


Figure 14. Mapping frequent contextual factors based on their relevance to pedagogy and technology. The size of the factor reflects its frequency in the surveyed literature (Abu-Rasheed, Weber, & Fathi, 2023).

The analysis of surveyed literature has generated several observations and conclusions regarding state-of-the-art research on learning contexts. The subsequent sections provide an objective summary of these observations to highlight key findings, patterns, and limitations identified in the literature. Additionally, challenges and open research opportunities are identified in relation to these findings.

Observation 1: Learning context between pedagogy and technology

Our review results reveal conceptual and implementation gaps between pedagogical and technical domains regarding learning context definition and factor capture. While linguistic context definitions from pedagogical and technical perspectives appear superficially similar, their practical interpretation diverges significantly during implementation. The literature on contextual factors and learning contextualization methods demonstrates that pedagogical research emphasizes high-level, conceptual, and occasionally abstract aspects of learning context. For instance, scaffolding learning highlights the teacher-intervention's role in supporting learning processes without defining intervention mechanisms, boundaries, or descriptive parameters. Similarly, defining learner context in vocational education settings through factors like "knowledge acquisition ability" presents abstract considerations requiring definitional clarity about "ability" itself before addressing what the contextual factor represents. This research argues that this philosophical abstraction tendency in learning theories and pedagogical science occurs naturally for two reasons: (1) the cognitive domain's nature necessitates abstraction and conceptualization levels when defining context as a construct; and (2) learning theories and pedagogical science implementation models inherently delegate detailed interpretation responsibility to educators who realize pedagogical approaches in practice, such as classroom settings.

Technical definitions of learning context similarly exhibit a distinct inclination toward purely practical context realizations, focusing on quantifiable, measurable factors describable through technologies that capture them. For example, location might be virtually represented as a network device node through numerical IP address values. Although this approach potentially offers an intelligent realization of pedagogical learning context concepts addressing connectivity factors from connectivism and social constructivism theories, most technical solutions fail to acknowledge these pedagogical foundations. Nevertheless, they continue utilizing IP addresses as virtual location indicators, motivated by improved accuracy in applications like social recommendation systems. In other words, the underlying rationale for adopting particular technical factors to represent learning context in technological solutions typically stems not from supporting foundational pedagogical goals, but because factors are easily measurable or have demonstrated utility in non-educational algorithmic implementations. Technical solutions adopt simplified measures not because they optimally represent pedagogical constructs, but because they offer implementation ease or proven effectiveness in other contexts. The implicit assumption that simplified contextual measures (e.g., IP addresses)

adequately correlate with targeted contexts (e.g., social settings) rarely receives pedagogical justification.

Several simplified factors, such as "reviews" or "like counts", implemented in context-aware educational recommendation systems originate from product recommendation algorithms for books or movies. No literature providing a clear pedagogical justification for similar factor applications in learning recommendations was observed. This approach becomes particularly evident when examining evaluation strategies for context-aware recommendation systems, which predominantly rely on large datasets rarely containing pedagogically-defined contextual factors.

It is worth noting that many technical learning personalization solutions build explicitly or implicitly on pedagogical foundations, as observed in the surveyed literature. However, standardized context definitions and consistent factor implementations remain notably absent across implementations.

Observation 2: Context dimensions as domain-specific or standardized schemas

Surveyed literature reveals considerable variation in contextual factor classification approaches. Classification schemes reflect diverse considerations, including pedagogical settings, implementation infrastructures, and domain-specific requirements. Classification schemes sometimes incorporate dimensions from other schemes, but with different definitions or granularity levels. For instance, location may constitute a specific factor within learner or environment dimensions in certain schemes, while representing an entire dimension encompassing multiple spatial factors in others.

Standardized context factor classification is essential for aligning context-aware solutions and enabling cross-system factor reusability. Unique definition, reusability, and interpretability represent key elements for implementing FAIR principles (Findability, Accessibility, Interoperability, Reusability) in scientific data management (Wilkinson et al., 2016), offering important guidelines for context-aware learning personalization implementations.

The standardization challenge for describing contextual factors and dimensions isn't new, Verbert et al. identified this issue in 2012, and apparently remains unresolved in the current state-of-the-art practice despite several attempts to propose generalizable context dimension schemes, including work by Zbick et al. (2016), Moebert et al. (2016), and Li (2019). One potential solution involves semantic web-based approaches for defining standardized vocabularies for context factors and dimensions. Utilizing standard definitions for elements like time or location, comprehensively described in semantic web schemas such as Schema.org from the World Wide Web Consortium (W3C), could enable global factor and dimension reusability across context-aware solutions.

Observation 3: Context capturing and quantification

Beyond definitional differences between technical and pedagogical domains, capturing and quantifying contextual factors originating from pedagogical approaches and learning theories presents further challenges. The ability to represent or model abstract contextual factors like "Ethics" is hindered by their inherently abstract nature. A widespread difficulty is observed in capturing and quantifying numerous contextual factors derived from pedagogical approaches and learning theories. This difficulty stems from: (1) inadequate methodologies for quantifying qualitative factors such as "parental involvement" in learning; and (2) uncertainty that quantification approaches preserve the pedagogical meaningfulness of contextual factors.

For example, if a qualitative factor like "teacher interaction" is quantified by counting student-teacher conversations, does this quantification accurately reflect the concept as intended in scaffolding learning theory? The answer requires pedagogical expertise. Consequently, when context-aware algorithms are designed and implemented by technical specialists without pedagogical guidance, quantification methods risk misinterpreting educational requirements. This observation emphasizes the necessity of collaboration between pedagogical and technical experts when designing and implementing personalization algorithms generally, and context-aware algorithms specifically.

Observation 4: Explicit and implicit learning context consideration

During literature search on learning context definitions and factors, numerous publications implicitly addressed learning context factors (e.g., in learner profile design) without explicitly labeling them as "learning context factors." Similarly, some publications describe learning context indirectly by characterizing learner situations or learning settings without explicit "learning context" terminology. This creates challenges for comprehensive literature identification using explicit search terms—a limitation of survey methodologies like ours.

It has been observed that publications implicitly addressing learning context frequently relate to user or learning resource profiling. Despite diverse applications, these publications typically include learner representations with metadata describing contextual factors. Examples include works utilizing ontological classes to characterize learner or learning material situations without explicitly framing these as contextual classes (Barria-Pineda et al., 2019; Chimalakonda & Nori, 2020; Ciloglulugil & Inceoglu, 2018; Ilkou et al., 2021; Pal et al., 2021).

One potential explanation for this observation relates to the absence of standardized context definitions suitable for universal adoption across user and resource modeling research. This suggests a research need for standardized learning context descriptions and

frameworks providing implementation roadmaps for considering contextual factors in pedagogical and technical methods.

A fundamental modeling challenge persists, which is that models inherently represent imperfect reality approximations. Consequently, differing implementation approaches may optimize for specific limited scenarios while performing sub-optimally elsewhere. A semantic web approach cataloging factor implementations could specify assumptions, limitations, and objectives for each approach, generating implementation ranges sufficiently covering factor meanings across relevant scenarios.

Observation 4 objectively addresses a limitation of the survey, which is related to the literature coverage and finding all potential papers that address the contextual factors. The search terms were selected to the best of our knowledge for collecting papers in multiple domains that are relevant to the learning context and contextual factors. However, other publications that implicitly address context factors may still be found. Therefore, a further extension of the literature coverage and knowledge base of identified papers holds the potential to include additional resources within the analysis presented in this review.

3.5 Related and similar reviews

During the reviewed period, several surveys on context-aware recommendation systems emerged, understandably due to growing attention to this recommendation system type, both commercially and pedagogically. While surveys focused on recommendation algorithms and implementations, literature surveys addressing contextual factors considered by such recommender systems, especially in TEL and PALE domains, were limited in state-of-the-art research, as shown in Table 1. This particularly applies when considering pedagogical foundations of context. After 2012, a gap exists in literature reviews focusing on context definitions and factors informed by pedagogy and learning theories.

In 2012, Verbert et al. (2012) thoroughly investigated recommender system context in TEL domain. Authors introduced a context framework for such recommenders, identifying context dimensions. Their survey analyzes learning context definitions in TEL from pedagogical and technical viewpoints, alongside factor variety used to capture context. Thus, authors offer a comprehensive understanding of contextual factors in TEL. In 2013, Hwang (2013) investigated context-aware ubiquitous learning, highlighting pedagogical and theoretical e-learning concept roles in shaping ubiquitous learning context definition and factors.

Following these surveys, the next survey focusing on pedagogical implications of defining contextual factors appears in Hemmler and Ifenthaler (2022). Authors revisit empirical-pedagogical research role in shaping PALE design and generating meaningful, context-aware recommendations. Their survey identifies 208 internal and external contextual factors across 27 dimensions. The survey also notes literature review scarcity

focusing on pedagogical aspects of learning context and factors. While authors focus on context indicators in higher education, workplace learning, learning analytics, competence acquisition, and competence assessment, as indicated by search terms adopted in their survey, researchers acknowledge survey time span limitation, only extending back to 2019. Authors also don't thoroughly discuss context definitions and identified indicators. Another recent work by Ahmad et al. (2020) introduced an indicator repository and a dashboard tool supporting educators in learning activity selection. Although authors primarily propose and evaluate their tool, they build an indicator selection based on a literature review conducted between 2011 and 2021. Authors don't discuss indicators themselves, their definitions, or pedagogical and technical foundations.

This research therefore builds upon important studies by Verbert et al. and Hemmler and Ifenthaler, extending the survey of contextual factors in TEL considering pedagogical and technical learning process aspects. Learning context definitions and dimensions were highlighted between 2012 and 2022, extending the work of Verbert et al. The review also addresses learning theories forming pedagogical foundations for defining context in TEL and selecting indicators and factors capturing that context. A focus is given to both technical factors and those derived from pedagogy, as both are essential for supporting pedagogically informed digital learning recommendations. Therefore, search terms utilized for literature search and data collection in this survey directly target contextual factors and dimensions in TEL and learning theories.

Table 2. Related reviews in the literature within the review period (Abu-Rasheed et al., 2023)

Review	Year	Reviewed period	Focus
(Verbert et al., 2012)	2012	-	TEL, recommender systems
(Hwang, 2013)	2013	-	Ubiquitous learning
(Hemmler & Ifenthaler, 2022)	2022	2007-2021	PALE, Pedagogy
This review	2023	2012-2022	Pedagogy, TEL, and recommender systems

3.6 Context representation for the knowledge graph construction

This chapter surveyed state-of-the-art learning context definitions and contextual indicators within TEL and context-aware recommendation systems. The goal of the review is to establish a rigorous foundation for reflecting the learning context structurally in the knowledge graphs, and algorithmically in the graph-based recommendation algorithm. The analysis of the literature from the pedagogical and technical perspectives allowed investigating:

1. learning context definitions in pedagogical and technological domains
2. schemes for contextual factors and indicators
3. pedagogical foundations underlying technological development of personalization algorithms; and
4. alignment and divergence between pedagogical and technological approaches to contextual factor utilization.

The two classification schemes introduced for contextual factors in this study, based on pedagogical vs. technological foundations, and based on learning platform representation, present a ground for specifying node and relation types in the knowledge graph, once the base-ontology is defined in the light of those classes. In other words, context factors that stem from the user, the learning resources, and the learning environment, are meant to be embedded in the ontology classes, properties, and class relations. The populated knowledge graph, in this sense, will serve as a structural source of information about the learner and their learning, enabling a CARS to generate contextualized learning path recommendations for the learner.

In the next chapter, the creation of the knowledge graph is discussed and presented through three use cases that were handled in the scope of this thesis. The integration of the learning context is elaborated in each use case, building on the results of the review and the nature of the application domain.

Part II

Knowledge Graph Based Learning-Path Recommendations

4 Knowledge Graph Construction for Context-Aware Personalized Learning

The design process of a knowledge graph is usually domain-specific and, depending on the involved domains, an interdisciplinary one. Its goal is not only to define the main components of the KG, i.e., its nodes, relations, and properties, but also to provide the basis for algorithms operating over the graph structure, for example, generating a learning recommendation or retrieving relevant knowledge. In this sense, the definition of KG components is a purposeful, long-term process that emphasizes that the KG's role in an intelligent system is not solely serving as a graphical database, but also an efficient source of information, and potentially knowledge, once queried properly by the graph-based algorithms.

The embedding of learning context factors reviewed in chapter three in the KG structure is then the step that enables the retrieval of contextual information and reflecting contextual knowledge on the graph-based algorithms. The methodology for context embedding in KGs begins at the design phase itself and materializes through the decisions made to define KG components. Building on chapter three, this chapter continues to address the first research question in the thesis:

“How can the integration of a learning context into knowledge graphs be achieved, structurally and algorithmically, to improve the personalization and adaptability of learning paths?”

By constructing contextualized knowledge graphs in three different use cases. The goal of investigating the three use cases is to answer the research question in a domain-based manner, connecting to the domain requirements and the learning settings within it, rather than falling into a mistaken generalization of the answer.

In this chapter, the construction of the knowledge graphs that were developed and used in the scope of this thesis is discussed in detail. The design considerations of the KGs are linked to the domain-specific nature of the use case. To that end, three use cases in three different, but overlapping, domains are presented. These are in the fields of:

1. **Lessons-learned discovery and retrieval for chip-design in semi-conductor industry:** in which the target group is design engineers, who require direct support for finding failure reports that are semantically relevant to a specific problem or failure they are facing, to learn more about the proper procedure of approaching that failure.
2. **Learning path recommendation for complex staff training in the elderly-care domain:** where the beneficiary group is composed of the nursing staff working in an elderly home, who face daily challenges that need to be solved based on thorough awareness of the multidimensionality and interdisciplinarity of the problem at hand. Therefore, the recommender system here generates a personalized, interdisciplinary learning path for the challenge a nurse is facing.
3. **Module path scheduling in higher education:** which is the use case that takes place in a formal education, unlike the previous two. This use case demonstrates that the techniques developed in this thesis are extendable to a broader range of learning settings, although their development was focused on VET. The adjustments to the graph structure and graph-based recommendation algorithms are reflected in the use case through the consideration of the formal curriculum structures in the KG, as well as the consideration of the time restrictions in the recommendation algorithm.

These three use cases were concluded and implemented in three different projects with a duration ranging between two and three years each. Each of the use cases represents a stand-alone, real-world challenge that requires a strong knowledge foundation for the intelligent and potentially explainable algorithms to operate on. This, in turn, is the reason for selecting the KG in this thesis as the semantically laid foundation for the information retrieval and learning path recommendation algorithms in the three use cases.

4.1 Knowledge graph design, construction, and requirements

The process of designing and implementing a knowledge graph leads ultimately to defining the types and properties of the graph's nodes and relations. However, the decisions made to select each node and relation type, as well as what should be a property for either of them, are highly dependent on several factors that affect how the graph is modeled, e.g., based on its base ontology, and how it is queried effectively by an algorithm. An example of those design considerations is the decision of adding a new property to a node type or defining a stand-alone node type that replaces that property. Take, for instance, a recommendation system in a VET setting, where learners come from different countries and speak, therefore, different native languages. In a simple graph model, the language of learning materials that are meant to be recommended for the learners can be modeled in the graph as a property "*language*" of the node type "*learning_material*". However, considering the algorithm operating over the KG, one can detect that there is a high likelihood that learners may prefer material in their own mother tongues. The recommendation algorithm in that case will have to perform the

following steps: 1) query all nodes from the type “*learning_material*”, and 2) query the property “*language*” from each node, and 3) compare the value of the property to the user-preferred language, to select only materials from that language in the learning recommendation.

A thorough design of the KG, on the other hand, may consider evaluating the likelihood of requesting learning materials in a specific language, and based on the importance and frequency of that use case, define a stand-alone node type “*Language*” in the schema of the graph. In this approach, each “*learning_material*” node in the graph will have a relation to its corresponding “*Language*” node. The process of querying all learning materials in the graph, which are in the user-preferred language, boils down in that case to: 1) locating the preferred language node, and 2) retrieving all learning materials connected to it. This process can be faster and more computationally effective in this simple scenario.

A similar argument is also related to the definition of graph relations, whose role of connecting graph nodes to each other plays an important role in design decision, such as connecting two nodes explicitly instead of inserting a property of the same value in both nodes, or using different relation types to connect two nodes based on different reasons. Generally, certain considerations can be made when designing a KG, which include, but are not limited to:

For graph nodes

- If there is a need to traverse through nodes and relations, defining separate node types is preferred, as properties cannot be traversed.
- Creating separate nodes when the information needs its own properties or relationships.
- Using nodes for information that will be queried or grouped independently.
- Considering implementing separate attribute nodes when multiple entities share the same attributes.

For graph relations

- Adding properties to relationships when they describe the connection itself.
- Converting join tables from relational models into relationships with properties.
- Considering relationship direction or using non-directional relations if possible and needed.
- Use descriptive relationship types that clearly indicate their purpose, e.g., “sells” instead of “has_a”, whenever possible.
- Properties on relationships enable more nuanced graph queries.

For graph properties

- Using properties for descriptive information that doesn't need its own relationships.
- Properties work well for attributes that are unique to a specific entity.
- Considering that properties are not recommended to store structural types (node, relationship, path).

An effective knowledge graph model is, therefore, designed with query patterns in mind.

4.2 Expert integration in KG construction

In addition to the structural considerations of the KG construction, domain-related considerations play an equally essential role in building a KG that can be well utilized in a real-world scenario. Reflecting on the KG design with query patterns in mind, it is clear that graph querying is the technical realization of a domain requirement or a use case. In the example above, the fact that VET workers speak different native languages is a domain-related aspect that affected the decision to design the graph model. Understanding those aspects in each domain requires the knowledge and experience of a domain expert, who is able to define the goals of the KG and the algorithms operating based on the KG, as well as the circumstances in which the overall system will be used.

To that end, a KG that is solely designed by technical developers is more subject to failing to represent domain-related requirements and considerations. Therefore, an approach for designing the KG with domain experts, from the very first phases of modelling, becomes a necessity for ensuring a solid graphical data structure that corresponds to the needs of the users in that domain.

In the three use cases handled in this thesis, several KG structures and taxonomies were developed to correspond to the particular nature and circumstances of the three domains. For example, in the chip-design use case, design engineers in the semiconductor industry had unique patterns of documenting chip failures, in which abbreviations of electronic-component names, chip modules, and electrical measurements were extensively used in the documentation. A learning material in nursing VET, on the other hand, was created by trainers in the field with comprehensive and clear textual content. Therefore, semantic-similarity pipelines that were developed for both use cases required different levels of support from the KG structural design, leading to the introduction of a special type of connecting nodes in the KG of the chop-design use case, which were not needed in the KG design of nursing VET.

The integration of domain experts in the design process of a KG is a complex process on its own. This is because it needs to establish an interdisciplinary communication between the domain experts and the technical experts. In educational

settings, regardless of being formal, non-formal, or informal, domain experts mostly relate to the pedagogy domain, which translates to less understanding of the technical details, e.g., of calculating a semantic similarity score between two sentences. Similarly, the technical experts mostly lack the knowledge about the pedagogical approaches to realize a learning theory in practice, leaving them unable to effectively incorporate the requirements from pedagogy in some cases. To ensure a sound and productive communication among the experts from these different disciplines, this thesis has identified three main components of the integration process of domain experts in the design of the KG:

1. Enhancing technical and pedagogical literacy for domain- and technical-experts, respectively. This step is meant to bridge a part of the knowledge gap among the disciplines, to open a path for productive exchange of knowledge, and effective definition of goals, requirements, and thus alignment steps, for the design of the knowledge structures and the overall system.
2. Searching and defining a common linguistic ground for communication between the experts from the different disciplines. This becomes critically important if two or more disciplines use the same terms with different meanings or interpretations of the meaning.
3. Designing the KG and the algorithms operating over it iteratively, in a way that allows validating and adjusting the intermediate design steps and decisions, with a minimal accumulation of design errors and system rigidity.

This thesis implemented those steps thoroughly and extensively during the design of the KG for each use case. While some of the steps gain more importance depending on the domain, they were essentially required for all involved domains and experts in this research. Moreover, this research identified that implementing those steps not only enhances the effectiveness of creating the KG but also enhances the level of satisfaction of the experts towards the final system. This is traced back to two main elements: 1) the enhanced understanding of the experts from all disciplines of the inner mechanisms and operations of the final system, and 2) their ability to contribute strongly to the different types of decisions made throughout the iterative design and implementation process. This effect is reported here as a side product of the process followed in the thesis, without being argued with as a specific contribution of this research, since it was not systematically evaluated in this thesis.

Three use cases for KG construction, which this thesis handled, are described in detail in the following sections.

4.3 Use case 1 - Lessons-learned recommendations for chip-design process engineers

4.3.1 Use case description

In a semiconductor chip-design scenario, increasing the visibility and findability of previous design failures has the potential to save time and material during a current design process. Previous design-related documents may implicitly or explicitly represent lessons learned from past design cases. Therefore, their visibility to the design engineer is essential for gaining the required knowledge needed to identify, test, and solve a design failure. This extends from solving one design failure at hand to building training programs for the engineers in chip design, where the training program is composed as a learning recommendation, based on the lessons learned from previously documented design failures.

Within the scope of the use case, an approach to improve the contextual connectedness and retrieval of design failure-case documents from multiple sources is developed. The solution utilizes text mining to extract information from the textual data in semi-structured documents. Multiple data sources, providing the textual data and other useful process-related information, are fused together in a knowledge graph in order to model their interlinking and increase their searchability. Relations of the knowledge graph are extracted with the introduction of an additional connector type of nodes to the graph, namely “Linking Nodes”, which acts as an intermediate node type intended to increase the likelihood of retrieving relevant search results by adding useful relations between data points. Linking nodes are based on a textual analysis of the domain-specific, semantic relevance of connected nodes. Intelligent classification methods are also utilized to extract relations between the failure cases, based on the domain-specific vocabulary used in the documentation process.

4.3.2 KG construction methodology

The construction of a knowledge graph consists mainly of two steps: defining the graph’s nodes and extracting the relations between them. In the environment of chip design documentations, a graph node is defined as a representation of an individual document. This representation is primarily achieved through the document’s vector embedding. However, since available datasets contain multiple document types that require different processing, graph nodes do yield a different meaning, depending on the specific document type they represent. An example of this concept is the structural information that accompanies textual data extracted from project descriptions. In this example, project documents provide not only the elements used in a certain microchip design, but also the hierarchical structure of the final chip. This information is added to those nodes in the graph that represent project description nodes.

In the constructed knowledge graph, which is conceptually illustrated in Figure 15, three types of document-representing nodes are identified:

- **Failure Case Nodes (FC):** Which represent the maintenance documentation in the form of reports of previous design failure cases. This type of node is the center of the search functionality in the use case, since enhancing the visibility of previous failure cases, and thus lessons learned, is the main objective of the developed algorithm.
- **Project Element Nodes (PE):** This type of node represents the documentation of semiconductor elements that are utilized for microchip design projects. Those elements are inserted into the design composition. The elements have an implicit hierarchical structure, expressing “is part of” relationships. They reveal a hierarchical structure of one project, as well as element patterns that appear in several projects. Therefore, structural information accompanies the textual description in this node type, enhancing the relation extraction process in the graph.
- **Product Specification Nodes (PS):** Which are the final and most detailed documentation of the designed product. This type of document includes the technical details of the designed chip and is organized in a human-readable form, to support design engineers as well as to inform product customers.

After defining node types, relation extraction is the next step in the construction of the proposed graph. In the knowledge graph, inter-node-type and intra-node-type relations can be introduced:

- **Intra-node-type relations:** This type of relation links document nodes of the same type together. It is extracted from the specific information available for each node type. For example, intra-failure-case relations, or FC-FC relations, are extracted from the textual patterns that appear in maintenance documents, whereas intra-project-element, or PE-PE, relations reflect the hierarchical structure of the elements in the microchip design, as a parent-child relationship.
- **Inter-node-type relations:** Which represent the relations between different types of nodes in the graph, such as FC-PS relations between failure cases and product specification documents. This type of relation is based on the shared structure of the nodes, i.e. the document vector format.

Although including those N-grams in the document vectors seems the suitable approach to utilize their added value, it introduced an amount of noise to the extracted vectors from TF-IDF within the context of the use case. To handle this issue, the structure of the constructed knowledge graph itself is used, and a new type of node is introduced to the graph that is mainly based on the N-gram approach. Therefore, the fourth type of nodes (linking nodes) serves the goal of creating new connections between graph nodes of other types, based on the discovered possible patterns in the textual data and the POS analysis.

Linking Nodes: This type of node is the part of the developed knowledge graph in the use case, which tailors the graph structure itself to reflect the data mining results and domain-specific features. Linking nodes play the roles of representing the domain-specific text mining techniques adopted to discover similarities between past and future design cases in the context of chip-design documentation. Mainly, N-gram and POS approaches correspond to documentation patterns, such as the writing styles of design engineers, and provide the content of the linking nodes.

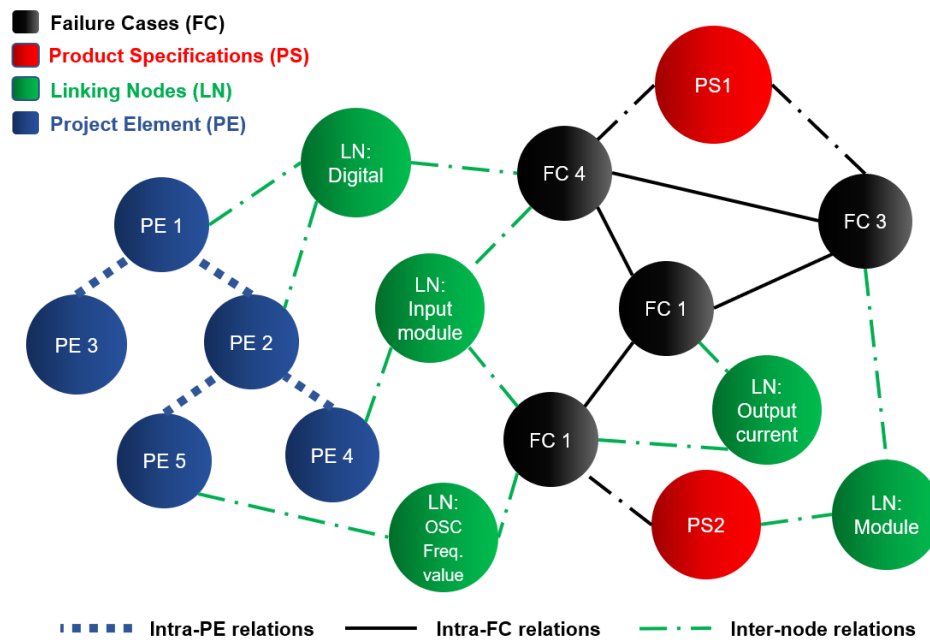


Figure 15. Multidimensional knowledge graph concept for use case 1.

4.3.3 Text mining pipeline for KG construction

Text mining is a collection of methods and algorithms intended for information extraction and knowledge discovery from structured, semi-structured and unstructured textual data. In the mining process, several techniques work individually and collaboratively to process raw data and utilize it in the knowledge extraction. Furthermore, the combination of outputs of individual methods can be used to further inform future models in an industrial context. The nature of the data handled by text mining determines the need for pre-processing stages, which prepare the textual input for the mining process. Being unstructured data that results from free input usually, several steps are needed to enhance the quality of the textual input, and thus the outcome of text mining.

In domain-specific applications, such as the chip-design process in this use case, extra pre-processing steps are integrated to handle this specific domain. The overall developed algorithm is shown in Figure 16.

Chip design documents have multiple sources and types. Specialized documents from different sources hold valuable information, which include, but are not limited to:

- **documents with a hierarchical structure:** which hold information about the design patterns and potential relations between different parts and modules of the designed chip.
- **unstructured documents:** including free textual user input that is not consistently structured. The nature of this type of document requires domain-specific natural language processing of its content.
- **documents with a column-based labeled structure:** effectively labeling textual content. This content can represent structured lessons-learned tables from previous design processes.

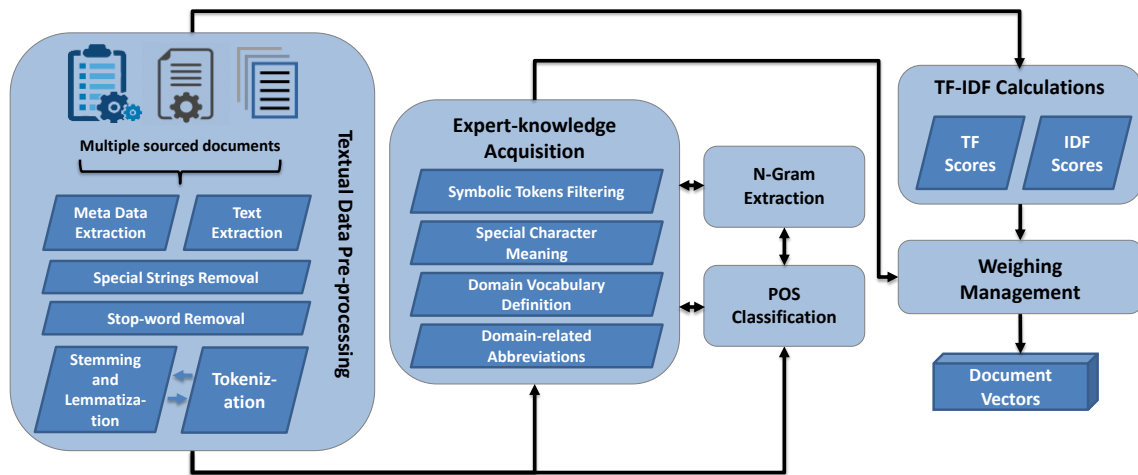


Figure 16. Domain-specific text mining pipeline for multi-source documents.

After defining the multiple formats of documents from the available datasets, the document content is extracted. Then, non-symbolic special characters are to be removed, along with specific patterns that are automatically generated from the database's software. The remaining text is then filtered for punctuation and tokenized. The tokenization process means recognizing individual entities in the texts (i.e., tokens), which can be single words or sentences, for example. In tokenization, some identifiers, such as punctuation, can be used to separate individual tokens. Tokens are then scanned to eliminate less meaningful tokens, which are referred to as stop words. Stop words can be: punctuation or pronouns, frequent words or, in some cases, non-frequent vocabulary (Zenkert et al., 2018). The resulting tokens are then lemmatized and stemmed in different phases of the procedure. Lemmatization is the process by which the tokens are converted back to their original linguistic form. If the token is a verb (e.g. designing), it is converted to its base form (design). In stemming, prefixes and suffixes are removed from the token in order to find its basic form as well.

Due to the technical language used in the chip design documentations, an additional part is added to the text mining pipeline, which utilizes the expert knowledge in defining domain-specific terminology. Resulting tokens from text pre-processing are the input for the expert-knowledge-supported part, where firstly, standard abbreviations are identified and isolated. Then, special symbols that are domain-specific, such as multiple underscores, are also reconstructed so that they are handled similarly amongst the

multiple document sources. With the elimination of duplicates and meaningless terms in extracted tokens, a representative model of each document can be constructed. Document models aim to represent a whole document as a vector of features, which is reliable enough to describe the extended content of the original document. Multiple approaches can be used for this modelling, including topic extraction, entity recognition, TF-IDF, and others. The choice of a certain method depends on the application fields and the nature of the available data. In the use case, a combination of approaches is hybridized to gain a better understanding and thus a more accurate modeling of the design documentation. This combination includes term frequency-inverse document frequency (TF-IDF) scores, N-grams, and part-of-speech (POS) classification.

The TF-IDF algorithm, Eq. 14, is utilized to extract the most descriptive tokens from a document, as well as from the overall sum of available documents. In the microchip design use case, term frequency (tf) is capable of detecting meaningful repeating words in a design document, such as the term “oscillator” or “processor”. However, this part of the TF-IDF also detects non-relevant frequent words, such as “volt”, which do not hold a unique meaning in the document, although it is considerably frequent. This is why the (idf) part is added to the equation, in order to balance the discovered terms identified in the (tf) part.

The inverse document frequency (idf) in this use case identifies the tokens that appear in a limited number of documents, among the total document datasets. Based on the assumption that those terms are most likely to represent the topic of the document, they are given a high score by the IDF part of the algorithm.

$$TF - IDF_{t,d} = tf_{t,d} \times idf_t = tf_{t,d} \times \log\left(\frac{N}{df_t}\right) \quad (14)$$

In Eq.14, t is the term in question, d is the document in which the term t is looked up, and N is the total number of documents. Scores for all tokens are calculated from the TF and the IDF parts. Then, a certain threshold determines the choice of tokens that will represent each document in a vector format. Document vectors represent a unified approach for describing the content of documents from multiple sources. This approach has been chosen for its benefits in the construction of the knowledge graph, as well as the fusion of information extracted from different documentation processes. This also supports the search functions to perform the same query on all data sources, since they have a coherent representation in the graph.

However, document vectors are not the only type of information extracted from the available data sources. Each of the available data sources contains extra information based on the source's nature itself. For example, a project description, as a documentation source, inherits information about the project's structure, which is used to create links between specific documents. Moreover, failure documentations that contain free textual descriptions of maintenance procedures also contain patterns that appear when the same person writes the case description, or when the workers in the design department use

certain terminology amongst each other. Those patterns can also be mined and utilized to extract relations between maintenance documents. In order to achieve this goal, an N-gram approach is used in the text mining algorithm. N-grams are combinations of N tokens that appear in the same pattern over the textual data. In the developed approach, the largest useful number of combinations that reveal possible patterns in the description of the failure cases is identified. Then, all N-grams equal or below this threshold, i.e., unigrams, bigrams, trigrams, etc., are utilized in order to enhance the relation extraction between documents that contain those N-grams.

The third method in the hybridized approach handles domain-specific terms as a classification approach. In this approach, a deep learning model is used to recognize the part of speech (POS) for each token. This is found to classify special terminology that design engineers use in the documentation.

Although obtaining data from multiple sources poses the potential to support textual data with other data modalities, it also reveals a main challenge concerning the fusion of those modalities into a single representation, as well as the extraction of relations between them. In this use case, a text mining-based algorithm provides the overall system with the ability to reduce the imprecision in textual data, represent the documents in meaningful token vectors, and extract possible patterns in the text represented by the extracted N-grams. Document vectors hold the potential to extract relations between the documents. However, searching for a certain document through those relations is not very different from a keyword search approach, which is a limited approach in this particular use case. Therefore, to enhance the connectivity, searchability, and retrieval of design documentation, a knowledge graph is utilized for its ability to represent relations between those documents based on several criteria.

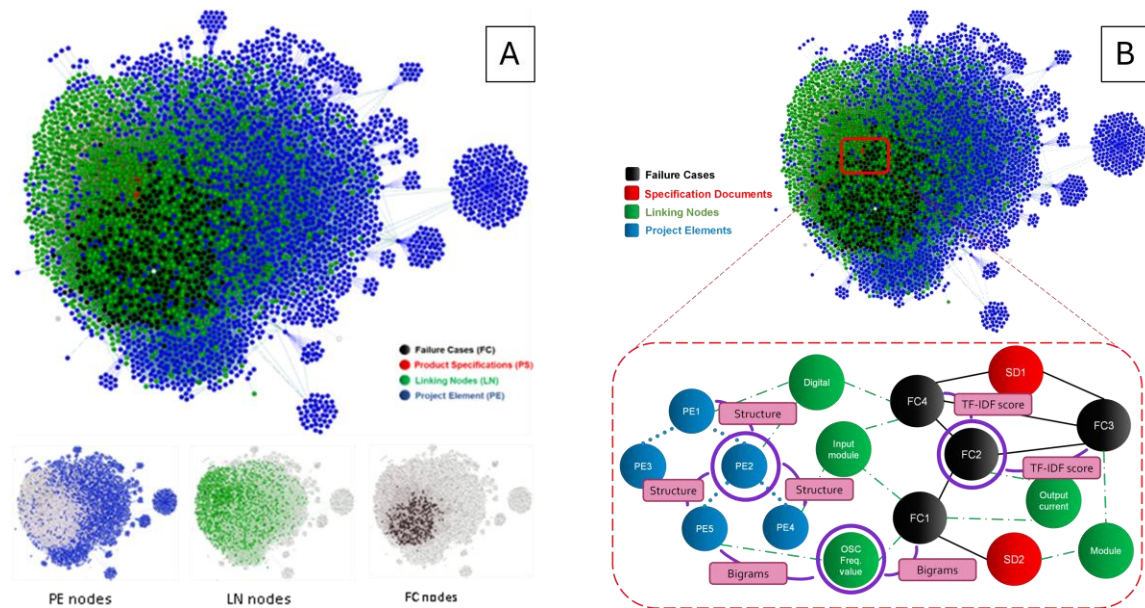


Figure 17. Construction of the multidimensional knowledge graph from the three node types (A), and the defined types of nodes and the relations among nodes of the same type and those of different types (B).

With this concept of knowledge graph structure, a portion of the complete graph is shown in Figure 17, where the different types of nodes and relations are highlighted within the graph. Within the large number of nodes and relations in the knowledge graph, the ability to retrieve documents most relevant to a specific design process becomes essential for design engineers. Therefore, search infrastructure has to be provided in order to support the ease of access to previous design experiences and lessons learned.

4.3.4 Graph-based Search, information retrieval, and explainability

Among the tasks that knowledge graphs have been utilized for, enhancing search and matching is a promising field, especially with the advances in search engines and semantic web. Supported by text mining, researchers such as Tiwari et al. (Tiwari et al., 2015), developed a concept to integrate the Resource Description Framework (RDF) with natural language understanding, in order to find a common format that supports human and machine understanding together, and thus support the representation and retrieval of information from heterogeneous and distributed data sets. The approach developed for the use case intersects with the previous examples in terms of utilizing domain-specific text mining in constructing the knowledge graph for search and information retrieval. However, in the context of the use case, domain information is integrated in the graph structure itself, in order to enable:

- Domain-specific relation extraction alongside the text-mining-based relation extraction.
- Domain-specific search and information retrieval.
- An easy integration of new knowledge in the graph, supporting a dynamic process of transferring lessons-learned from past and present to future design scenarios.

Graph-based search plays a considerable role in the proposed system in retrieving not only directly related search results, but also relevant results that are not shown to the user with traditional keyword searches. Search functionality in this use case is illustrated in Figure 18. Search strategies define a direct hit result, which is retrieved from graph nodes corresponding to a direct match to the search query. In Figure 18, the node *FC2* represents this direct hit. The strategy then utilizes graph shortest-path calculation, starting from the direct-hit node, in order to find relevant nodes to the search query. Shortest paths represent a certain length of the defined relations between two nodes. Therefore, the well-defined types of graph relations can be a key to finding those relevant results in the constructed graph. Depending on the length of the shortest paths, relevant search results, including nodes *FC1*, *PE5* and perhaps also *PS2*, can be retrieved and made visible to the design engineer, based on their search preferences.

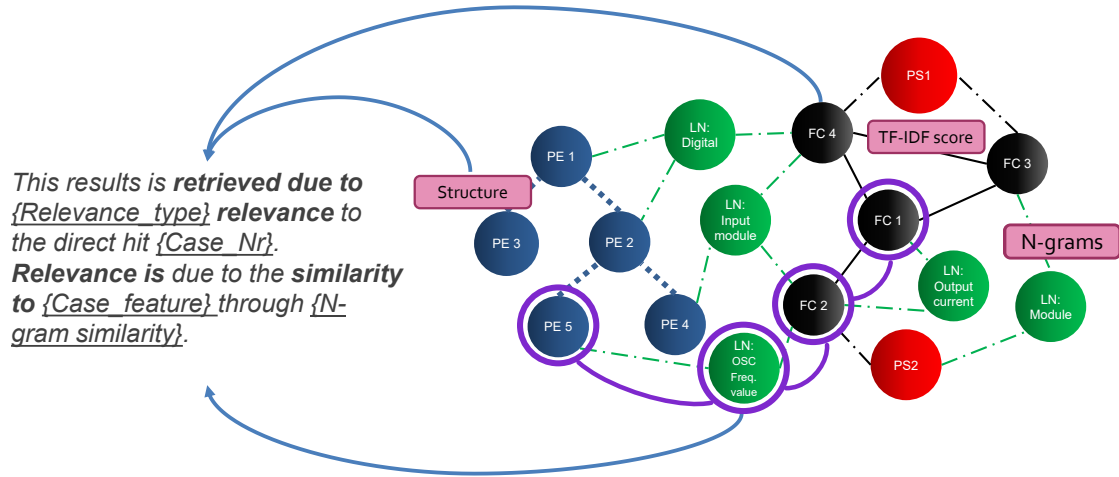


Figure 18. Graph-based search and explanation strategy.

Different types of nodes and relations, as well as the shortest calculated paths, are also designed to provide a mean of result explanation to the user. Explainable retrieved results are meant to support the user's ability to assess the result relevancy to their own design scenario. In the context of the use case, explainability is utilized to support a feedback loop, in which the expert user can effectively evaluate the retrieved results, in order to enable future relation analysis and correction in the graph structure. This is not only possible during the testing phases of the system, but also as a continuous feedback and enhancement process during the actual use of the deployed graph search engine, especially with new lessons learned being added dynamically to the graph.

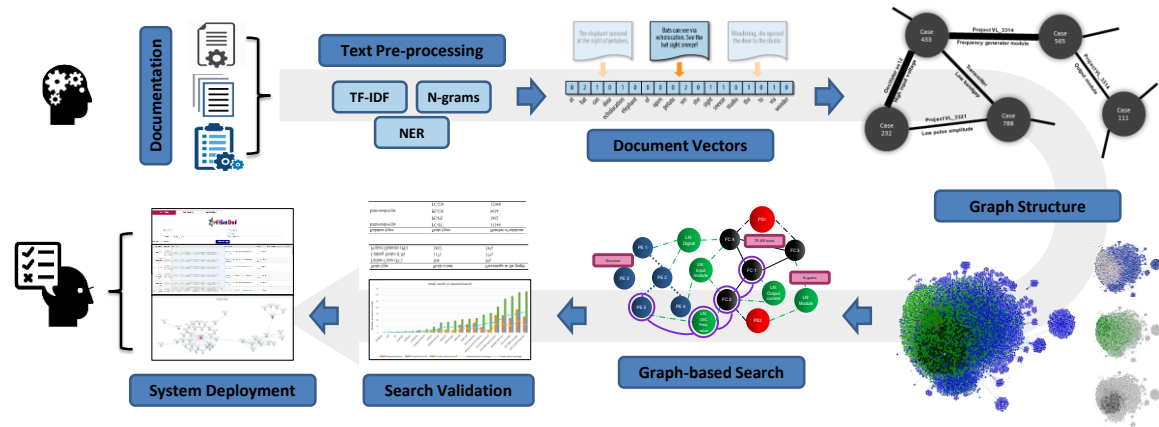


Figure 19. Overall KG design and utilization in explainable lessons-learner recommendation.

Graph-based explainability used the information collected from node types, relation types and the linking nodes themselves, in order to fill the gaps in several predefined, human readable templates. This allows the user to understand the reasoning of the system, in terms of why each search result has been retrieved.

Explaining the reasons behind retrieving graph-based results is not limited to the textual formats, but can be supported by a visual explanation that is constructed from the overall knowledge graph. It can visually illustrate the connections among design

documents, as well as their different types of relations to the search terms. In the developed overall solution for the use case, see Figure 19, search strategies take the form of a search engine, which provides the possibility to retrieve relevant design cases from a current one. The engine allows users to search previous lessons learned with a short description of either the failure itself, the module in which the failure occurs, or other failure-related details.

4.4 Use case 2 - Learning path recommendation for staff training in elderly-care domain

4.4.1 use case description

The current landscape of vocational education and training (VET) in healthcare presents significant challenges that necessitate innovative approaches to learning and competency development (Buchmann, 2022). A fundamental issue in this domain is that initial training acquired through formal education frequently becomes inadequate for evolving job requirements in real-world healthcare settings. This discrepancy between formal education and practical needs is particularly pronounced in the context of non-clinical care, where professionals must continuously adapt to changing regulations, technologies, and care methodologies while maintaining high standards of patient care.

Healthcare professionals operate within complex environments characterized by dynamic knowledge requirements and interdisciplinary interactions. This use case focuses specifically on non-clinical care workers across diverse professional groups within elderly care organizations, including nursing staff, social services personnel, administrative employees, housekeeping staff, trainees, and auxiliary support staff. This heterogeneous professional demographic represents a significant challenge for traditional educational approaches, which often struggle to address the specific needs of such diverse learner populations simultaneously.

The educational challenges in this domain are multifaceted. Learning takes place within this context, where multiple learning objectives frequently overlap based on job-specific requirements. Simultaneously, learners' backgrounds, existing knowledge states, and learning preferences differ considerably across professional roles and experience levels. Addressing these challenges requires educational content that is meaningful to learners, relevant to their setting, and tailored to their individual needs. To that end, traditional linear or hierarchical approaches to content organization and delivery often fail to capture the interconnected nature of knowledge in healthcare settings, where professionals must synthesize information from multiple domains to provide comprehensive care.

To address these challenges, this use case explores how digital learning environments can be transformed into more contextualized and personalized training spaces. The study focused on mandatory training topics that all professional groups in

non-clinical care must regularly complete, specifically: dementia care, hygiene protocols, fire safety, and data protection regulations. What distinguishes this approach from conventional e-learning systems is the intentional interweaving of these topics to create learning paths that transcend simple memorization of facts. By creating intersections, such as dementia-data protection or hygiene for people with dementia, the approach enables learners to examine content through different disciplinary perspectives while ensuring the material remains relevant to all staff regardless of their position or department.

This interdisciplinary approach realizes an inclusive educational philosophy while fostering interdepartmental understanding, recognition, and appreciation among different professional groups. It maintains explicit target-group orientation through specific topic combinations that relate to each learner's context, enabling a form of personalization that extends beyond mere content selection to include contextual awareness of the learner's situation and the appropriate presentation of learning materials.

The research is grounded in situated (Lave & Wenger, 1991) and subject-oriented learning theories, which define context as the situation in which learning happens. This includes both the learner's situation (location, previous knowledge, etc.) and the learning object's situation (type, length, implementation scenario, etc.). Context is described by factors that can relate two or more learning objects to each other if they serve the same learning situation, creating a network of meaningful connections that reflect the complex reality of healthcare practice.

4.4.2 KG construction methodology

The construction of the KG is conducted to extend a taxonomical structure that provides a five-level hierarchical framework for organizing learning content, see Figure 20. This taxonomy was selected because it is "sufficiently simple and semantically distinctive for the construction of a solid graph data model". The taxonomy organizes learning content into a carefully structured hierarchy that facilitates both traditional learning progression and the more complex semantic relationships that underpin contextualized learning experiences.

The taxonomy consists of five distinct levels, each representing a different granularity of learning content and educational purpose:

1. **Journey:** Represents an overarching learning goal (e.g., "Dementia in elderly care"), serving as the broadest organizational unit and establishing the primary educational objective.
2. **Course:** Functions as a container representing one specific aspect or dimension of the broader learning goal, organizing related content into coherent units.
3. **Topic:** Represents concrete knowledge elements within a course that address specific areas of learning, providing focused content on particular subjects.

4. **Educational Package:** Groups related and sequentially organized Open Educational Resources (OERs) into coherent learning units, establishing logical progression through the material.
5. **Educational Content:** Represents a concrete OER that can be directly studied by learners, providing the most granular level of content engagement.

For simplification, levels 2-5 (Course, Topic, Educational Package, and Educational Content) are collectively referred to in the following as "Learning Objects" (LOs). This taxonomical structure provides significant advantages for modeling learning content as it allows for the representation of two distinct types of relationships: hierarchical relationships (parent-child connections that establish the organizational structure) and semantic relationships (contextual connections between learning objects that may cross different domains or hierarchical branches).

While hierarchical data models are commonly used in digital learning platforms, they present several limitations in the context of healthcare professional education. Traditional hierarchies organize learning content into groups of materials that deploy a hierarchical nature (e.g., curricula organized into courses, lectures, topics, and materials), which is suitable for representing the details of a learning objective but offers limited support for learning individualization or cross-domain contextualization.

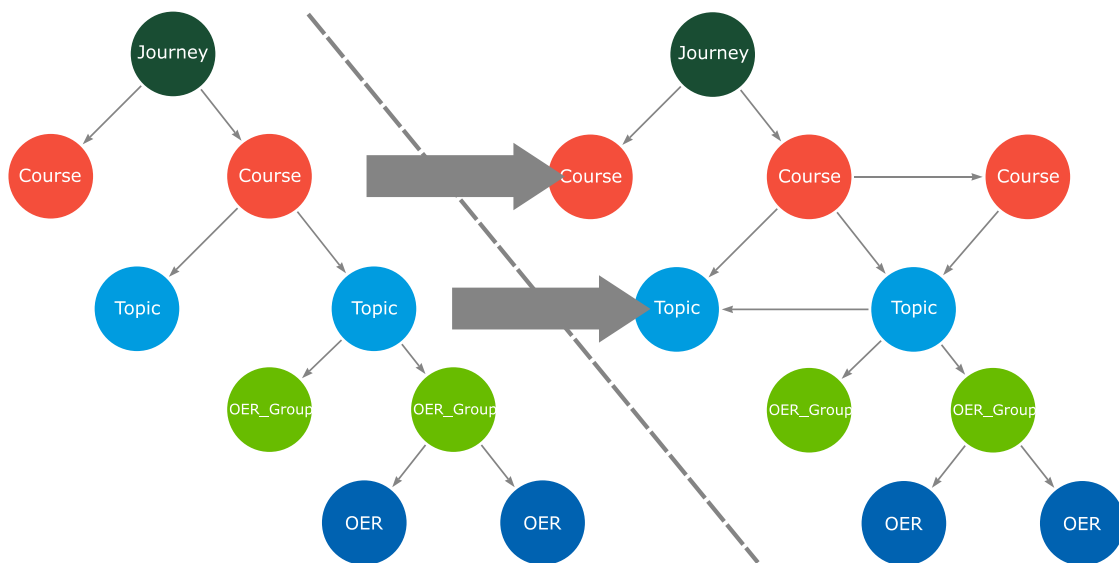


Figure 20. Hierarchical and network models for the learning materials in the use case (Buchmann et al., 2024).

Hierarchical models struggle to indicate learning recommendations that span different curricular areas but may be required in a single application area or learning context. For example, a healthcare professional dealing with dementia patients may need content from formal dementia care curricula, communication skills training, and data protection regulations—areas typically separated in hierarchical structures but intimately connected in practical application.

Use case 2 - Learning path recommendation for staff training in elderly-care domain

Furthermore, hierarchical models fail to capture the rich, interconnected nature of healthcare knowledge, where concepts from different domains frequently overlap and interact. This limitation is particularly problematic in healthcare settings where professionals must synthesize knowledge from multiple domains (medical, social, ethical, and administrative) to provide comprehensive care.

To overcome these limitations, the research proposes a transition from hierarchical data models to knowledge graph structures that better represent the content and context of learning objects, as well as their interrelationships. This approach enables the representation of complex learning contexts where knowledge from multiple disciplines may be relevant to a particular learning objective or professional challenge.

The semantic connections established in knowledge graphs are required for the representation of complex interdisciplinary relationships that are crucial for developing higher-order thinking skills. By providing contextual connections between learning materials across different domains or subject areas, the system supports learners in developing a more nuanced and integrated understanding of complex topics.

The knowledge graph construction extends beyond the original hierarchical taxonomy model by incorporating rich semantic relations discovered through tailored text mining processes. The approach developed in this use case connected courses and topics to transform the hierarchical learning content organization into a semantically interconnected network that better represents the contextual relationships between different learning materials.

The node types in the knowledge graph directly correspond to the five levels of the taxonomy described earlier. Each node represents a specific learning object at its respective level in the hierarchy (Journey, Course, Topic, Educational Package, or Educational Content). This one-to-one mapping ensures that the original taxonomical structure is preserved while enabling the addition of semantic enrichment.

Each node functions as a self-contained entity that encapsulates all the relevant information about a particular learning object, allowing it to operate independently within the broader knowledge network. These nodes serve as the fundamental building blocks of the graph, establishing the entities that can be connected through various relation types. The node structure preserves the pedagogical intent of the original content organization while enabling more flexible traversal patterns through the addition of semantic relations.

The knowledge graph implements two primary categories of relations that connect the nodes throughout the structure:

1. **Hierarchical relations:** These preserve the original parent-child relationships from the taxonomy as created by content developers. For example, a Course node would have hierarchical relations to the Topics it contains, and a Topic would have hierarchical relations to its Educational Packages. These relations maintain

the organizational integrity of the learning content and reflect the intentional structural design created by educational experts during content development.

2. **Semantic relations:** These are new connections established through text mining processes, labeled with the "has_semantic_relation_to" relation type. These relations create direct and indirect connections between learning goals (Journeys) and learning objects that might not be directly connected in the hierarchical structure. For example, although the Journey "Dementia in elderly care" might not have a direct hierarchical connection to the Journey "Communication in elderly care," semantic analysis can identify several indirect paths through connected Courses and Topics that serve similar learning contexts. These semantic relations are particularly valuable as they enable the discovery of cross-disciplinary learning paths that might not be immediately obvious from the hierarchical structure alone.

The combination of these relation types creates a rich network structure where learning objects can be connected not only vertically within their own hierarchy but also horizontally across different domains and subject areas. This enables the representation of complex learning contexts where knowledge from multiple disciplines may be relevant to a particular learning objective or professional challenge.

Node and relation properties

The nodes in the knowledge graph contain extensive metadata that enriches their representation and facilitates semantic analysis. Key properties include:

1. **Textual metadata:** Titles and descriptions in multiple languages (primarily German and English), which provide the semantic basis for relation extraction. These properties contain the most comprehensive textual information about learning objects and serve as the primary input for the text mining processes.
2. **Similarity scores:** Values derived from text analysis, which quantify the strength of semantic relationships between nodes. These scores help determine the relevance of connections between learning objects and can be used to rank or filter recommended content.
3. **Contextual information:** Additional properties integrated within the graph, such as learning objectives, difficulty levels, or relevant professional contexts. These properties provide valuable supplementary information that enhances the representation of learning objects and their relationships.

These properties not only provide valuable information about individual learning objects but also serve as the foundation for the text mining processes that establish semantic relations between nodes. The textual properties, in particular, are critical for

identifying meaningful connections between learning objects that might not be explicitly defined in the original taxonomical structure.

4.4.3 Modeling context in the knowledge graph

The resulting graph structure enables modeling the context of learning objects through their connections to other objects that appear in the same learning scenario, creating a network of contextually related materials that can be traversed by recommendation algorithms. This context-aware representation allows the system to identify personalized learning paths that connect learners with relevant learning content based not only on explicit learning goals but also on implicit contextual relationships and individual learner characteristics.

For example, the knowledge graph might reveal that certain topics in dementia care have strong semantic connections to specific data protection considerations, allowing the recommendation system to create personalized learning paths that integrate these related areas in a coherent and meaningful way. This integration reflects the reality of healthcare practice, where professionals must simultaneously consider multiple domains of knowledge when addressing specific care scenarios.

The knowledge graph structure also enables the representation of learning contexts through densely connected partial networks of nodes that serve similar learning scenarios, see Figure 21. These networks can be identified through the analysis of textual descriptions that reveal semantic similarities between learning objects from different domains or hierarchical branches. By capturing these contextual relationships, the knowledge graph enables learners to achieve higher knowledge levels that facilitate comprehension, inferencing, and problem-solving, as opposed to learning isolated facts that may only enable basic memorization.

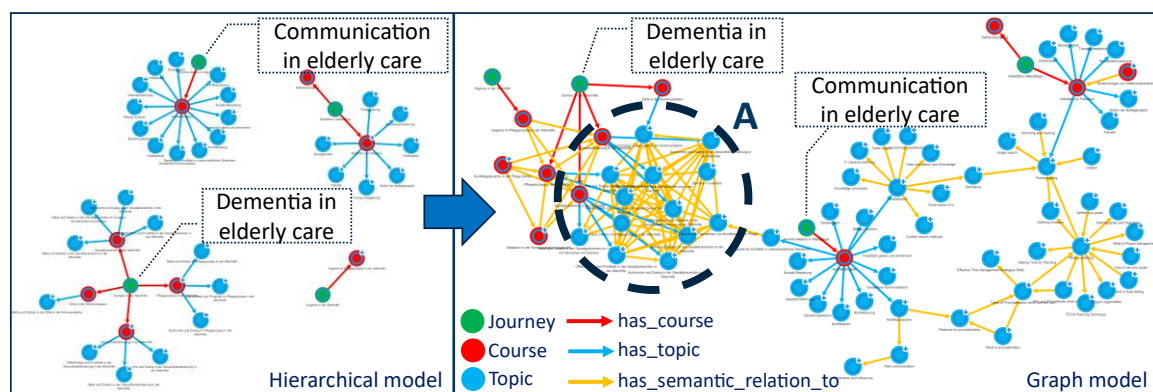


Figure 21. First three levels of the KG structure, representing the role of semantic relations in creating the KG from the hierarchical one (Abu-Rasheed, Dornhöfer, et al., 2023).

4.4.4 Text mining and analysis for semantic relation extraction

A customized text mining pipeline (TMP) has been created in this use case to extract semantic relations between learning objects based on their textual descriptions, creating a rich knowledge graph that represents the complex relationships between different learning materials, see Figure 22. This process involves multiple steps and leverages several advanced natural language processing technologies to identify meaningful connections that might not be immediately apparent in the original hierarchical structure.

The proposed TMP is specifically designed to maximize the utility of learning objects' textual metadata. This pipeline processes both short texts (titles) and longer texts (descriptions) in multiple languages (primarily German and English) to extract semantic relationships. The pipeline consists of several key components organized in a sequential processing flow:

1. **Language detection:** The process begins with language identification using a pre-trained model. This step is important for determining which language-specific processing should be applied to each text. The system recognizes whether content is written in German or English and routes it through the appropriate processing path.
2. **Text cleaning:** Once the language is identified, the text undergoes cleaning processes that remove special characters and strings that might negatively affect similarity algorithms. Importantly, the structure of sentences is preserved during this process since text similarities are calculated at a semantic level rather than using simple bag-of-words approaches. This ensures that the semantic integrity of the text is maintained throughout the analysis process.
3. **Parallel processing for titles and descriptions:** The pipeline implements different processing paths for titles (short texts) and descriptions (longer texts) due to their different characteristics:
 - **For titles:** The system creates title embeddings directly using the SentenceBERT model and the Spacy library for natural language processing. These embeddings capture the semantic essence of the titles in a numerical form that can be compared using vector similarity measures.
 - **For descriptions:** Due to their greater length and complexity, descriptions undergo an additional step of topic extraction before embedding creation. The KeyBERT model extracts the main topics present in the description, and then embeddings are created for each extracted topic separately. This approach allows for more nuanced comparison of longer texts by breaking them down into their constituent topics.

4. **Cross-language comparison:** When comparing texts in different languages (e.g., German and English), the system faces the challenge that embeddings created in different language spaces may not be directly comparable. To address this, the pipeline incorporates automatic translation of German titles using the DeepL API translation service before calculating similarity. This ensures that semantic comparisons can be made across language boundaries, which is particularly important in multilingual learning environments.
5. **Similarity calculation:** The pipeline implements different similarity calculation approaches depending on the text type:
 - **For titles:** Direct cosine similarity is calculated between title embeddings, providing a straightforward measure of semantic similarity between short texts.
 - **For descriptions:** The system creates an intersection matrix to determine which topic-pairs from the two topic sets should be compared, and then calculates the average similarity across all topics. This approach accounts for the more complex structure of longer texts and ensures that similarity is assessed based on the relevant topics rather than superficial textual characteristics.
6. **Relation creation:** The final similarity scores from both title and description comparisons are used to determine whether a semantic relation should be created between two learning objects. This decision is based on experimentally defined thresholds that have been validated by domain experts to ensure that only meaningful relationships are included in the knowledge graph.

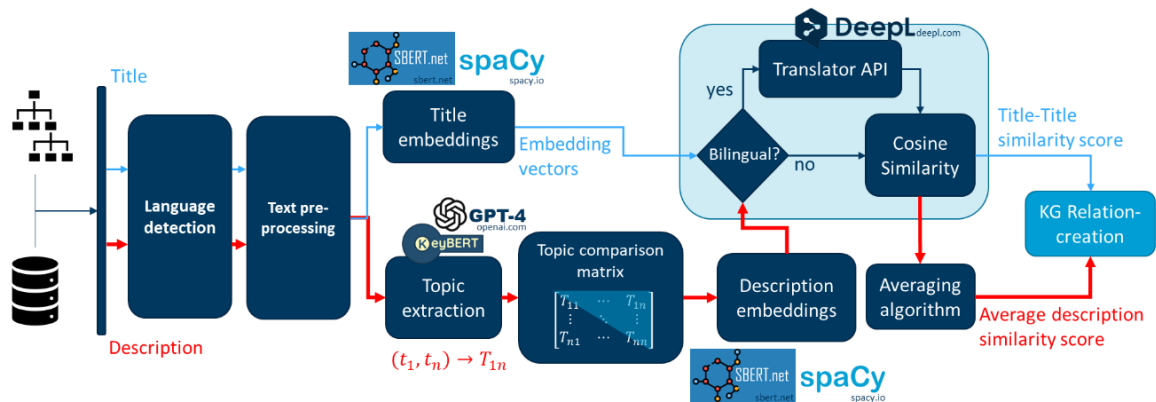


Figure 22. Text mining pipeline (TMP) and semantic similarity calculation for title and description properties of the learning objects.

4.4.5 Enhancement with large language models

In response to the significant advancements in generative AI technologies, the research methodology was extended to incorporate Large Language Models (specifically

GPT-3.5 and GPT-4) to enhance the relation extraction process. This integration serves two main objectives:

1. **Contextual content analysis:** These models are used to better understand the textual content of learning materials within their specific context and extract more precise topics from the text based on the position of the materials in the knowledge graph. This contextual understanding helps identify more accurate and relevant semantic relationships between learning objects.
2. **Explanation generation:** The LLMs are also used to generate explanations for the extracted relationships between learning materials, improving the transparency and interpretability of the recommendation system. These explanations help learners understand why specific content is recommended and how it relates to their learning goals.

The integration of LLMs brings significant advantages in terms of natural language understanding, but also introduces challenges related to the generation of potentially irrelevant or inaccurate content. To mitigate these risks, the research implements several strategies:

- **Task limitation:** The tasks assigned to LLMs are carefully constrained to avoid scenarios that might lead to the generation of random text, hallucinations, or other problematic outputs. By defining narrow and specific tasks, the system reduces the risk of generating misleading information.
- **Contextual prompting:** When interacting with LLMs, the system uses comprehensive contextual information derived from the knowledge graph structure to guide the text generation process and increase output accuracy. This contextual grounding helps ensure that generated text remains relevant and aligned with the specific learning materials being analyzed.
- **Template-based explanation generation:** Rather than using complete LLM-generated outputs directly, the system creates explanation templates that incorporate only selected parts of the generated text. These templates are designed to be filled with specific information that can be verified against the knowledge graph, reducing the risk of presenting inaccurate information to learners.

4.4.6 Explainability components

The research also developed explainability components that allow users to visually explore how their learning materials connect within the knowledge graph. These visual explanations help learners understand why specific content is recommended and how it relates to their learning goals, enhancing transparency and trust in the recommendation system.

The explainability approach consists of both textual and visual components:

1. **Textual explanations:** Generated using a combination of template-based approaches and LLM-generated content, these explanations provide clear rationales for recommendations, answering questions such as "Why is this learning path connected to the learning goal?" and "How do the elements of the learning path relate to each other?"
2. **Visual explanations:** Interactive graph visualizations that display the recommended learning path within the broader knowledge graph context. Users can interact with these visualizations by clicking on nodes or relationships to reveal additional information through tooltips, providing deeper insight into the semantic connections between learning materials.

These explainability components address a critical challenge in AI-driven educational systems: the need for transparency and understandability in automated recommendations. By providing both textual and visual explanations, the system helps learners develop a more nuanced understanding of the relationships between different learning materials and how these relationships relate to their learning goals. The overall system developed for this use case is shown in Figure 23.

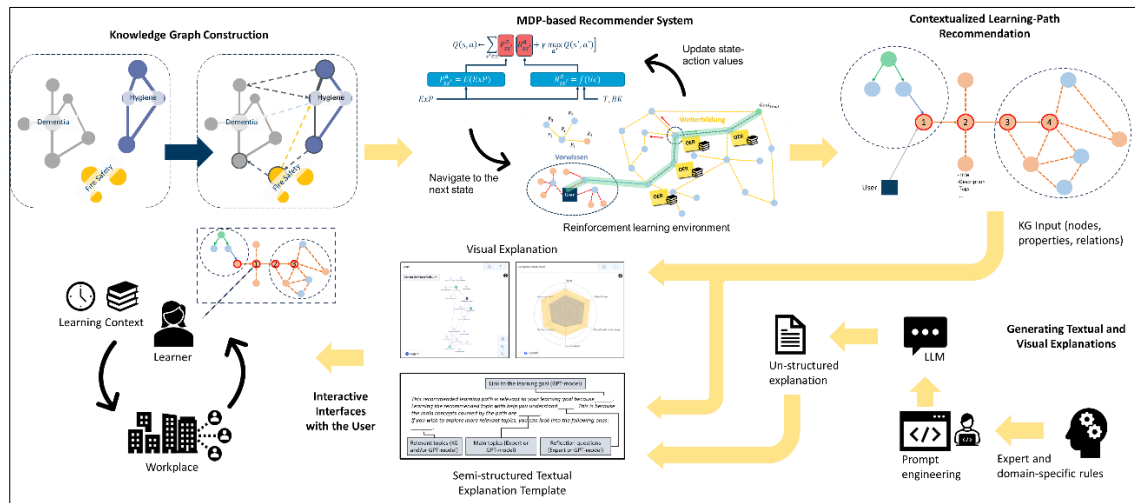


Figure 23. Overall KG design and utilization in creating explainable learning path recommendations for nursing VET.

4.5 Use case 3 - Higher education module scheduling for learning personalization

4.5.1 Use case description

This use case addresses learning in higher education and aims at improving learning personalization through curriculum modeling and knowledge graph completion. The primary objective is to develop personalized learning-path recommendations by linking

Use case 3 - Higher education module scheduling for learning personalization

university subjects and their topics to domain models. This approach facilitates the integration of learning modules across faculties and institutions, enhancing the adaptability of educational paths for students.

Higher education systems often struggle to meet individual student needs due to their structured nature. Students enrolled in programs with diverse backgrounds or international cohorts may encounter redundant content they have already studied, leading to inefficiencies in their learning paths. To address this issue, it is essential to tailor educational pathways based on a student's prior knowledge, completed coursework, career orientation, and personal goals.

However, several obstacles impede this level of personalization:

1. Educators often lack sufficient time to analyze students' prior knowledge or cross-reference content from other courses.
2. There is no standardized representation of courses or study modules across universities or even within the same institution.

The process utilizes large language models (LLMs) to assist human experts in extracting detailed topics from lecture materials. The research focuses on creating domain, curriculum, and user models for university modules and stakeholders. To demonstrate the approach, two study modules, namely Embedded Systems and Development of Embedded Systems Using FPGA, were used to construct a KG that organizes the curriculum and connects it to domain models.

The methodology was evaluated using both qualitative feedback from domain experts and quantitative metrics assessing graph quality. Experts validated the relevance and accuracy of the proposed model, while structural metrics confirmed the quality of the KG. This use case highlights the role of LLM-assisted KG completion methods in effectively connecting related courses across disciplines and thereby enhancing learning personalization, as well as the role of the KG in connecting learning modules in formal learning settings in comparison to informal or non-formal ones.

4.5.2 KG construction methodology

To overcome the challenges addressed above, a flexible representation of higher education curricula is proposed, supported by intelligent algorithms that assist teachers in analyzing educational materials efficiently.

The foundation for constructing a KG lies in developing an ontology tailored for higher education that integrates three key models: curriculum, domain, and user models, see Figure 24. These models ensure interoperability among diverse educational programs while maintaining granularity in content representation.

Use case 3 - Higher education module scheduling for learning personalization

- **Curriculum model:** Represents educational content as defined by universities, including classes such as "Topic" and "Sub-Topic." A "Topic" refers to abstract concepts taught during sessions, while "Sub-Topic" accounts for detailed content described through tangible means like lecture slides.
- **Domain model:** Captures high-level knowledge domains (e.g., Automotive or Embedded Systems) and their sub-domains (e.g., Self-Driving Cars within Automotive). This hierarchical structure allows for organizing content at varying levels of abstraction.
- **User model:** Focuses on students as primary stakeholders, incorporating factors such as background knowledge, learning goals, preferences, academic performance (e.g., exam results), and other parameters relevant to learning personalization.

The ontology provides a standardized framework for representing educational content across institutions while enabling meaningful comparisons between courses.

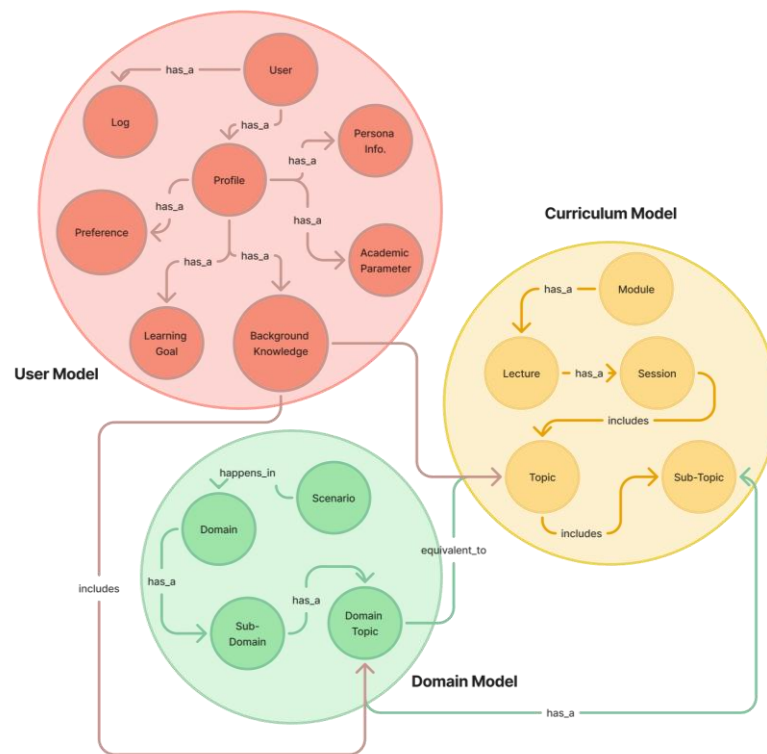


Figure 24. The higher education ontology with user-, curriculum-, and domain-models.

KG construction process involves two primary steps: node creation and relation extraction.

Node creation: Nodes are populated with data derived from educational content (topics/sub-topics), domain models, and user profiles based on the ontology. Each node represents a distinct element within the curriculum or domain structure.

Relation extraction: Relations between nodes are created at two levels:

Use case 3 - Higher education module scheduling for learning personalization

1. Hierarchical relations are derived directly from the ontology's structure (e.g., linking topics to sub-topics within a session).
2. Semantic relations are identified by analyzing similarities between topics/sub-topics across different lectures using natural language processing techniques.

Semantic similarity analysis utilizes both NLP pipelines and LLM-based methods to detect connections between nodes based on their titles and descriptions. For validation purposes, a semi-supervised approach is employed where human educators review a random sample of predicted relations within the KG. This iterative validation process refines parameters for relation extraction and ensures high-quality graph construction.

Once constructed, the KG serves multiple purposes:

1. It acts as a homogeneous database supporting personalized learning-path recommendations through recommender systems.
2. It provides contextual information for future topic extraction tasks when analyzing similar educational content.

4.5.3 Automated pipeline for content extraction and classification

To alleviate the time-intensive task of classifying lecture materials into topics and sub-topics manually, an automated pipeline was developed, see Figure 25. This pipeline begins with extracting textual data from various sources such as lecture slides, manuscripts, and video recordings. For videos, transcription is performed using OpenAI's Whisper model, a highly accurate speech recognition tool capable of zero-shot recognition tasks.

Once textual data is extracted, it is processed using OpenAI's GPT-4o model for topic extraction. The LLM is prompted to identify key concepts taught during each session:

- **Topics** are defined as overarching abstract concepts covered in lectures.
- **Sub-Topics** represent detailed explanations found within lecture materials (e.g., slides).

The hierarchical relationship between topics and sub-topics is explicitly defined in LLM prompts based on ontology rules to ensure accurate classification. Extracted concepts are classified into topics and sub-topics according to their definitions within the ontology framework. Titles and descriptions associated with these concepts are analyzed further to identify semantic similarities across different study programs or universities.

Descriptions generated from detailed lecture materials provide high-quality contextual information at the sub-topic level. These descriptions enhance interoperability among curricula by enabling meaningful comparisons between courses.

Use case 3 - Higher education module scheduling for learning personalization

Given the sensitivity of educational content analysis, human validation plays a critical role in ensuring quality:

- Teachers review LLM-generated predictions for topics/sub-topics and make final decisions regarding their classification.
- Educators also contribute to refining LLM prompts by providing definitions specific to their domain (e.g., terminology unique to certain fields).

Validated content forms the basis for KG construction while maintaining high standards of accuracy. Based on the KG construction, the overall system developed for this use case is illustrated in Figure 26, where the use case serves a problem-based learning scenario, providing explainable learning recommendations for specific user requests that include descriptions of the challenge the user intends to solve through the recommendation.

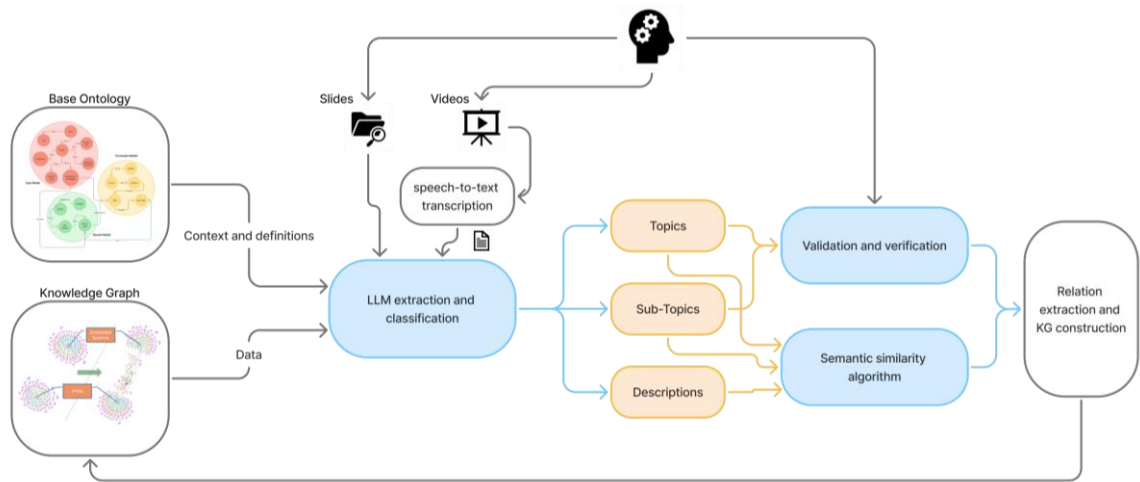


Figure 25. Pipeline components for the transcription, extraction, classification, and KG construction, based on a human-AI collaborative approach.

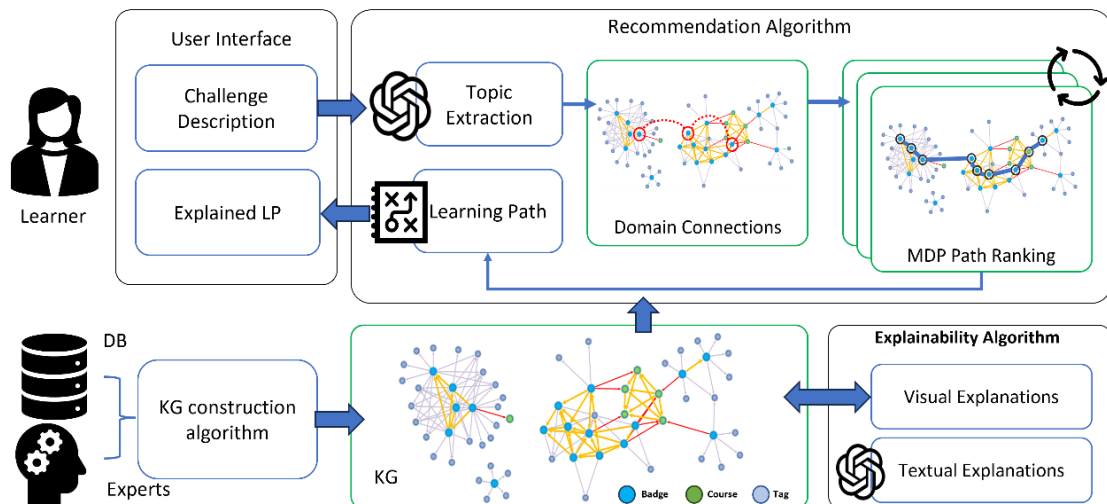


Figure 26. Overall system for problem-based learning path recommendation (Abu-Rasheed et al., 2024).

4.6 Discussion

This chapter demonstrated the construction strategies of the KG in three educational use cases. The goals of diversifying the use cases aimed to study the ability of the methods proposed for KG construction in reflecting the requirements from different domains, as well as in different educational settings.

As the construction strategies of the KG are part of the first research question, the implementation in each use case provided an important insight into the answer that the CARExKG method proposes to this question. This is because the process of selecting graph nodes, relations, and properties in the three use cases overlaps on several points, yet it is sufficiently dynamic to account for the unique goals, constraints, stakeholders, pedagogical requirements, and learning settings of each of those use cases. From the implementations demonstrated above, several observations can be made about the roles of KGs in providing the structural foundation for personalizing learning path recommendations:

- The role of stakeholders in the design of the KG is essential for a domain-specific application. The information and insights that TEL developers can collect from relevant stakeholders should be integrated in the early phases of KG, and ideally, in the overall system design. This is because the types of graph nodes and relations, along with their properties, will only reflect generic, surface-level representations of the domain if domain stakeholders were not taken into consideration while selecting those nodes, relations, and properties.
- The definition of the learning context in each domain requires considering the differences in the importance of different contextual factors in relation to the use case at hand. Therefore, starting by defining the learning context, in the sense of its meaning and factors, becomes necessary for building graph structures that can thoroughly represent that context, and account for the difference between the factors, as well as their importance in that specific use case.
- The balance between thoroughness and simplicity in creating the ontological or taxonomical schema of the KG is a step of the KG design that becomes more important for complex educational settings. This is because of the substantial implications that this design decision has on the complexity of graph querying algorithms, as well as the ability to extract explicit and accurate information from the KG at the same time.
- The range of algorithmic approaches to calculate node similarities, and thus extract graph relations, is not only dependent on the types of graph entities, but also has a strong connection to the goals of the educational stakeholders. Text-based semantic similarity algorithms can be considerably supported by input from the stakeholders, in terms of defining clearly what semantics should be addressed in the text, and how they may reveal different types of relations, e.g., between the LOs or the learners.

5 Markov Decision Process (MDP) over Knowledge Graphs for Learning Path Recommendation Systems

The construction of the knowledge graphs in chapter 4 establishes the foundation for building a reasoning algorithm over the graph structure. To construct personalized learning paths for the target users, a graph-based CARS is developed to generate the path recommendations. This thesis proposes an approach for a contextualized implementation of Markov decision process (MDP) over the KG, to generate long-term learning path recommendations. The selection of MDP is based on its alignment with:

1. situated learning theory requirements, in terms of integrating the learning context in the learning process
2. knowledge graphs as a data structure, which can serve as the learning environment for MDP
3. explainability requirements of the constructed learning path recommendation.

Through the graph-based MDP, this chapter aims to answer the second research question of the thesis:

“In what way can graph-based learning path recommendations be contextually tailored to improve learning?”

The answer to RQ2 is approaches with graph-based MDP because it provides the mathematical framework for modeling sequential decision-making under uncertainty. When combined with knowledge graphs, MDP offers a capability for developing intelligent systems that can reason over structured knowledge and make contextually appropriate recommendations.

In the following sections, the design of the contextually aware recommendation system in CARExKG framework is described in detail.

5.1 Recommendation problem formulation

In the proposed implementation of MDP over the KG, the recommendation problem can be broken down into two sub-problems: *path finding* and *path ranking*. Each part of the recommendation problem is strongly linked to the KG structure, which provides the infrastructure for navigating the learning content and selecting the elements of the recommended path.

In the scope of this thesis, the MDP-based recommendation algorithm is developed for the VET learning setting (see section 4.4). This use case represents a context-rich learning setting that combines dynamic user contexts to real-world challenges, which users face during their daily tasks. Thus, the user perception of the learning recommendation, and later its explanation, is accompanied by a high level of criticality, due to time constraints and workload. This research utilizes that aspect to evaluate the ability of the recommendation system to enhance learner’s performance in this setting, in terms of solving a work problem of a complex and interdisciplinary nature.

5.2 Graph path-finding algorithm

In the KG structure described in section 4.4, KG paths result from the semantic similarity relations among “Course” and “Topic” nodes in the KG. The hierarchical nature of the learning content also results in graph paths, flowing from the Journey nodes to the OER ones, through Course and Topic nodes. Since the two lower levels of the hierarchy are required, by design, for learning a topic in a course, they do not enable thus personalizing a learning trajectory. Therefore, the search for graph paths is limited in this thesis to the top three levels of the taxonomy: “Journey”, “Course”, and “Topic”. Journey nodes in this approach are the representation of a learning goal, which is achieved through a certain learning path. Recommended courses and topics in the learning path are then provided to the learner along with all the OERs they include. This approach is meant to ensure that a learner can receive all needed learning materials to master a certain topic.

Definition 1: Recommended learning path (LP) is defined as the ordered subset of courses $c_i \in C$, and topics $t_i \in T$, where C, T are the complete sets of courses and topics in the KG, respectively. A learning path LP always begins with a Course or a Topic node type, hereafter referred to as the landing node; it ends at a Journey node that represents the targeted learning goal.

Definition 2: A landing node in the KG is defined as the node $l_i \in (C \cup T)$ that has a direct connection, through semantic similarity or otherwise, to a node $c_{context} \in C_{user}$ or $t_{context} \in T_{user}$ which exists in the context vector of a user.

Landing nodes play an important role in connecting a User node to a Journey node in the KG, see Figure 27.

courses, which was calculated from the average time needed to finish all OERs in 20 courses and the average learning time available for the target users in an elderly home.

To implement this approach for controlled path finding in the KG, Dijkstra's algorithm (Dijkstra, 1959) for finding the shortest path in the graph is utilized. With a controlled minimum and maximum length of the potential shortest path, Dijkstra's algorithm is utilized to extract a subset of graph path, which will be then used by the path-ranking algorithm as a navigation environment.

The goal of finding the shortest paths in this approach is not to find the fastest ways to achieve the learning goal, but rather to accomplish two main tasks:

1. To ensure that the path in the learning environment of MDP is one that connects a learner to their learning goal.
2. To select from the graph only paths that comply with the practical requirements of the real-world learning setting, such as the time limitation.

These two conditions contribute to the effective construction of the learning environment for the MDP-based path ranking by reducing the computational complexity and load that long, overlapping, and non-realistic learning paths impose on the algorithm. Listing 1 represents the process for identifying the set of learning paths in the KG, building on the customized shortest path algorithm.

Listing 1. Graph path-finding algorithm

Algorithm 1: Learning Path Identification Algorithm
Input: userProfile (learningGoal, contextVector), knowledgeGraph (KG)
Output: PossibleLearningPaths
<i>landingNodes</i> $\leftarrow$ EmptyList();
<i>goalNode</i> $\leftarrow$ LocateNodeInKG(userProfile.learningGoal, KG);
foreach <i>node</i> in <i>userProfile.contextVector</i> do
if <i>NodeExistsInKG</i> (<i>node</i> , KG) then
Add <i>node</i> to <i>landingNodes</i> ;
else
<i>similarNode</i> , <i>similarityScore</i> $\leftarrow$ FindSemanticallySimilarNode(<i>node</i> , KG);
if <i>similarityScore</i> $\geq$ <i>similarityThreshold</i> then
Add <i>similarNode</i> to <i>landingNodes</i> ;
end
end
end
<i>PossibleLearningPaths</i> $\leftarrow$ EmptyList();
foreach <i>landingNode</i> in <i>landingNodes</i> do
<i>path</i> $\leftarrow$ FindShortestPath(<i>landingNode</i> , <i>goalNode</i> , <i>maxPathLength</i>);
if <i>path</i> is not Null AND <i>Length</i> (<i>path</i>) $\leq$ <i>maxPathLength</i> then
Add <i>path</i> to <i>PossibleLearningPaths</i> ;
end
end
return <i>PossibleLearningPaths</i> ;

5.3 Graph path ranking

The second problem that the recommendation system addresses is ranking all potential learning paths that may lead a learner to their learning goal, in order to select the most suitable path for them individually based on their learning context. Here, the highest ranking path is defined as the optimal path, since the algorithm increases path ranking based on its alignment with the user's learning goal and context.

To achieve the goal of ranking graph paths based on the learning context, MDP is utilized to guide the navigation strategy over the KG and assign progressive weights to each learning path as the algorithm examines each transition between path states. Practically, a virtual agent is defined to perform those transitions. The agent is assigned the task of navigating all possible learning paths, which were identified by the path-finding algorithm. The agent then assigns scores to path states and actions between those states, according to a contextualized reward and penalty function. After exploring the complete set of possible learning paths, the agent will have ranked all possible learning transitions for an individual learner, resulting in the identification of the most suitable learning path for them. In the following sections, the operational details of implementing the MDP-based recommendation algorithm in the CARExKG method is described.

5.3.1 The knowledge graph as the MDP learning environment

As discussed in section 2.6, the implementation of MDP requires defining the environment in which the process is implemented. The implementation of MDP in this research utilizes the KG as the direct environment for implementing MDP and reasoning through it. As a network of interconnected users and learning materials, the interaction between a learner and the learning materials is represented through graph relations that connect those materials to the user's profile. A user profile is designed to include the materials that a user has finished learning, in the form of previous knowledge that the user has. In this sense, the learning path recommendation becomes the process of selecting a n ordered sequence of learning materials from the KG, which is meant to be added to the user's previous knowledge through finishing the learning path.

The advantage of this approach is that no intermediate environment needs to be defined for the MPD reasoning. Using the KG directly as the MDP environment reduces the complexity of the recommendation algorithm, which not only enhances its computational efficiency, but most importantly improves its transparency and explainability. This, in turn, is one of the overall goals of this thesis, due to its importance in facilitating a sound learning process.

To define the MDP environment, environment states, actions, and a transition function are to be specified. Then, the implementation of the process follows a state-action value update based on Q-learning.

MDP States

The states of the MDP environment are defined in this system as the KG nodes, which include all taxonomy levels from Journey to OED nodes. MDP states are acquired from both layers of the KG, the static and the dynamic ones. This means that the courses and topics that a user learned previously are also included in the environment, where they are represented as a subset of the complete course set C and the complete topics T , see definition 2.

MDP Actions

An action in this algorithm is defined as the transition between two learning states, if and only if:

1. they belong to set C or set T , and
2. an edge exists between the two states.

Graph edges stem from the taxonomy defined for the KG, which include, in case of the VET KG, the relations: `has_course`, `has_topic`, and `has_semantic_similarity_to`.

The navigation in the KG is limited to Course and Topic nodes in the KG due to the construction process of the KG itself, in which the child nodes of a topic are all required to master that topic. This leaves no need to personalize on the OER level, but rather on the Topic and Course levels.

MDP Transitions

Defining the transition probability in this approach is a part of the algorithmic contextualization procedure that the thesis proposes. Here, the transition probability $\mathcal{T}(s'|s, a)$ to state s' from state s via action a , is defined as a function of two factors:

1. The exploration preference of the learner: where learners who prefer to explore more content during learning are more likely to deviate from the recommended path and try new content, see Figure 28. Therefore, the probability of taking an alternative action a_{alt} at a state s , which leads to an alternative state s'_{alt} that does not belong to the learning path, becomes higher.
2. The existence of both states s and s' on one of the learning paths identified by the path finding algorithm. In this case, the probability of transitioning to state s' from state s is assigned a higher value, due to the structural relevance between the two states in the KG, which reflects their strong connection in terms of leading to the same learning goal.

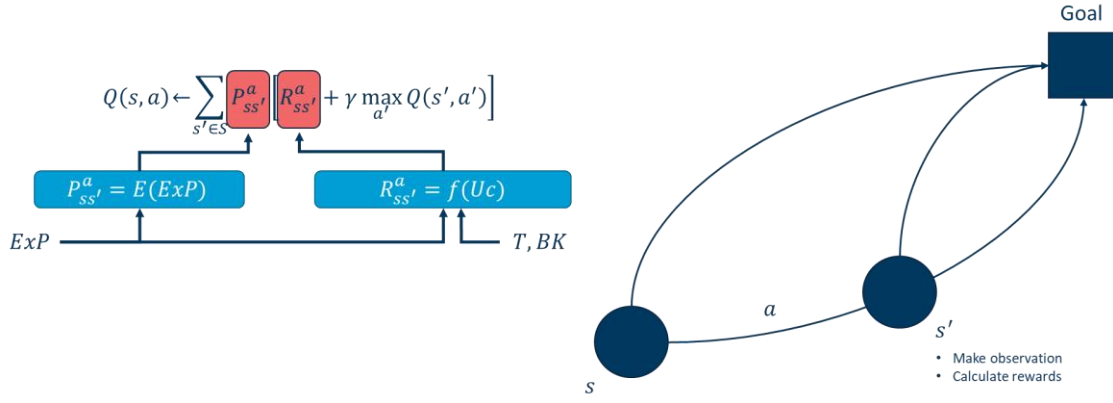


Figure 28. Integration of contextual factors in the state-action value function

5.3.2 Implementing State-Action Value Updates

To update state-action values $Q(s, a)$, the reinforcement learning technique Q-learning is utilized in this research. The selection of Q-learning is based on three main reasons:

- Its computational efficiency.
- Its sampling efficiency, for cases where KGs are relatively small.
- Its interpretability in comparison to policy-update methods.

The update of state-action values is based on Bellman's equation, Eq. 15, which updates $Q(s, a)$ as:

$$Q(s, a) = Q(s, a) + \alpha \left[r + \gamma \max_{a'} Q(s', a') - Q(s, a) \right] \quad (15)$$

Where:

- r : Reward for taking action a .
- α : Learning rate.
- γ : Discount factor.
- s' : Next state after taking action a .

Design of the reward for contextual knowledge graph navigation

The second part of Bellman's equation represents the reward (or penalty) that a state-action pair gains after a certain transition. For example, if state s' is the final state, here the learning goal, then the reward is the maximum that a transition can reach. However, if the state s' is further from the learning goal than state s , the reward is dependent on the exploration preference of the learner. In case the learner did not wish to explore the learning content in the graph, the reward may become zero, or negative representing a penalty. On the other hand, taking a longer path might be a result of the

previous knowledge of the learner, meaning that the state-action value in this case should increase, as it corresponds to the learner’s knowledge. This interplay between different learning-context factors requires, therefore, a sound integration of these factors in the reward function, to account for their levels of importance, and the different influences they have on the learning path recommendation.

To address this point, this thesis proposes an approach for designing weighted, contextualized reward functions, which aim to fulfill three main criteria:

1. Allow the representation of a variable number of context factors in one function.
2. Allow a clear differentiation between the importance of those factors in different learning settings.
3. Enable a flexible representation of each factor, defined by the domain and use case requirements.

Eq. 16 represents the reward function as a sum of all weighted contextual factors in a learning setting.

$$R_{ss'}^a = \sum_{i=1}^n W_x * F_x = W_{BK} * F_{BK} + W_{Exp} * F_{Exp} + W_T * F_T \quad (16)$$

This equation demonstrates the construction of the reward function using three sample contextual factors:

- Background knowledge
- Exploration preference
- Time available for learning

While these factors are used in the CARExKG method implemented in the VET use case in this research, the extension of $R_{ss'}^a$ is inherently possible due to the nature of the weighted polynomial. For the rest of this section on the design and implementation of the MDP-based reasoning, these three main context factors are used, since they were the ones selected by VET domain experts as the most important factors in that use case.

Definition 3: Background knowledge BK is the cumulative amount of learning materials, represented by graph classes “course” and “topic”, which the learner has finished learning on the learning platform.

Definition 4: Time available for learning T is the time the learner can dedicate to following the learning path, described as the number of weeks, and the number of hours available for the learning in each week.

Definition 5: Learning exploration preference Exp is the tendency of the learner to deviate from the recommended learning path and explore new learning content in the knowledge graph

In the knowledge graph, BK as a factor is determined by the courses in the user context, which are identical or semantically similar to the courses in the KG. Therefore, it is represented as in Eq. 17.

$$F_{BK} = \left[\frac{\|\sum_i Sim(BK_i, txt(s'))\|}{count(BK)} \right] \quad (17)$$

The Exploration preference factor is directly related to the distance from the learning goal, and thus represented as in Eq. 18.

$$F_{Exp} = \left[Exp * DfG + \frac{1 - Exp}{DfG} \right] \quad (18)$$

Time available for learning is related to the average time of courses selected in the learning path, and thus described as in Eq. 19.

$$F_T = \left[\frac{1}{\|1 - (t(s') - Avg(T_p))\|} \right] \quad (19)$$

Substituting Eq. 17, 18, and 19 in Eq. 16 gives the total reward function for the use case, as shown in Eq. 20.

$$R_{ss'}^a = W_{BK} * \left[\frac{\|\sum_i Sim(BK_i, txt(s'))\|}{count(BK)} \right] + W_{Exp} * \left[Exp * DfG + \frac{1 - Exp}{DfG} \right] + W_T * \left[\frac{1}{\|1 - (t(s') - Avg(T_p))\|} \right] \quad (20)$$

5.3.3 Algorithm Overview

The implementation of the recommendation algorithm builds on the defined environment, and utilizes the contextualized reward function to rank all paths identified in the path finding step. The following steps are conducted to perform the ranking task:

1. Initialization: Initialize $Q(s, a)$ for all states and actions. Start at a source node in the KG.
2. Action selection: Use an exploration strategy (e.g., epsilon-greedy) to select actions based on $Q(s, a)$.
3. Transition and reward: Execute the selected action and observe the next state and reward.
4. Update state-action values: Apply the Bellman equation to update $Q(s, a)$.
5. Repeat from step 2: Continue until convergence or task completion.

The recommendation algorithm is described in Listing 2.

Listing 2. MDP-based recommendation algorithm

Algorithm 2: MDP-Based Learning Path Selection Algorithm

Input: knowledgeGraph (nodes, relations, properties), userProfile,
possibleLearningPaths
Output: optimalLearningPath
optimalLearningPath $\leftarrow$ Null;
highestAvgValue $\leftarrow -\infty$;
foreach *path* **in** *possibleLearningPaths* **do**
 stateActionValues $\leftarrow$ EmptyList();
 finalState $\leftarrow$ LastState(path);
 reward(finalState) $\leftarrow$ 100;
 foreach *state* **in** *reverse path* **do**
 action $\leftarrow$ ActionToPreviousState(state, path);
 value $\leftarrow$ ContextualizedRewardFunction(state, action, userProfile,
knowledgeGraph);
 Add value to stateActionValues;
 end
 avgValue $\leftarrow$ Average(stateActionValues);
 if *avgValue* > *highestAvgValue* **then**
 highestAvgValue $\leftarrow$ avgValue;
 optimalLearningPath $\leftarrow$ path;
 end
end
return optimalLearningPath;

5.4 Discussion

In reflection on the second RQ, the implementation of the MDP-based recommendation algorithm points out that solid and transparent reasoning can be utilized to generate personalized learning paths if the reasoning algorithm is designed contextually and when the reasoning environment is well established based on contextual and pedagogical foundations.

Graph-based MDP inherits the strength of reinforcement learning through the iterative exploration and ranking steps, yet it still offers an explainable reasoning, since its steps can be traced back to specific logic for ranking the paths in the KG. Moreover, its design, which is based on learning context factors, governing the transition between the states and the ranking of each state-action pair, provides a user-centric and pedagogically-oriented realization of the path selection for the recommendations step.

In addition to a contextually designed KG, as shown in chapter 4, the transparent nature of the KG supports the transparency of the MDP algorithm by providing additional information about the learning content in the recommended path, and the connections between the elements of the path, among other structural information that is directly extracted from the KG.

It is worth noting here that the proposed implementation of MDP in this thesis accomplishes an additional practical value-added, which is the reduction of computational

resources needed to run a full navigation phase by the virtual agent. Dividing the implementation of MDP into path-finding and path-ranking phases allowed for more control over the definition and construction of the MDP environment. This, in turn, enabled the reduction of the size of the KG-path set, which would have been normally the learning environment for the agent. By applying pedagogical and domain-specific rules to control the length of the learning path, the number of navigation iterations is considerably reduced to a more meaningful and less computationally demanding one. While this was not the priority of the MDP implementation, it was noted during the experiments in this research that the selection of path-length limits, based on a realistic learning journey for the learner, had supported an faster execution of the recommendation algorithm and thus allowed providing the recommendations to the user without long waiting times after requesting the recommendation.

Part III

Explainable AI for Learner Agency over Learning Recommendations

6 A Framework for Explainable Recommendation Systems over Knowledge Graphs

6.1 Introduction

While reasoning over knowledge graphs provides the means for generating an accurate prediction, the user's adoption of this prediction is not only related to its accuracy but also influenced by their ability to understand how suitable that prediction is for their needs. In an educational setting, adopting a learning path recommendation is a decision that results into a commitment to dedicate the time and effort for learning. Depending on the length of the learning path, this commitment may last hours, weeks, or months. Therefore, enabling the learner of making an informed decision on the learning recommendation is crucial to allow long-term adoption, acceptance, and trust in the recommendation algorithm. This concept extends to other educational stakeholders, including the educators, education management, and parents. The educational stakeholder's understanding of the recommendation logic, and the underlying methods used to generate it, enhances their awareness of its capabilities and limitations, which in turn translates to their ability to use the recommendation effectively and safely.

Explaining how a recommendation algorithm works is, however, not always a straightforward procedure. This is because a wide range of intelligent algorithms, such as those using machine learning (ML) and deep learning (DL), often function as black-box models, with opaque reasoning processes underlying their predictions. This opacity can negatively impact user understanding of the recommendation, and thus their trust and acceptance of the retrieved information (Polley et al., 2021). This is particularly problematic in critical domains, such as education, where transparent reasoning is a requirement, not only a feature, for incorporating predictions into the decision-making processes. To address this challenge, this chapter introduces a framework for generating explainable recommendations over the knowledge graph structure. The chapter contributes to the answer of the third research question of this thesis, as it proposes a graph-based foundation for integrating open-box data structures with black-box algorithm in the explainable recommendation task, in a way that allows deducing the reasoning that led to generating a certain recommendation for an individual user.

The key concept underpinning this framework is the use of the semantic knowledge graph as a single source of truth (SSoT). The KG allows the fusion of multiple information sources within the graph, which can then be queried by intelligent RS and explainability algorithms to provide users with comprehensible recommendations. Moreover, the framework is designed as a foundation for developing domain-flexible explainability of CARS. The term domain-flexible here refers to the framework's capacity to be domain-agnostic through transferability across multiple domains, while retaining the ability to represent each domain's specific requirements within the structure of the KG, and thus the explainable CARS algorithm.

The selection of KGs as the basis for the framework is based on several advantages:

1. Knowledge graphs excel at contextualizing entities through semantic relations to other entities within the graph. This contextual richness enhances both the reliability of the CARS algorithm and the meaningfulness of generated explanations.
2. Unlike relational databases, knowledge graphs enable direct implementation of graph theory methods for querying of KG elements. This structure facilitates the use of basic statistics, such as element degrees of connectivity, as well as other complex centrality measures (CMs) for potential future algorithmic extensions. These capabilities enhance the relevance of retrieved results and generated recommendations based on user queries by considering more nuanced measures of connectivity.
3. Knowledge graphs provide paths between entity pairs, consisting of intermediate nodes and relations. By utilizing the concept of graph walks, i.e., navigating these paths according to predefined rules, the intelligent algorithm can connect more relevant content to user queries. As these graph paths reflect a part of the reasoning behind entity connections through multiple graph-hops, they can be used for explaining the retrieval of one graph entity based on the retrieval of other transitively connected ones.

Since explainability is a domain- and context-dependent task the explainability framework developed and used in this thesis offers the required methods for implementing domain and expert requirements within explainable CARS algorithm. This enables cross-context implementation through the explicit integration of requirements engineering concepts into the KG model creation process. In this sense, the KG serves as a unifying structure, linking domain requirements, explanations, and contextual data. During KG construction, see Figure 28, information from domain requirements, expert knowledge, and textual data sources is modeled through graph nodes and relations. The resulting structure is then used to generate explanations for the CARS algorithm recommendations.

The knowledge embedded in KG relations addresses the challenge of explaining black-box IR algorithms (Abu-Rasheed, Weber, Zenkert, Dornhöfer, et al., 2022; Y. Li et al., 2022; Tiddi et al., 2020) by providing sufficient information to generate

understandable explanations that reflect the reasoning behind model's predictions. KG-based explanations can be constructed visually or verbally, enhancing their comprehensibility for end-users.

To achieve framework transferability, the knowledge graph construction process is designed to capture high-level domain requirements from both domain experts and data sources. Once defined, these requirements are uniformly embedded in the KG and subsequently retrieved using the same CARS algorithm, as the algorithm operates over the KG structure rather than raw data sources.

6.2 TExKG: a transferable, graph-based, explainability framework for RSs

To address the challenges of explaining intelligent RS algorithms while considering domain requirements, a structure composed of static and dynamic components is developed, see Figure 29. Static and dynamic components correspond to both domain-specific and domain-agnostic elements of a use case, respectively. Domain-specific components represent the domain's data, requirements, and expert-defined rules. Domain-agnostic components serve two primary functions within the framework:

1. They link domain-specific components to the knowledge graph, enabling a transferable graph construction process replicable across domains.
2. They perform explanation generation, information retrieval, or recommendation tasks based on the KG structure.

Domain-agnostic components integrate the intelligence of RS algorithms without compromising their explainability, due to their direct connection to the transparent knowledge graph representation. The resulting framework, which links all components through the KG, demonstrates high flexibility in addressing the needs of multiple domains. Moreover, it enables flexible adaptation to various AI algorithms used in information retrieval or recommendation tasks, as they depend on the data provided by the KG itself.

The knowledge graph forms the core of the framework, enabling the extraction and fusion of information from multiple sources, such as databases, domain experts, and requirement lists, into a queryable knowledge structure through defined nodes and relations. Once constructed, the graph serves as a graphical knowledge base, providing RS and explainability algorithms with the necessary information to generate explainable predictions for users. Since explanations are derived from the knowledge graph, the framework can generate high-level explanations even for black-box RS algorithms.

The explanation format is dependent of the available information that can be retrieved from the algorithms and the knowledge structure. Therefore, the KG allows generating both visual and textual explanations for a learning recommendation. Textual

explanations can be reconstructed from the data in graph nodes, relations, and properties, using different techniques, such as explanation templates (Abu-Rasheed et al., 2024). Visual explanations, on the other hand, can utilize the graph structure itself to generate network visualizations of the recommended learning path in the KG, or can be generated independently in other formats, such as Pie charts, radar charts, etc., based on the information retrieved from the KG.

To accommodate any requirements from the user or the domain expert, e.g., regarding the explanation shape, content, structure, level of detail, etc., a dedicated components is utilized to collect and represent those requirements in the form of a list of rules that govern all aspects of the explanation. This list is then accessed by the explanation algorithm to adjust the explanation to the needed from the user and the domain expert.

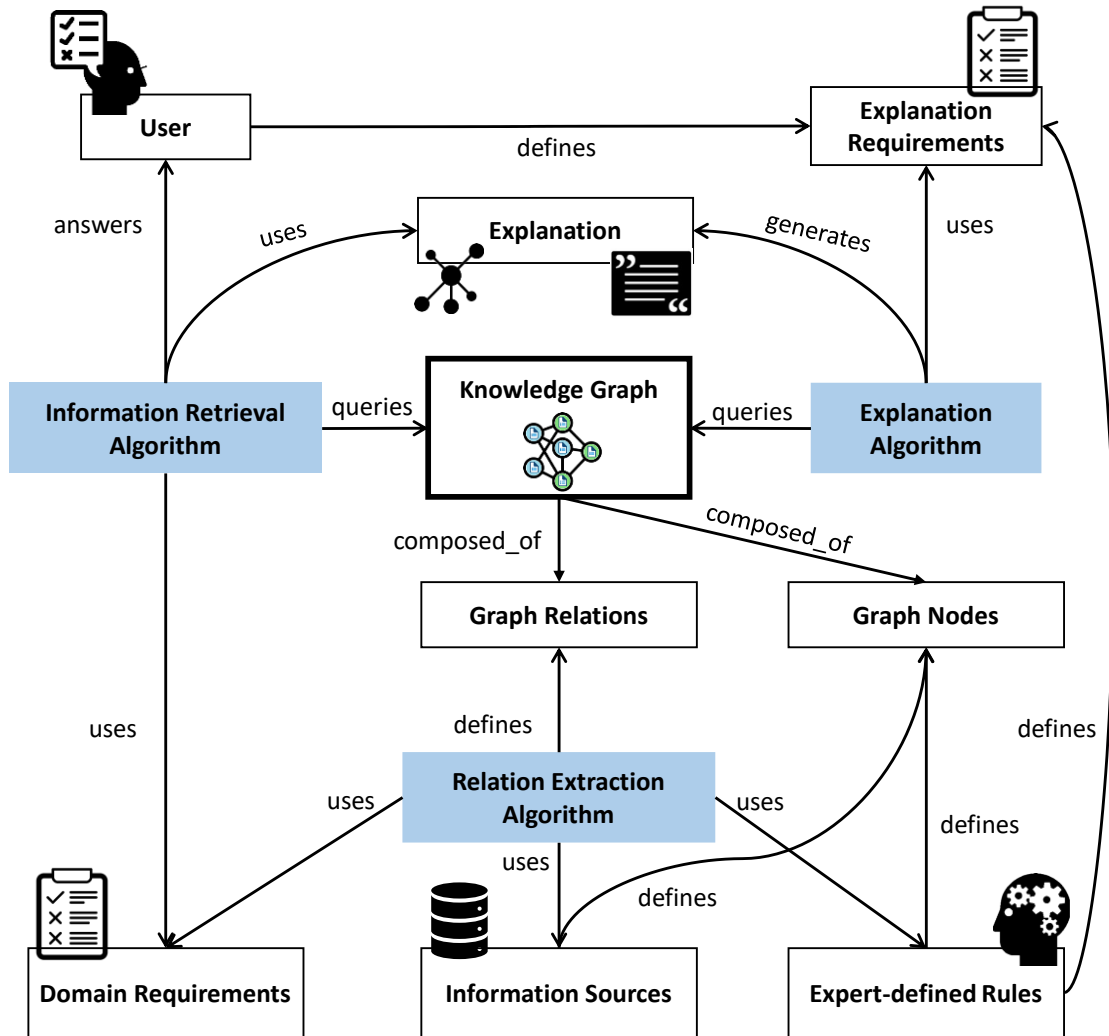


Figure 29. Proposed framework for transferable, domain-oriented, explainable RS, based on knowledge graphs. Domain-specific elements are shown in white, while domain-agnostic algorithms are shown in blue (Abu-Rasheed, Weber, Zenkert, Dornhöfer, et al., 2022).

6.2.1 Knowledge graph construction

The construction of the knowledge graph is accomplished by defining its nodes and relations. Knowledge graphs excel at representing multiple process dimensions through:

1. Defining specific node types within the graph structure (Abu Rasheed et al., 2021).
2. Expanding relation definitions to connect nodes of similar or different types while establishing varying relevance levels between nodes (Shalaby et al., 2017).

In this framework, graph nodes are defined based on two foundational pillars: 1) information sources, and 2) expert-defined rules. Each pillar contributes to determining the node's type, content, and attributes.

Graph-nodes definition

In educational environments, information typically spans multiple data sources, which include lecture slides, reference books, lecture recordings, exercise sheets, as well as additional online resources. The framework enables several information sources to define different node types within the graph, with each type representing an individual information source, or a part of the information it includes, while graph relations establish the connections between those sources. This approach enables nodes to represent the characteristics of domain-specific information while maintaining integration across information sources within a unified knowledge structure, namely the KG.

Expert-defined rules are employed to reflect domain expert knowledge within the graph construction process. This step is essential in sensitive domains, such as education, since the inherent knowledge that experts have, through their long experience, is not explicitly written in the education materials, but implicitly communicated to their students through teaching. A form of representing this implicit knowledge through rules is adopted in this framework. These rules follow an "IF...THEN..." structure, supporting decision-making regarding node types, content, graph relations, as well as the explanations. Reasoning with these rules may employ either forward-chaining or backward-chaining approaches as defined by (Cheng et al., 2018; Yu et al., 2021; Zhang & Li, 2019).

Various methods can extract knowledge from domain experts, including questionnaires, interviews, and formal reports. The experience collected throughout this thesis during the implementation of this concept suggests that interviews with domain experts effectively acquire knowledge that can be directly translated into system rules. For example, interviews with career coaches have shown that the organizational role of the staff member is one of the factors used to define the length of the training they should follow, i.e., the learning path recommendation in the case of this research. Backward chaining inference subsequently allows the deduction of graph node types and content to represent information sources appropriately. Figure 30 illustrates an example of integrating expert-defined rules and domain requirements in the creation of graph nodes

and relations based on the documentation style of engineering staff in chip-design-related learning setting.

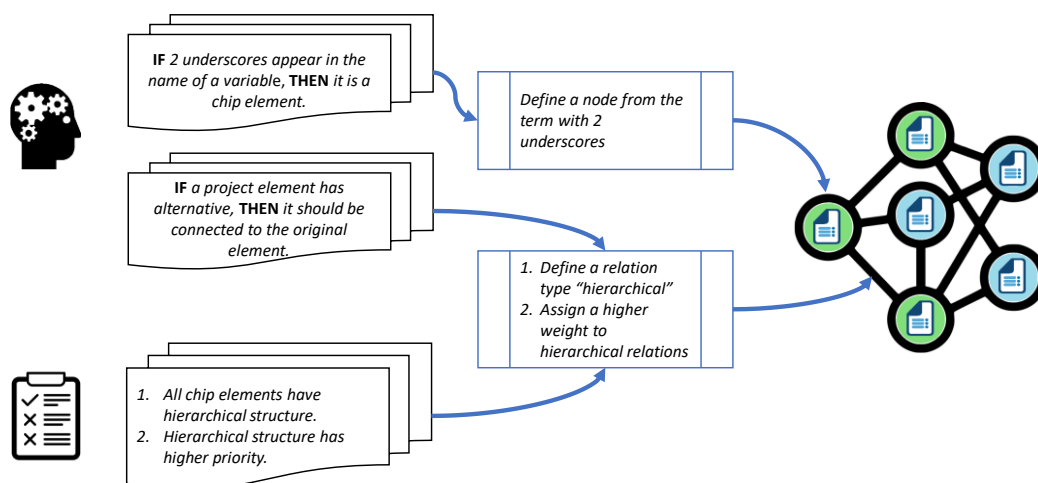


Figure 30. Example of integrating expert defined rules and domain requirements in the creation of graph nodes and relations (Abu-Rasheed, Weber, Zenkert, Dornhöfer, et al., 2022).

Graph-relations extraction

Graph relation definition employs both expert rules and data sources while incorporating domain requirements at the same time to frame the variety and goal of KG relations. Independent domain-requirements component is used in the framework to represent domain-specific needs, which are not directly expressed in the information sources or through expert-defined rules. For example, terminology variations across domains may include entities with different contextual meanings that are not explicitly articulated. A term such as “user autonomy” is defined slightly differently between a technical domain, a psychological one, and a medical one. This definition is essential in its own context to generate meaningful learning recommendation, as the term is interpreted by NLP algorithms and text mining pipelines. These domain- and context-related differences are modeled as domain requirements within the framework, which are made visible to any recommendation and explainability algorithm through their integration in the KG structure.

Domain requirements may take various formats, including lists, dictionaries, or statements. They are utilized within relation extraction and information retrieval processes. The relation extraction (RE) algorithm identifies connections between nodes by analyzing node types and content reflecting information source characteristics. For document nodes containing textual content, relation extraction may calculate textual similarities between node pairs, establishing relations when similarity scores exceed predefined thresholds.

Experts may also define relations directly, with the relation extraction algorithm translating expert rules into connections between corresponding graph nodes. Through information sources, expert rules, and domain requirements, the relation extraction

algorithm incorporates domain-specific characteristics while maintaining domain-flexibility, allowing its application across multiple domains when their requirements are understood.

6.2.2 Graph-based information retrieval for the RS

Similar to the relation extraction component, the recommendation system module is designed to be domain-agnostic, enabling its reuse across domains while preserving domain-specificity considerations. Domain-specific information is made available to the RS algorithm through the knowledge graph. The RS also has access to domain requirements, which can play a role in adjusting the queries it makes to the KG. An RS system predicts a learning path for a certain user, based on the KG. Then, it obtains the path explanations from the explainability algorithm.

6.2.3 Graph-based explainability

To ensure explanation generation capability regardless of whether algorithms are open-box or black-box, the explainability algorithm operates independently, requiring only knowledge graph queries and expert-defined input to generate the final explanations.

Explainability algorithms vary depending on the learning context, and domain- and expert-requirements. Textual explanations can then be generated in the form of semi-structured templates, which include predefined information slots that are complete with information from the KG and other controlled sources, such as large language models (LLMs). The graphical knowledge representation in the KG provides visual explanation capabilities, illustrating node relevancies, graph communities, and learning path surrounding. Various visual attributes, such as color and size, enhance the interpretation and understanding of retrieved information. Figure 31 demonstrates an example of generating textual and visual explanations from the knowledge graph and based on expert-controlled LLMs.

Explainability, however, is not only a matter of developing the infrastructure for the information retrieval and recommendation algorithms, but also a matter of developing the user interfaces and the interaction models between the users and the recommendation system. Therefore, a comprehensive learning and explainability interface is designed based on the recommendation and explainability concepts described in chapters 4 and 5, offering a detailed overview for the learner on the reasoning behind a learning path recommendation. In the following chapter, the interface design, its explainability components and sub-systems, as well as its role in supporting learner agency over the learning recommendation, are presented.

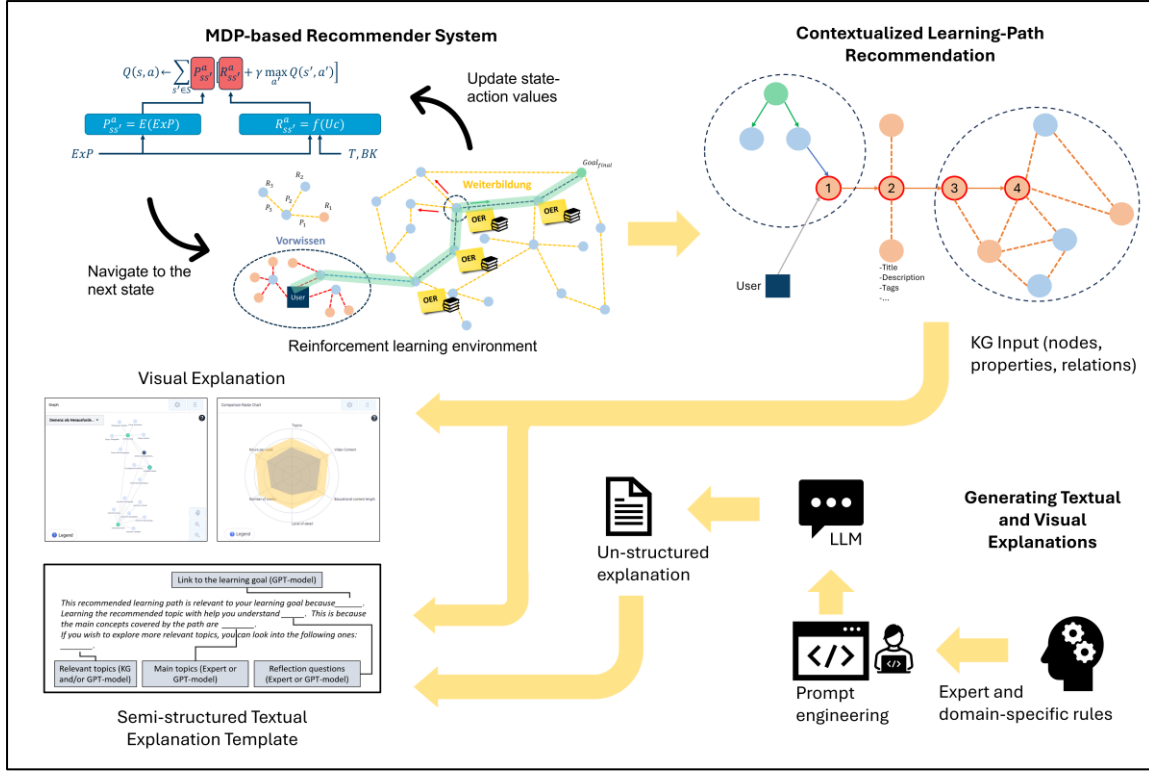


Figure 31. Visual and textual explanations generated from the information and structure of the knowledge graph.

6.2.4 Contextualized implementation of the framework

The framework has been implemented in multiple use cases, reflecting its flexibility and potential for adaptation to the requirements from the domain, experts, and target users. In the following, three variations of the framework are described.

Emergency Response Use Case

In the emergency response scenario, the TEExKG framework was applied to improve the speed and quality of on-site medical decision-making during rescue operations. The system is built around a knowledge graph encoding formal medical protocols (expert-authored action-sequence manuals and standard operating procedures) for first aid, which serves as the core knowledge base and is curated entirely under expert guidance, see Figure 32. Paramedics interact with this knowledge graph via a wearable device that delivers context-aware action recommendation, for example, suggesting the next treatment steps, providing relevant medical information, and indicating medication dosages, based on real-time situation recognition. Given the life-critical nature of emergency care, the explainability requirements demand absolute accuracy and clarity: the framework's explanations must have zero tolerance for error and are generated through rule-based interpretations of the situation-detection algorithm's output, drawing from multiple trusted sources in the knowledge graph. Domain experts were integrally involved at every stage of this use case: they fully supervised the construction of the

knowledge base (no automated knowledge extraction was allowed), defined all medical rules and relations, and retained final authority over the system’s recommendations, ensuring that the AI’s suggestions and explanations align strictly with established medical knowledge.

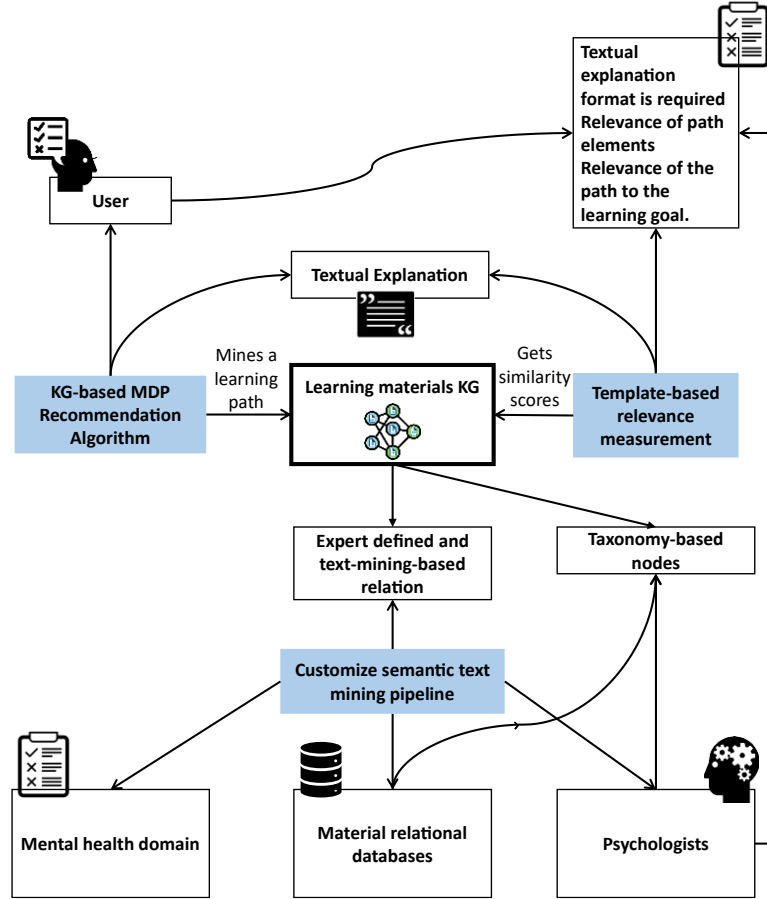


Figure 32. Framework implementation in the emergency response use case (Abu-Rasheed et al., 2024).

Elderly Care Use Case

In the elderly care training scenario, TExKG was implemented to support the on-the-job education of care home staff by providing intelligent learning recommendations with explainable guidance for managing complex daily care challenges. The knowledge graph in this case is constructed from a wide range of educational resources, such as caregiving handbooks, clinical guidelines, and training curricula, with semantic relations between topics identified via natural language processing and then validated by domain experts. Domain experts contribute extensively by establishing rules and constraints for graph construction and by adding expert annotations, thereby ensuring that the integrated content remains pedagogically sound and contextually appropriate. Using this expert-enriched knowledge graph, a graph-driven reinforcement learning algorithm produces personalized learning recommendations: some offer immediate, targeted solutions to urgent caregiving problems, while others suggest broader learning pathways for long-

term skill development. Correspondingly, the explanation component is designed to be sensitive and educationally oriented; it delivers semi-structured textual explanations that highlight the relevance and similarity of the recommended materials to the caregiver’s situation, while incorporating reflective questions crafted by pedagogical experts to stimulate the user’s understanding. This hybrid explanation approach (combining knowledge graph insights with expert-provided context) ensures a high level of clarity and minimizes the risk of misunderstandings in a domain where the well-being of vulnerable elderly patients is at stake. Figure 33 illustrates the implementation of the framework, taking into consideration those factors.

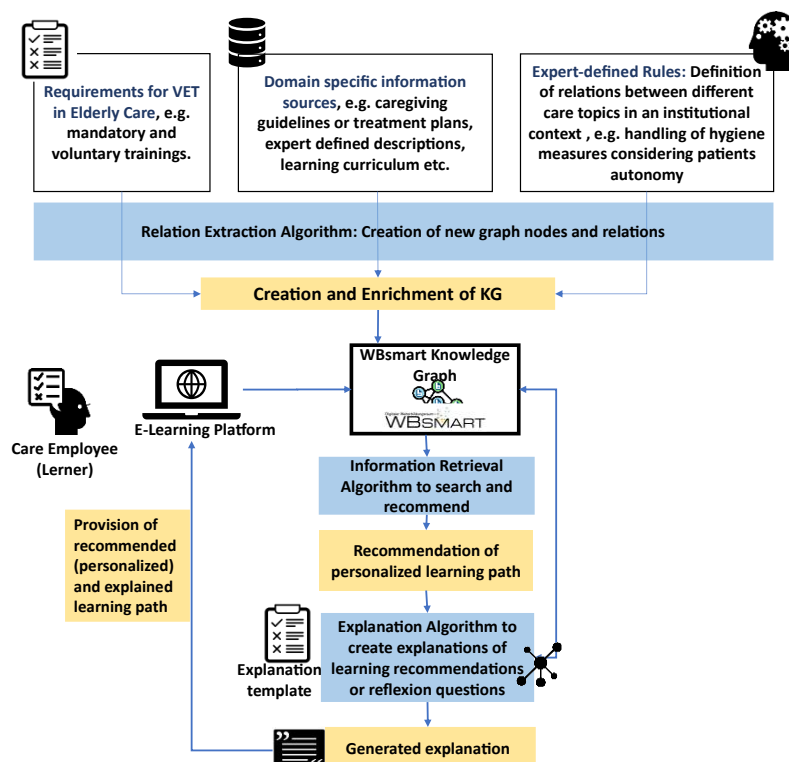


Figure 33. Framework implementation in the elderly care use case (Abu-Rasheed et al., 2024).

Mental Health Use Case

In the mental health support scenario, TExKG was deployed to aid early-career researchers in improving their well-being through guided self-learning, offering personalized content recommendations alongside supportive, mentorship-like explanations, see Figure 34. The framework’s knowledge graph was populated with expert-curated educational materials on mental health and well-being, organized according to a predefined taxonomy, and a Markov Decision Process-based algorithm was used to generate tailored learning paths through these resources. Given the high sensitivity of the mental health domain, domain experts imposed strict boundaries to ensure the system’s recommendations remained purely educational and never strayed into clinical diagnosis or therapy advice. Accordingly, the explainability module was heavily guided by expert-defined rules and templates: it uses an expert-crafted textual template to justify each recommendation by explaining how the suggested resources are semantically

related to each other and why they are relevant to the user's self-improvement goals. All explanatory content and terminology were co-designed with mental health professionals to remain non-therapeutic and supportive in tone, effectively allowing the system's explanations to serve as a surrogate for a human mentor's guidance while rigorously respecting the domain's clinical boundaries.

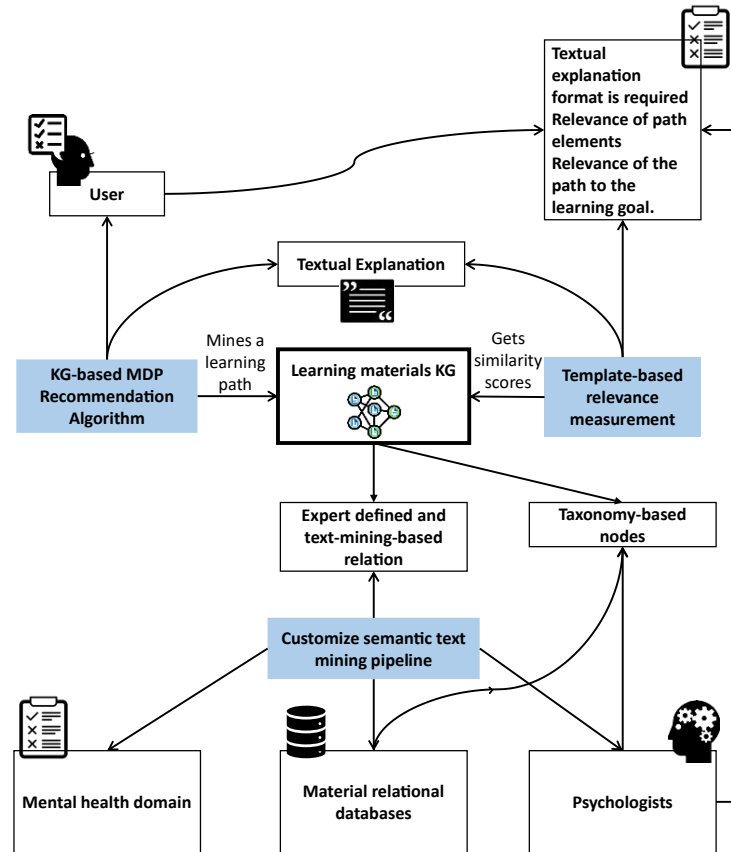


Figure 34. Framework implementation in the mental health use case (Abu-Rasheed et al., 2024).

6.3 Discussion

The design of the TExKG framework has been conducted to support CARExKG method with a flexible and comprehensive approach to enhance the explainability of intelligent RS algorithms. It is noticeable that the design is general in nature, i.e., it does not define specific components for storing the extracted information from data sources, or a certain shape of the rules extracted from domains of applications of domain experts. This aspect requires making additional decisions from the deployers of the framework, to make sound selections on that level of detail. This, however, is intended by design, to make the framework as flexible and domain-agnostic as possible, while preserving the option to include concrete domain-related information in the information sources or the algorithms.

While the framework shows a high level of flexibility in implementation, it is also important to address its limitation, which materializes in the form of rigidity towards

Discussion

replacing the KG with other types of data structures, such as relational databases. For a high value-added from the framework, the KG is required as a center of the implementation, which means that data points are expected to be representable by graph nodes, and connectable by graph relations. On the other hand, this dependence on graph structures itself offers considerable advantages, which led to considering the KG as the main type of data structures for the framework. In case other data structures are inevitable in a certain application, a part of the system needs to be added to the framework to convert that structure into a graph-compliant one, such as representing primary and foreign keys in a relational database with graph relations and properties.

7 Explainability for Learning: Supporting Learner's Agency with Interactive, Conversational, and KG-based Explainability

Explaining educational recommendations to learners showed promising potential to enhance the learner's acceptance of the learning recommendations (Ooge et al., 2022). Explaining the recommendation is not only meant to clarify why certain content is recommended but also to support the learner's ability to make an informed decision about following the recommendation. This ability is greatly influenced by the type and volume of information learners receive from the explanations. Multiple types of explanations have been investigated in recent years. Hybrid and multimodal explanation approaches were found to increase the learner's satisfaction (Pecune et al., 2019) and engagement with the system (Tsai & Brusilovsky, 2019). With the new capabilities of LLMs, an emerging form of conversational explanations holds the potential to offer a better understanding of the recommendation by engaging the learner in a question-and-answer session about the recommended content. While the use of chatbots in education is not new, the use of LLM-powered chatbots in generating learning explanations is still under investigation, due to the great limitations of LLMs in a sensitive field like education. This reveals an essential and urgent need for creating tools that support researchers and educators in experimenting and evaluating different types of explanation modalities and approaches, as well as the different hybridization approaches amongst them. Our main contribution in this research is developing such an experimental tool, in the form of a modular and interactive interface, which allows the delivery of multimodal and chatbot-supported explanations.

An interactive, experimental, web-based interface is proposed as a tool for delivering textual, visual, and chatbot-supported explanations of learning-recommendations. The proposed design is based on a pedagogical analysis of requirements expected from a learning-explanation. A focus is put on providing the means to explain:

- 1) the recommendation's connection to the learning goal,

Discussion

- 2) relations amongst the recommended topics,
- 3) connection to the user-profile, and
- 4) connection to the teacher's original structure of the materials.

Additionally, open-ended questions are supported through a chatbot. Our components, see Figure 35, are designed to enable experimenting with different combinations of explainability methods, by selecting which components are visible to the learner. The following explainability components were offered for experimentation: Textual explanations: which are provided in a variable-length text area. By incorporating Markdown, this component allows for versatile text presentations. Tag-based explanation: this is a more specific form of textual explanation but allows the user to click on the tags to reach further information using hyperlinks. Hierarchical structure: which allows a clear overview of a recommendation's hierarchy, such as the one created by a human educator. It offers expandable elements and clickable titles with hyperlinking.

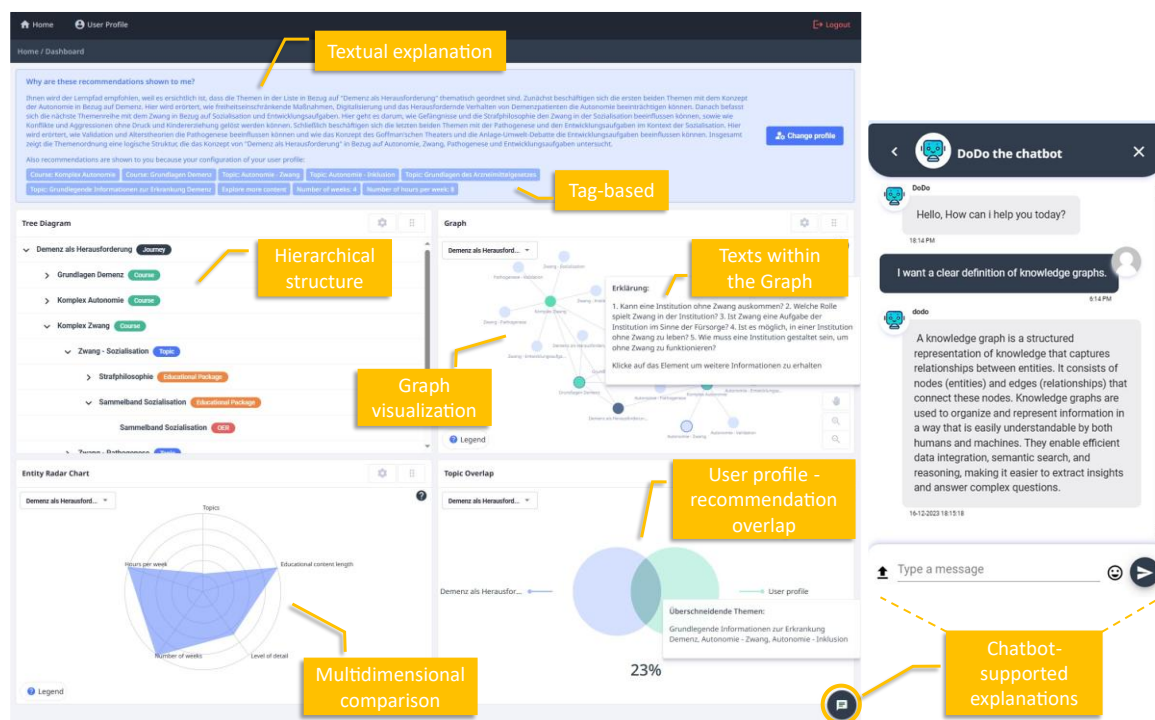


Figure 35. Proposed interface with the complete view of explainability components (Abu-Rasheed, Weber, et al., 2024a).

Graph-based visual explanations: It provides interactive graphs for visual explanations, which can also show textual elements on the graph. Radar charts: offering the potential to create multidimensional comparisons, to explain, e.g., coverage of the recommendations or overlap with preferences.

Venn diagrams: which provide a more focused overlap view, with a limited number of dimensions, but with interaction features including clicking, hover-over, and numeric or textual overlays.

Chatbot-based explainability: This component offers flexible support for chatbot use in explainability. The interface used LLMs powered by OpenAI API. This support of LLMs is also achieved by the design of chat messages to include contextual information and human-defined rules, offering more control over the LLM's output and thus reducing the risk of hallucinations and irrelevant responses. The chatbot element enables researchers and educators to utilize the information on other elements of the interface, to enrich the LLM's prompt and context. It also provides a connection to the database, to acquire additional information about the recommended items. A multi-agent chatting support is also offered, allowing the user to chat with other users, mentors, or even different LLMs. This provides high flexibility for experimenting and evaluating chatbot roles in recommendation explainability.

7.1 Conversational explainability

The extent to which students commit to a learning recommendation is closely tied to their comprehension of the rationale behind its suggestion, as well as their ability to adapt it based on that understanding. Among various explainability strategies, chatbots present an opportunity to engage students in dialogue akin to conversations with peers or mentors. However, despite advancements in generative artificial intelligence (GenAI) and large language models (LLMs), the capabilities of chatbots remain insufficient to fully substitute human mentors. Consequently, this study proposes a framework for leveraging chatbots as intermediaries in conversational interactions and as sources of controlled and limited explanation generation. This approach aims to capitalize on the strengths of LLMs while simultaneously mitigating their associated risks. The proposed chatbot, powered by LLMs, is designed to assist students in grasping the reasoning behind learning-path recommendations. A knowledge graph (KG) serves as a curated repository of information, enabling the regulation of the LLM's output by defining the context within its prompts. Additionally, a group chat mechanism is introduced to connect students with human mentors either upon request or when tasks surpass the chatbot's predefined scope. The chatbot's efficacy is assessed through a user study, which serves as a proof-of-concept and identifies potential requirements and limitations in employing chatbots for conversational explainability.

Providing students with continuous support, guidance, and feedback throughout their learning journey is a crucial concept rooted in social constructivism and Vygotsky's Zone of Proximal Development (Vygotsky, 1978), along with scaffolding theory (Wood et al., 1976). Nevertheless, one of the primary challenges in online education lies in ensuring access to experienced peers or mentors, especially given the abundance of online learning resources. This has led to increasing interest in research on recommendation systems (RS), explainability techniques, and automated feedback systems within the domain of technology-enhanced learning (TEL) and decision-making in educational contexts.

Evidence suggests that presenting learners with explanations and supplementary information regarding educational recommendations significantly enhances their acceptance of such suggestions (Ooge et al., 2022). Explainability in learning recommendations fulfills two key objectives: elucidating the basis for specific content suggestions and empowering learners to make informed decisions about whether to follow automated recommendations. The effectiveness of this decision-making process is heavily influenced by the type and quantity of information provided within these explanations. Recent studies have explored diverse forms of explanations, with increasing interest in utilizing LLMs for conversational explanations that involve multi-step dialogues to deepen learners' understanding of recommended content. Although chatbots have been employed in educational settings previously (Wollny et al., 2021), the application of LLM-powered chatbots for generating explanations related to learning remains underexplored due to significant limitations inherent to LLMs in educational contexts. Wollny et al.'s survey categorized chatbot tasks into assisting (20%) and mentoring (15%), with mentoring further divided into scaffolding, recommending, and informing methods. These classifications align with Yan et al.'s nine categories for LLM applications in education: profiling and labeling, detection, assessment and grading, teaching support, prediction, knowledge representation, feedback, content generation, and recommendation (Yan et al., 2023). Yan et al.'s study also highlighted practical and ethical challenges facing LLMs in education. They noted that most surveyed papers did not exceed Technology Readiness Level (TRL-2) and lacked transparency beyond Tier 1 according to Chaudhry et al.'s three-tier transparency framework (Chaudhry et al., 2022). This indicates that transparency was limited to researchers and practitioners without extending to learners or educators.

Concerns regarding the use of ChatGPT in education are echoed by Fullan et al. (2023), who identified drawbacks such as unoriginal responses that may lack depth or fail to inspire exploration or creativity due to their linearity. Despite these issues, Fullan et al. emphasized the potential benefits of such technologies in education while asserting that outright prohibition is not a viable solution. Consequently, active research efforts are underway to address inconsistencies, hallucinations, and biases associated with LLMs in educational applications. Strategies include fine-tuning models using labeled domain-specific datasets and contextualizing model prompts using existing data structures without requiring labeling. Sequeda et al. (2023) demonstrated that enriching LLM prompts with knowledge graph data significantly improves response accuracy. Building upon this concept and leveraging KGs' ability to model educational content effectively, this study seeks to enhance an LLM-based chatbot's capacity for generating relevant explanations related to learning recommendations.

7.2 Methodology for utilizing LLMs

To tackle identified challenges, a chatbot module integrated into a web application for learning-path recommendations was designed. This module leverages LLMs for

providing information on learning recommendations while offering students access to human mentors when needed. Four strategies were implemented:

1. **Task Scope Limitation:** The chatbot's responsibilities are confined to tasks less prone to errors or irrelevant outputs.
2. **Dialogue Design:** User intent is confirmed through re-prompting mechanisms before proceeding with responses.
3. **Contextual Prompt Enrichment:** KG-derived contextual data enriches prompts with user-specific information such as prior conversation states or descriptions of learning materials.
4. **Mentor Support Integration:** Students can access mentor assistance either directly or via fallback mechanisms when queries exceed chatbot capabilities.

The chatbot serves as the central component within the system architecture, Figure 36, acting as an intermediary between users, LLMs, and KG databases through its dialogue manager equipped with intent classifiers for predefined action categories.

Supported tasks include:

- Explaining recommendation rationales.
- Providing details about content on recommendation pages.
- Outlining benefits of specific learning materials.
- Highlighting relationships between recommended materials using KG data.
- Offering additional insights into recommended materials.
- Relating recommendations to users' personal contexts.

Queries outside these categories prompt rephrasing requests or mentor support suggestions via state-machine confirmation processes.

Mentor requests can be initiated at any time during chat sessions and are forwarded to available mentors registered within the system. Upon mentor acceptance, new sessions are created where mentors interact directly with students while utilizing chatbot assistance when necessary, Figures 37 and 38. Features such as image/PDF uploads enable multimodal interactions supported by advanced LLM capabilities like GPT-4.

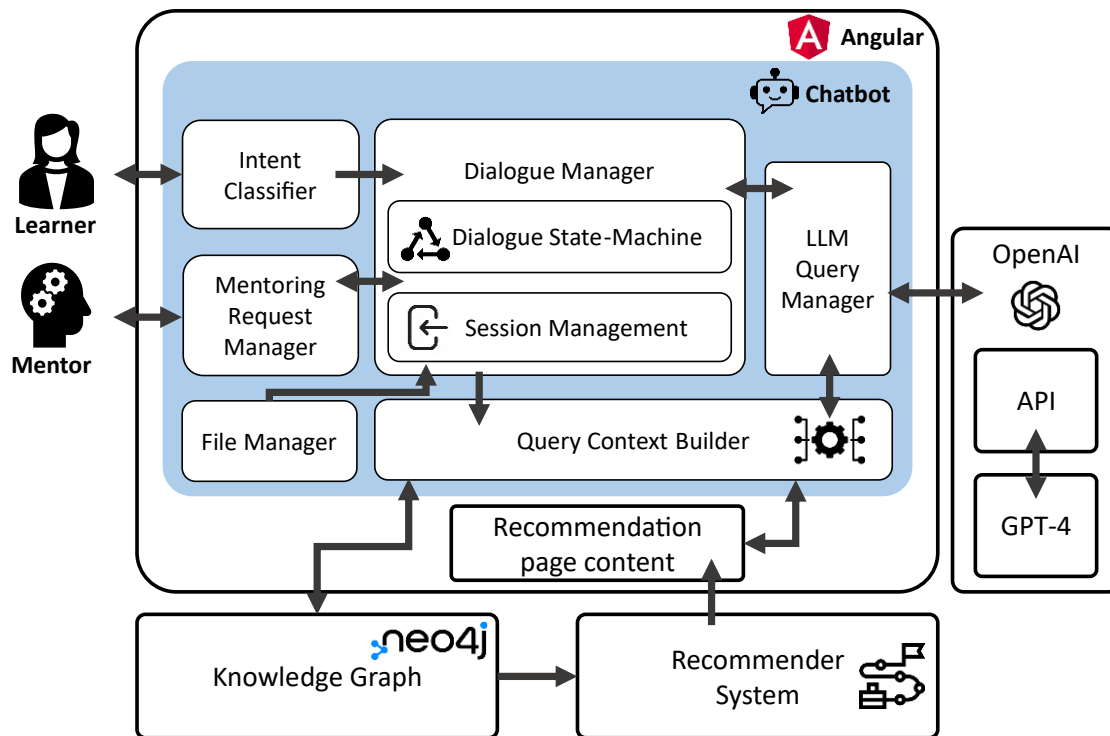


Figure 36: Chatbot system architecture within the Angular Web-App, and its connections to the KG, recommendation algorithm, and the LLM (Abu-Rasheed et al., 2024).

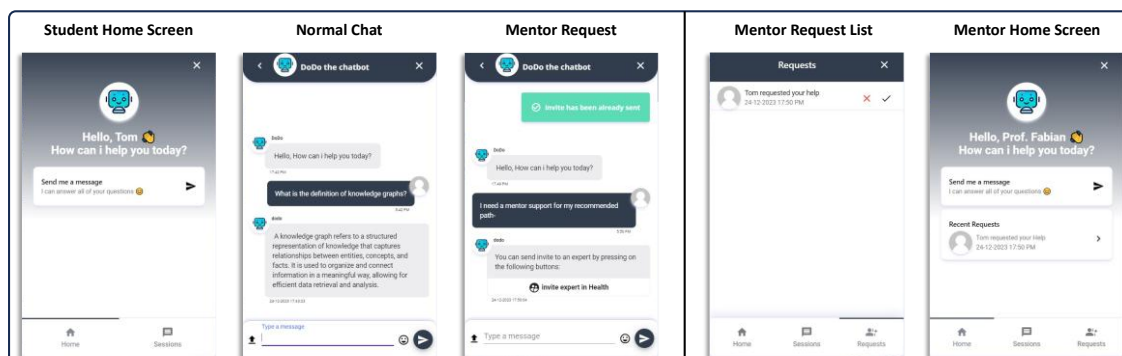


Figure 37: Chatbot interface for user-chatbot sessions and mentor requests (Abu-Rasheed et al., 2024).

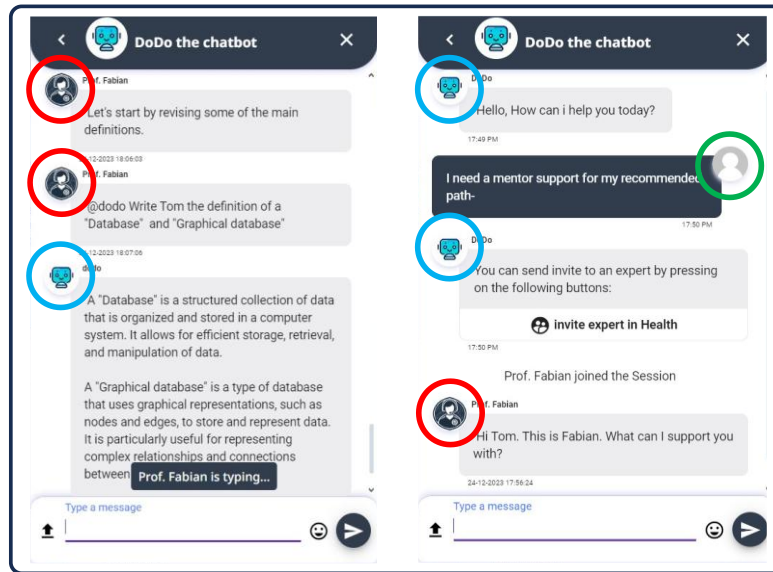


Figure 38: Group-chat session, including the user (green), mentor (red) and chatbot (blue) (Abu-Rasheed et al., 2024).

7.3 LLM prompt contextualization

Our third strategy for designing the chatbot is based on the concept of considering the context of a user query to generate a more relevant LLM response. While this concept has shown potential for influencing the output of the LLM, it is not indicated here that more context is necessarily better for the LLM’s output, because it is essential to select the type and volume of contextual information carefully for a positive influence on the output. For the conversational explainability task, the chatbot needs sufficient information about the learning path and its individual elements. It is also important to ensure that the LLM is not generating answers or explanations that a human educator will not accept or use with their students. Therefore, pedagogy experts and educators were enrolled in the design phase of prompt contexts, see Figure 39. Experts and educators were interviewed to extract sets of rules that the LLM should follow, such as the length of the response, the type of information to be provided (or not to be provided), and the definitions that the chatbot should consider when generating a response, such as the exact meanings of domain-specific terminology. The integration of domain experts in the design of our system is based on the explainability framework in (Abu-Rasheed, Weber, Zenkert, Dornhöfer, et al., 2022), which describes a set of roles of experts in the KG- and explanation-design processes.

In addition to experts, the other main source of context information in our system is the KG itself. The KG in our case represents a human-curated data structure, which provides comprehensive information on the learning materials and their relationships to each other. Relations among learning materials can form graph communities, in which a learning material is connected to others that appear in the same domain or application area (Abu-Rasheed, Dornhöfer, et al., 2023). For example, a course on “data analysis” can be created by a computer scientist and thus connected to other similar courses in the

domain of computer science, such as “data visualization”, “database management”, and “Python modules for data science”. The same course can be created by a health expert, and thus connected to another set of courses, such as “patient data privacy”, “digital health records” and “understanding X-ray scans”. Explaining the course “data analysis” is considerably dependent on the context in which it appears. In the KG, course relations can reveal that context, through the other courses connected to it. this potential is used to enrich the LLM prompt’s context with information about the course connections, and the potential KG communities to which it belongs. The KG as a data structure also provides information about the individual courses, the topics they are composed of, and the learning materials in each topic. This includes their metadata, similarity scores to other materials, which relation extraction (RE) algorithms calculated, and the hierarchical connections to the other taxonomical levels, which reflect their curricular format.

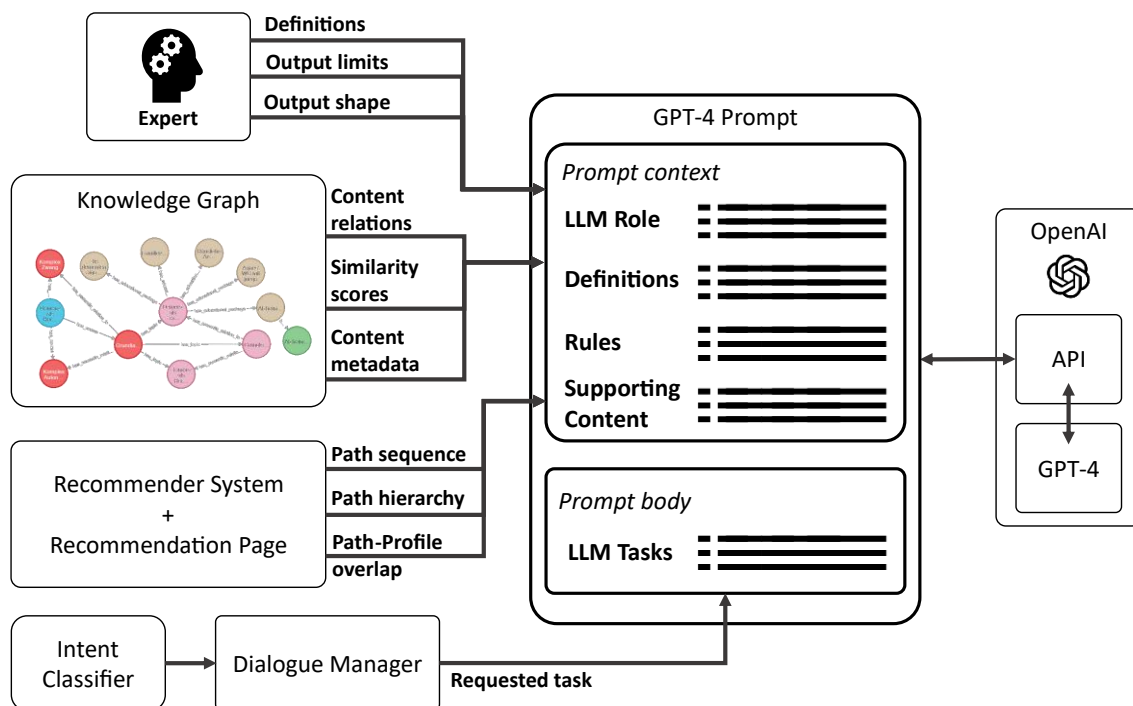


Figure 39: Constructing the prompt’s context from multiple information sources (Abu-Rasheed, Weber, et al., 2024b).

Recommended courses and their learning materials are ordered in the learning path by the recommendation algorithm. The path is also a part of the contextual information that the context builder extracts from the RS and the web page on which the recommended path is shown. Including the web page is needed since our Web-App presents the recommendation in multiple ways: textual, structural, and visual. One of the tasks allowed in our chatbot is asking questions about those formats.

To build the prompt’s context from these recourses, the context is divided into four sections: the **roles** that the chatbot plays, the **definitions** from the domain, the **rules** that are to be followed in generating the explanation, and the **additional content** that is retrieved from the KG. This context is then added to the main task, or set of tasks, that the dialogue manager defines based on the user’s request. It is noticeable that this

arrangement of the final prompt demands a larger volume of text to be transferred to the LLM. This is a compromise that was found necessary to reduce the risk of generating irrelevant outputs, even if the output is not wrong. Depending on the LLM used, the allowed size of context may differ, limiting the amount of contextual text that can be added to the prompt.

Part IV

Evaluation, Results, and Discussion

8 Evaluation Strategy, Experiments, and Results

To evaluate the proposed CARExKG method, a comprehensive evaluation strategy is adopted, in which each part of the system is evaluated individually, and the overall system is evaluated then as a whole.

The evaluation strategy is designed as a series of experiments, which are designed to evaluate the three parts of the system developed based on CARExKG method:

- The knowledge graph
- The recommendation system
- The explainability features

For the implementation of the experiments, quantitative and qualitative tests are conducted. The overall system is then evaluated in a user study with the staff members in two elderly homes in Germany.

The following sections present these experiments and highlight the relevance of the achieved results to the thesis hypothesis and research questions.

8.1 Knowledge graph evaluation

Knowledge graphs that were constructed within the three use cases in chapter 4 are evaluated within the context of their implementation. The evaluation of the knowledge graph offers a foundation for validating the answers to RQ1 and RQ2, in terms of the design and structural representation of the learning context in the knowledge structure, as well as the foundation for the algorithmic representation and integration of that context in the learning recommendation.

Three experiments are conducted for the evaluation of the KG, corresponding to each of the three use cases.

8.1.1 Experiment 1: Lessons learned information retrieval

To evaluate the value added from the KG to the design engineer, the first experiment defines a set of domain-specific key performance indicators (KPIs), which reflect the need for higher visibility of information regarding similar previous design cases. Defined KPIs include quantitative measures, which reflect the change in number of results through the utilization of graph relations.

A graph-based search is conceptualized and utilized in this experiment to exploit the connectivity of graph nodes for the purpose of enhancing document visibility to the user. The experiment is set for the application domain of microchip design, utilizing design documentation and maintenance cases. Two retrieval scenarios are designed for this purpose:

- direct search for a failure cases using search terms.
- indirect search in the graph upon the insertion of a new element in the project design.

The latter case is meant to highlight relevant maintenance documents and failure cases, which are related to an inserted component within a specific design structure, to the design engineer during the design phase.

The knowledge graph search mechanism that is adopted includes, for both cases, two stages. Following a search input from the user, the algorithm will find matching nodes and allocates firstly nodes in the graph that have a direct relation to the search term node with a distance of one. Those nodes are referred to as the base nodes. Relations are based on the document vector, N-grams or the structural information. Secondly, the algorithms allocated nodes that are connected to the ones discovered in the first step. Those new allocated nodes do not relate directly to the search term, but they are connected to it via other types of relations that have been extracted in the graph construction process.

In this mechanism, the number of search results can be controlled through leveraging several parameters. Two experimental parameters were utilized in this experiment, which are search depth and relation weights.

Search depth is defined as the number of intermediate relations that separate a base node from the furthest node connected to it. Within this definition, if the search depth is equal to 1, search results will contain base nodes, that are directly matching the query, and those that are connected to them. The deeper the search is, i.e. the more relation steps are followed between the nodes, the higher the number of results is. In fact, the number of results increases exponentially with the increase of search depth. Therefore, other parameters or search strategies are needed to govern the number of research results provided to the user.

To handle this issue, relation weights are utilized in this work, alongside the search depth, to order search results by relevance and provide the user with a reasonable number of connected documents. In the passive search scenario, triggered by adding components to the design process, this method forms the base of a simple recommendation process, which offers the user hints about relevant maintenance documents that are related to an inserted element and, in general, the current design.

Figure 40 represents a preliminary outlook on the quantitative change in the number of retrieved results, through the comparison between a normal key-word search in the database, and the search through extracted relations in the knowledge graph. With a group of 22 sample searches, the numbers show an increase of relevant results from the graph-database, which are retrieved through their semantic relations. It is noticeable also that direct search results themselves are different from the key-word results, since they are improved by the implemented text mining techniques. On average, a 75.6% increase of result retrieval, and thus visibility to the user, is achieved with the adoption of the proposed solution.

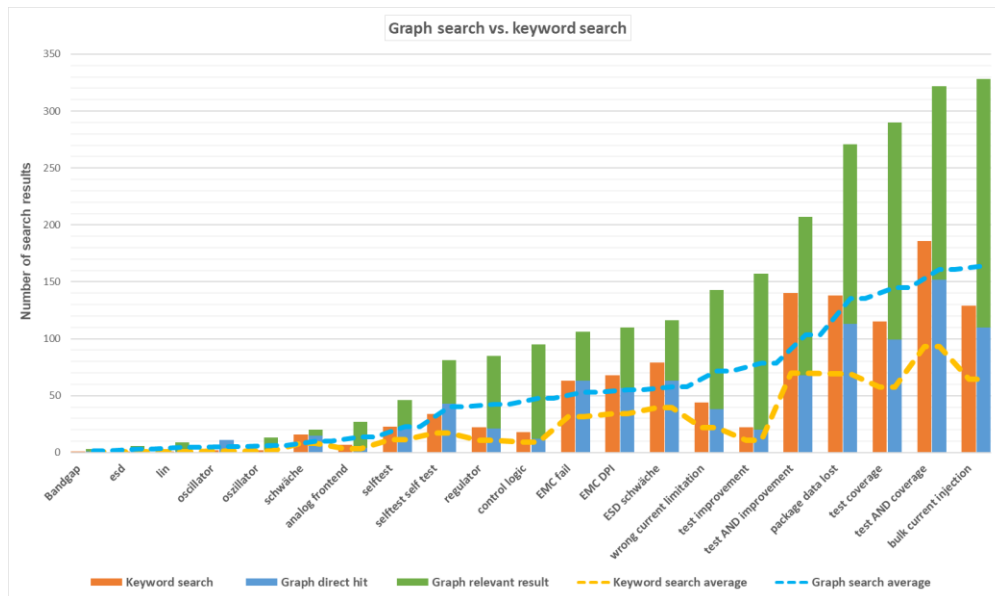


Figure 40. Quantitative comparison of search results (Abu-Rasheed, Weber, Zenkert, Krumm, et al., 2022).

Table 3 provides an overview of the relations extracted and utilized in the construction of the knowledge graph. Both inter-node-type and intra-node-type relations are shown in the table, and they highlight the condensed interlinking between failure case node in the graph. This is due to the fact that retrieving previous failure cases, as lessons learned, is the main objective of the graph search. This is also clear from Table 4, where the percentages of node numbers, relative to the total node count in the generated graph, is documented. It is noticeable that within the existing dataset, the number of maintenance documents, represented by failure case nodes, is relatively low compared to other types of documents. This is a usual limitation facing industrial datasets, due to the level of digitalization in the factory and the confidentiality of available data. However, through

Knowledge graph evaluation

the text mining of maintenance reports' textual content and the thorough relation extraction between them, they still possess the highest number of links in the graph, which positively influenced their retrieval rates, reaching between 28% and 43% more visibility, depending on the search configurations.

Table 3. Number of relations introduced by different types of nodes

Relation types	Node types	Number of relations
Intra-node-type	FC-FC	11146
	PE-PE	1862
Inter-node-type	PE-LN	9634
	FC-LN	12988

Table 4. Number of nodes in the graph

Node type	Node count	Percentage in the graph
Failure Cases (FC)	304	9%
Linking Nodes (LN)	1171	37%
Project Elements (PE)	1862	54%

Utilizing relation weights, as well as the depth of search as parameters to set the configurations of the search process, the number of search results for three search cases is illustrated in Figure 41, alongside the number of total relations in the graph, before and after inserting LN and PE relation types into the graph. For a unified search term and one level of search depth, the resulting number of maintenance documents that are visible to the user has increased by 28% when introducing the newly defined relations, which are based on the linking node and the intra-PE approaches. With the same number of relations, increasing the search depth to the second level has increased the visibility of documents by up to 43%, compared to the original case with only FC relation.

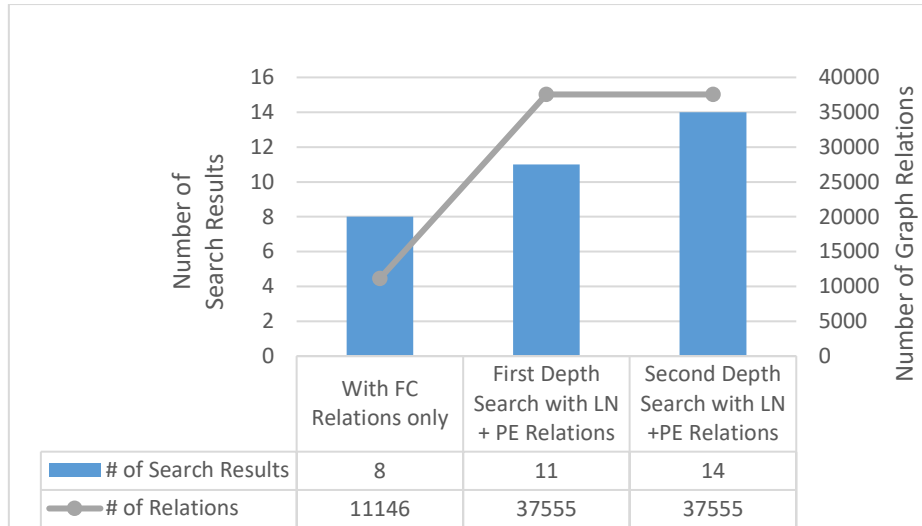


Figure 41. Effect of LNs and structural relations on the search results and graph connectivity (Abu Rasheed et al., 2021)

8.1.2 Experiment 2: Personalized learning in VET

From a wide range of KG evaluation metrics, which may have different indications in different domains, it is important to consider those metrics, whose indications can reflect the role of the KG in the learning process in TEL and VET. KG evaluation should also address the quality of the extracted relations, as a foundation for evaluating the quality of KG structure. Therefore, the evaluation strategy in the second experiments is also implemented quantitatively and qualitatively. Qualitative evaluation is accomplished with domain experts to validate the relation-extraction process and results, as well as the KG's role in contextualizing the learning recommendation. Quantitative evaluation is designed two-fold. The first evaluates the quality of extracted relation in the KG. It focuses on the text mining pipeline and evaluates its ability to connect LOs semantically in comparison to expert-created links. The second considers network metrics (Guéret et al., 2012) and quality-control methods (Zaveri et al., 2016), where those are classified in the domain of linked open data (LOD) into six dimensions: representation, accessibility, intrinsic, dataset dynamicity, contextual, and trust. To select relevant and meaningful metrics in the scope of this research, the results of (Giabbanelli et al., 2019), (Shin & Jeong, 2021), and (Kim & McCarthy, 2021) are utilized, which show that effective evaluation of KG's ability to represent learning context and the learner's mental model can be accomplished through metrics that correspond to contextual dimension (relevancy and completeness) and intrinsic dimension (interlinking, density, average degree). This work builds on those findings and use the metrics of *average degree centrality*, *clustering coefficient*, *weakly connected components*, *betweenness centrality*, and *closeness centrality*.

Evaluation Dataset

For the evaluation, a dataset of expert-curated OERs is used. The data set is provided by Tavakoli et al. (Tavakoli et al., 2022). Expert-curated content provides a high level of credibility to the baseline model, which is essential for a fair evaluation. After data exploratory analysis (DEA), LOs that do not have educational content associated with them are filtered out, and the duplicated and isolated ones are removed. The resulting hierarchical model used for creating the KG included 122 Journeys, 432 Courses, 767 Topics, 2565 Educational packages, and 7358 OERs.

Evaluation of the Semantic Relation Extraction

The goal of semantic relation extraction is to connect LOs that appear in the same or a similar learning context. This goal aligns with the best practices of content creators, who connect an LO to another one if both are needed in the same learning situation. This logic is utilized to evaluate the semantic similarity scores that the text mining pipeline calculates among Courses and Topics. First, the average of textual similarity scores among the LOs in each Journey (i) J_i^{sim} is calculated, by comparing each LO to all other LOs in Journey i . Then, the similarity score of each semantic relation ($SR_{i,j}^{sim}$) is compared to the average similarity of the journeys i, j that it connects

$$J_{i,j}^{avg_{sim}} = \frac{J_i^{sim} + J_j^{sim}}{2} \quad (21)$$

If the text mining pipeline is able to extract semantic relations with similarity scores that are comparable to the expert-curated relations, one can assume that the new semantic relation is meaningful. Comparability here is defined as being within the same range of both Journeys' similarity J^{sim} . This means that $SR_{i,j}^{sim} \geq J_{i,j}^{avg_{sim}}$.

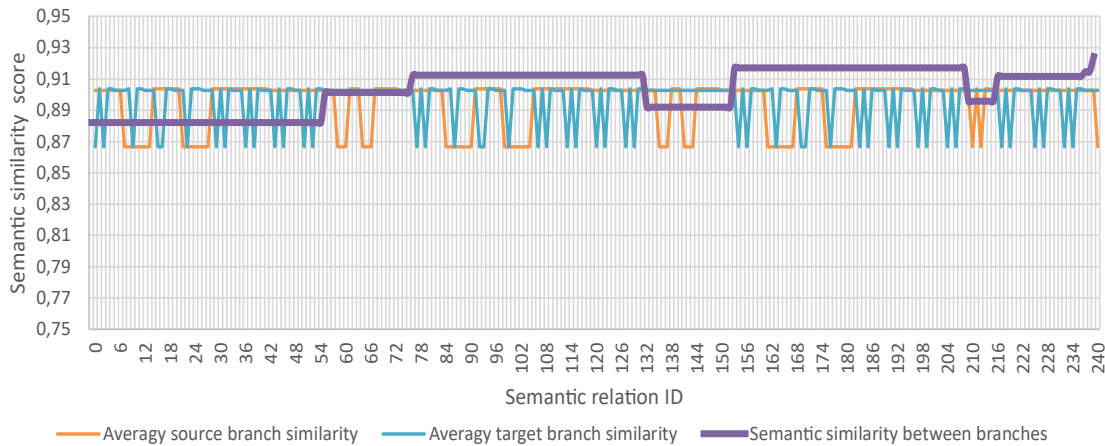


Figure 42. Semantic-relation textual-similarity scores (Purple), compared to the Journeys (Blue and Orange) connected by those relations through the text mining pipeline (Abu-Rasheed, Dornhöfer, et al., 2023).

Figure 42 shows the scores of a random sample of 240 semantic relations in the KG, alongside the semantic similarity of each Journey-pair they connect. From the results, one can conclude: 1) the similarity scores calculated within each Journey range between 86% and around 90%. This confirms the basic assumption that expert-curated LOs in Journeys are semantically similar. 2) 79% of the semantic similarity scores are either equal or higher than score averages calculated within Journeys, while the remaining 21% is slightly below the average. This indicates a high likelihood that the semantic relations between different Journeys in the KG are as meaningful as the relations that experts created within each Journey.

KG Evaluation Metrics

From the wide range of graph quality metrics, the selection in this experiment is based on the meaning of each metric in the scope of TEL and VET. Selected metrics reflect certain aspects of the KG, when interpreted from a pedagogical point of view in TEL and VET. The comparison here is conducted between the hierarchical structure and the proposed KG, in the light of the value-added that KG structure offers based on the interpretation of each graph quality metric. In the following, the selected metrics and their domain interpretation are elaborated. Table 5 shows the constructed KG results in comparison to the original hierarchical model.

Table 5. KG quality evaluation against hierarchical model. Preferred value-trends in VET and TEL use-case are explained

Evaluation Metric	Hierarchical data model	KG	Preferred value trends
Average Degree Centrality	1.079	2.262	increasing
Clustering Coefficient (Number of communities)	253	541	increasing
Clustering Coefficient (Average modularity score)	0.779	0.636	decreasing
Weakly Connected Components	63	35	decreasing
Betweenness Centrality	1.57	15.1	increasing

1. Average Degree Centrality (ADC):

Degree centrality measures a node's popularity in a graph (Freeman, 1978). It evaluates the connectedness of a node through its incoming and outgoing relations. In an educational use-case, ADC reflects LO's connectedness to other objects that take place in the learning scenario. The reason is that LOs are either connected by experts, or through a semantic relation. While the degree centrality in a hierarchical model represents the LO's relation to objects from the same Journey, the increase of that object's degree centrality in the graph model shows that new relations have been found to other Journeys, meaning that the LO may appear in more learning contexts. The degree centrality is calculated for

Knowledge graph evaluation

all graph nodes, with and without semantic relations. Then, the (ADC) scores are averaged over the KG. The results show that the ADC has doubled in the KG model.

2. *Clustering Coefficient (CC):*

CC reflects how strongly a node belongs to a community in a graph. Communities are clusters of nodes that are densely inter-connected. A community is considered more modular if it is easy to separate it from other communities. Communities of LOs in our KG represent the context of those objects. Therefore, the higher the community count is, the more contexts are represented. Modularity, on the other hand, reflects the potential to separate a learning context from another. Therefore, a lower modularity score is preferred, since it means that learning contexts are well connected to each other. Evaluation results show that the proposed KG increased the detected communities, due to the semantic relations, and decreased graph modularity.

3. *Weakly Connected Components (WCC):*

WCC detects sets of nodes in the graph that are loosely connected to other graph parts. Therefore, a high WCC score reflects poor connectedness of LO groups in the data model. Using Monge and Elkan's algorithm (Monge & Elkan, 1997), an effective reduction was found of the loosely connected node groups in the KG.

4. *Betweenness Centrality (BC):*

BC describes the impact of a node on the flow of information within that graph. It is especially relevant to VET and TEL, since LOs with high BC work as bridges that link individual learning goals to each other, to solve a more complex problem (Trumpower et al., 2014). Using Brandes and Pich's algorithm (Brandes & Pich, 2007), our results show that the KG has an average BC value of 15.1, which is about 10 times the BC score of the original hierarchical model.

Qualitative expert-evaluation

This experiment also evaluates the KG construction and contextualization approach with domain experts in two focus groups. The groups included experts in VET and researcher-training programs. A total of nine experts participated and answered three questions on 1) The role of connecting learning goals in learning contexts, 2) KGs as contextual data models in their domain, and 3) Using LO's textual description to mine contextual relations. All participants emphasized the need for connecting individual learning goals for solving real-life, job-related problems. Six participants pointed out that highly connected nodes in the graph can represent transferable skills among different domains. Three experts found a direct use of connecting multi-lingual LOs in their daily work. They also pointed out that a recommendation of multi-lingual content should avoid repetition of the same content written in different languages. Experts also agreed that textual content and description of LOs are important sources of extracting contextual

information. They indicated that content creators should roughly understand the way algorithms extract semantic relations, so that they can enrich the descriptions of their LO content with useful contextual information, allowing better connectivity in the KG. The pedagogical experts also addressed the dynamic change of the learner's context, and the development of the learning domain, which needs to be handled through a continuous update of the KG.

8.1.3 Experiment 3: Personalized learning in higher education

knowledge graph from the content of two university modules, namely Embedded Systems and Development of Embedded Systems Using FPGA. The two courses are taught in the University of Siegen, Germany, for master- and bachelor-level students. These courses were chosen because they are taught separately, despite the potential overlap between them. Each course is composed of 9 sessions, spanning the lecture period of one academic semester. Each weekly session covers a diverse number of topics, which differ in their level of detail.

Expert Evaluation of the Ontology and Pipeline

The evaluation is conducted with module lecturers, to ensure high-quality feedback based on their experience with the module's content and learning goals. Evaluators included one professor and two PhD assistants, working in the Embedded Systems Institute in the same university. Qualitative feedback from the lecturers about the base ontology and the overall acceptance of the approach for human-AI collaboration is collected. Feedback is collected through group meetings with the lecturers, in which the ontology was presented and evaluated from the perspective of use cases they face in their teaching.

Moreover, quantitative input from the lecturers regarding the accuracy of extraction and classification of topics, sub-topics, and descriptions through the pipeline is collected. To achieve this quantitative feedback, each lecturer checked a comprehensive list of extracted topics, their sub-topics, and the sub-topic descriptions, and evaluated the following aspects for each sample:

1. The correctness of extraction.

Lecturers also evaluated the structural extraction and verified whether the extracted sub-topics were correctly positioned within the corresponding topics.

A total of 1197 extraction samples were evaluated by the experts (173 Topics, 512 Sub-Topics, and 512 Descriptions). Table 6 shows the results of the Precision, Recall, and F1 measure for *Topic* and *Sub-Topic* classes in both evaluation modules. For the *Description* class, the Precision value is calculated to verify whether the long description text that was generated automatically fits the corresponding sub-topic.

Knowledge graph evaluation

The proposed pipeline and LLM-supported extraction of topics and sub-topics has shown a high capability of detecting the topics and their corresponding sub-topics from the lecture materials, see Table 6. A slight difference can be noticed between the two modules, which is traced back to the level and style of structuring that each lecturer uses to organize their lecture recordings. While some recordings only included the lecturer’s explanation of the session’s content, other recordings were simply made of the live, online session, including the discussions with students, their questions, and the lecturer’s answers. The difference between these two approaches had a considerable influence on the concept extraction from transcribed videos in the first iteration of LLM extraction and classification. The model extracted concepts from the side discussions, questions, and answers, as well as from the lecturer’s highlights about topics that will be covered in the course. Those topics, however, were not in fact discussed in the corresponding session. To account for this effect, a prompt engineering approach is used to set explicit rules and limitations for the way each video transcript is handled. The prompt engineering process also defined the ontology classes to the LLM, to clarify the contextual meaning of the *Topic* and *Sub-Topic* classes.

Table 6. Precision (p), Recall (R), and F1 measure values for Topics and Sub-Topics of the evaluation modules

	Topic			Sub-Topic			Description
	<i>P</i>	<i>R</i>	<i>F1</i>	<i>P</i>	<i>R</i>	<i>F1</i>	<i>P</i>
ES^a	0,99	0,94	0,96	1	0,97	0,98	1
FPGA	0,97	1	0,98	0,89	0,99	0,94	0,89

^a. Embedded Systems

It is important to highlight here that the differences among lecturers in the ways they record their lectures are vast and therefore cannot be only accounted for in the prompt engineering steps. To that end, diversifying the input of the model, e.g., by supporting the video transcriptions with structured lecture slides or references, will be important to ensure a good quality extraction. Moreover, the KG itself can play the role of a supporting data source, e.g., in a retrieval augmented generation (RAG) approach, to extend the context of the LLM’s prompt and increase the relevance and accuracy of the extraction and classification tasks.

2. KG-Structure

To evaluate the resulting knowledge graph structure, graph quality measures: Average Degree Centrality (ADC), and Graph Modularity are used. These measures validate our research goal of discovering connections between learning contents from multiple modules. Average degree centrality evaluates the connectedness of a graph node through its incoming and outgoing relations. In this use case, ADC highlights the role that the semantic similarity algorithm plays in increasing the density of relations among graph nodes. Graph modularity, on the other hand, evaluates the potential to separate learning

contents in the graph from each other. Therefore, a lower modularity score is preferred in our case, since it reflects stronger connections between similar contents.

In the resulting KG, an increase of the ADC value from 0.9 to 1.03 is calculated, as well as a slight modularity decrease from 0.769 to 0.767. This slight change can be traced back to the small number of semantic relations added between the two modules, due to the relatively small size of the sample data. Moreover, this is also a result of the high similarity thresholds that have been set for the semantic relation extraction algorithm, to ensure high-quality relations between the topics and sub-topics.

Despite this relatively small improvement on the KG structurally, qualitative expert feedback pointed out that the automatic topic extraction and the possibility to semantically connect topics from different courses is a considerable help for the lecturers in re-structuring their course contents. This is accomplished by being aware of similar content in other courses and modules, which might be overlapping, or can be re-used. The evaluators pointed out that the insights they received from the pipeline's output did not only support a better structuring of their courses and modules, but also the personalization of the learning experience for individual students, who have different levels of knowledge, scientific backgrounds, and career goals

8.2 Graph-based recommender system

8.2.1 Experiment 4: Personalized Nursing VET

The evaluation of the recommendation system in VAT learning setting was conducted during fall 2023 through spring 2024 with the staff of two elderly homes, to evaluate enhanced learning components such as "explainability" and "priming" within algorithm-assisted mentoring systems. The test sought to refine earlier findings and explore new experimental setups regarding the sequencing and dynamic personalization of learning pathways.

This experiment aimed to investigate how different combinations of learning recommendations ("correct" vs. "reversed") and additional supportive features ("explainability" and "priming") influenced learners' reflective capacities, resistance levels, and motivation.

User Sample

The study involved six randomized groups, each exposed to distinct learning conditions:

- Group 1: Learning recommendations without "explainability."
- Group 2: Learning recommendations with "explainability."
- Group 3: No learning recommendations with "explainability."

Graph-based recommender system

- Group 4: Correct learning recommendations alongside "priming" examples.
- Group 5: Reversed learning recommendations without "explainability."
- Group 6: Reversed learning recommendations with "explainability."

Each group consisted of participants who were representative of the target population in terms of demographics and educational background. The inclusion criteria ensured a broad range of learners, from those with limited prior exposure to digital learning environments to experienced users.

Data Collection Methods

- Pre- and post-surveys measured participants' attitudes, resistance levels, and self-reported reflexivity improvements.
- Observations captured real-time user engagement with content and adaptive recommendations.
- Thematic interviews provided qualitative insights into learner experiences, resistance behaviors, and motivations.
- Algorithmic monitoring logged interaction data, including time spent on tasks, recommendation follow-through rates, and progression patterns.

Experimental Setup

- Both "correct" recommendations (aligned with learners' expected pathways) and "reversed" recommendations (deliberately misaligned sequences) were tested to analyze the impact of sequencing on outcomes.
- "Explainability" features provided contextual guides and transparency about why certain content was recommended.
- "Priming" introduced reflective examples and contextual questions to elicit deeper engagement with the learning material.

Results and Findings of the Test

The results of the test revealed important insights into how learning recommendations, sequencing, and supportive features influenced learner outcomes:

- **Impact of Learning Recommendations:**

Groups that received correct recommendations consistently showed higher levels of reflective thinking compared to groups without recommendations, see Figure 43. Even reversed recommendations had a positive impact on reflection compared to control groups with no recommendations, although these effects were less pronounced.

- **Role of Explainability:**

"Explainability" clearly reduced learner resistance across most groups. Participants felt more confident and motivated when they understood the rationale behind specific recommendations. In groups with reversed recommendations, "explainability" mitigated confusion and frustration, enabling learners to adapt and engage more constructively, see Figure 44.

- **Effect of Priming:**

Groups exposed to "priming" examples (correct recommendations paired with reflective prompts) demonstrated the highest levels of reflexivity and engagement. Participants in these groups developed more complex understandings of the material and were able to articulate reflective insights beyond surface-level learning.

- **Sequencing and Its Impact:**

The experiment with reversed recommendations unexpectedly revealed an important insight: learners were still able to achieve reflective outcomes under misaligned learning sequences, provided that "explainability" was present. Correct sequencing was more effective overall, but the supportive features allowed learners to navigate even disruptive pathways with minimal resistance.

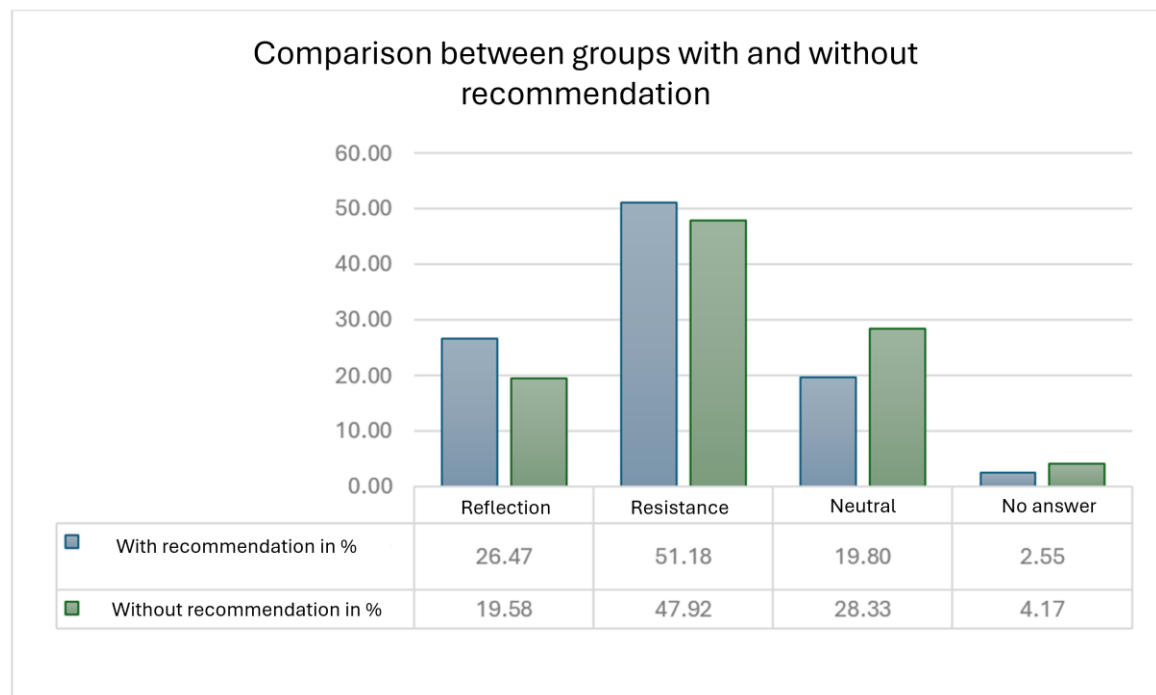


Figure 43. Groups comparison results with and without recommendation (Abu-Rasheed et al., 2024).

Observations and Reflections from the Results

The findings of this test highlighted several key observations and reflections that have implications for the future design of algorithm-assisted learning systems:

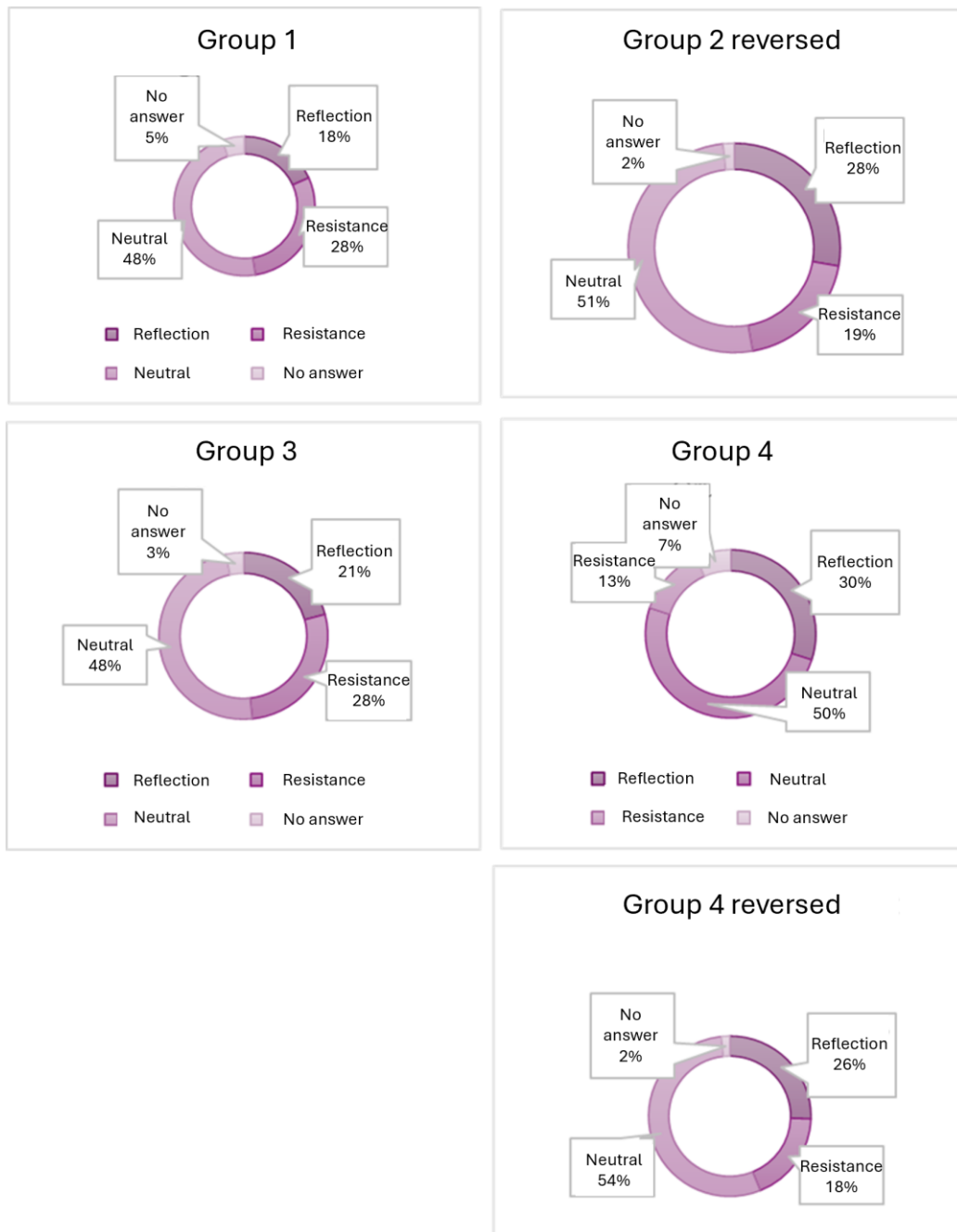


Figure 44. Group comparisons on the four evaluation dimensions in pre- and post-tests (Abu-Rasheed et al., 2024).

- Algorithmic recommendations and reflexivity:** Learning recommendations are a crucial factor in fostering reflective thinking, but their effectiveness heavily depends on alignment with learners' goals and expectations. Correct recommendations maximize engagement and comprehension, whereas reversed recommendations, while disruptive, can still yield positive results if learners are adequately supported.
- The value of explainability:** Explainability emerged as an essential component for mitigating resistance. Providing transparency about the logic behind recommendations enhanced learners' trust in the system and reduced feelings of

defensiveness or confusion. This, in turn, allowed participants to focus more on reflective learning.

- **Priming as a catalyst:** Priming examples significantly boosted reflective engagement. By situating content within relatable or illustrative contexts, learners were able to connect abstract concepts to their own experiences, fostering deeper understanding and emotional resonance with the material.
- **Interplay between motivation and content design:** Resistance levels were inversely correlated to reflexivity when recommendations were present, demonstrating that motivated learners tend to engage more deeply. The design of content plays a critical role in balancing cognitive engagement and emotional willingness.
- **Optimal conditions for reflective learning:** The ideal setup for expansive, reflective learning appears to be a combination of correct recommendations, explainability, and priming. This triad addresses cognitive, emotional, and motivational dimensions, providing a holistic approach to personalized learning.

This experiment demonstrated how algorithm-assisted mentoring systems, augmented by supportive features like explainability and priming, can enhance both individual reflexivity and overall learning outcomes. These findings underscore the importance of thoughtful content design, adaptive sequencing, and transparent algorithms in creating effective digital learning environments. Future iterations should continue to explore diversification in user groups, refine sequencing strategies, and expand the role of reflective scaffolding for broader applicability across educational contexts.

8.3 TE_xKG framework

8.3.1 Experiment 5: evaluation use cases

Use case1: chip design information retrieval

Graph-based explainability framework is evaluated in this use case in terms of their ability to represent the domain and generate meaningful, transferable explanations and search results from the KG.

To test the knowledge graph's ability of representing the domain, an ontological structural-evaluation metric is used, namely the class richness (García et al., 2010; Tartir & Arpinar, 2007), which is defined in Eq. (21).

$$\text{Class Richness} = \frac{\text{No.ClassesWithInstances}}{\text{TotalNo.Classes}} \times 100\% \quad (22)$$

Class Richness measure reflects how the graph covers the instances of the use case. The underlying assumption for using this measure is that the elements of the framework, expect for the algorithms, can represents classes of an ontology. If a class is capable of

representing a domain, then it will have instances defined from that domain. The more classes having instances, the more representative is the proposed framework of the domain of interest.

Evaluation of the Explainability Algorithm

Testing the explainability algorithm is done through four different metrics. Firstly, the availability of information for generating an explanation based on the knowledge graph is measured. *Information availability* is used in this context to measure the quality of the explanation. It represents the amount of information that the framework provides about the retrieved results from the KG. The more information collected about the algorithm’s reasoning, the more comprehensive the explanation will be, and thus the higher the explanation quality is. To quantify this measure, the explanation templates are that are used for generating the textual explanations are utilized. The explanation is considered of a high-quality if all the slots of the explanation template can be filled with information from the KG. Otherwise, the explanation is considered of a low-quality. The *information availability* measure is then defined as the ratio of high-quality explanations to a total number of explanations generated by the explainability algorithm, as in Eq. 22.

$$Information\ Availability = \frac{No.FullyGeneratedExplanations}{TotalNo.Explanations} \quad (22)$$

In addition to the information availability, the *mean explainability precision* (MEP) is calculated (Abdollahi & Nasraoui, 2016; Y. Chen & Miyazaki, 2020) for the explanations that are associated with the IR output, as defined in Eq. 23. *MEP* measures the average proportion of explainable results of an IR algorithm, or recommendations of a recommender system.

$$MEP = \frac{1}{U} \sum_{u=1}^U \frac{N_{exp}}{L} \quad (23)$$

Here, U is the number of users, N_{exp} is the number of explainable results, and L is the total number of results.

The third and fourth metrics focus on the quality of the explanation itself. For that purpose, *bilingual evaluation understudy* (BLEU) (Papineni et al., 2001) and *Rouge-L* (Lin, 2004) scores are used to calculate how much a generated explanation represents the explained result. BLEU score measures the closeness between a machine-generated text and the original human-defined one. Therefore, it is used to measure how a generated textual explanation represents the original information that the KG paths have provided to the explainability algorithm. Rouge-L score measures the ability of a machine generated text to summarize the original information. In this experiment, this score represents the ability of the explanation to summarize the information that led to retrieving a certain result from the KG structure.

While F_1 measure, *Class Richness*, *MEP* score, as well as *BLEU* and *Rouge-L* scores are all well-defined measures that are used in the literature, the *Relevance* and *Information Availability* metrics are developed to correspond to the nature of the proposed domain-oriented framework, and to demonstrate its ability of performing accurate information retrieval and generating comprehensive explanations for the different domain-specific environments it is used for.

The implementation of the framework is done within two evaluation use cases. A focus is given to the strategy of implementing the framework's components in each case, the results of evaluating the framework's domain coverage, and its transferability.

For the *Class Richness* metric, entities in the framework are considered as classes. Algorithm boxes are neglected since they do not follow a standard ontological representation. By defining the instances of each component in the use case and calculating the overall *Class Richness* of the knowledge graph, a value of 88% class coverage is achieved. This value reflects a high coverage rate of the proposed framework over the domain-specific requirements and their corresponding components in this use case.

To evaluate the IR and explainability algorithms, a sample of 300 random searches through the IR algorithm is performed. First, the *Relevance* score is calculated for each retrieved search result, to evaluate the performance of the IR algorithm. Then, the *Information Availability*, *BLEU* score and *Rouge-L* score are calculated to measure the quality of the explanation generated for each result, Figures 45 and 46.

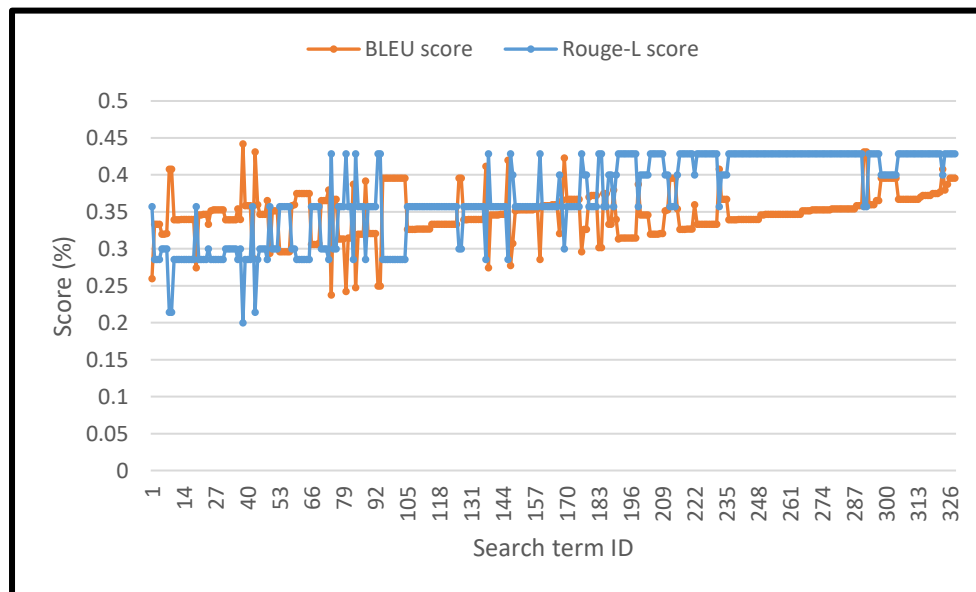


Figure 45. BLEU and Rouge-L scores for the search results. The figure represents the scores for the short search queries. The average BLEU score for the short search queries (on the horizontal axis) reached 37.2%, and the average Rouge-L score was 25.1%.

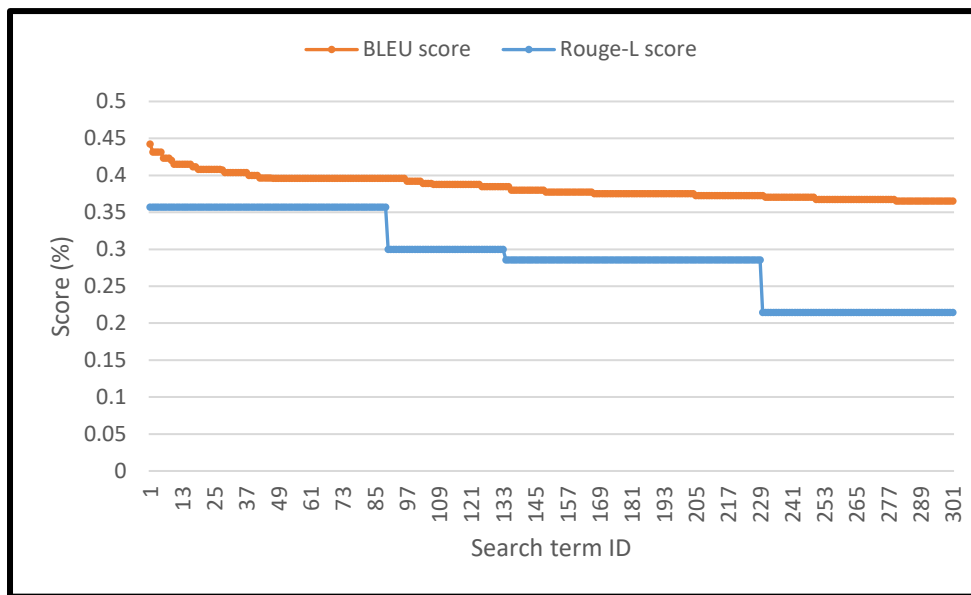


Figure 46. BLEU and Rouge-L scores for the search results. The figure shows the scores of the long search queries. The average BLEU score for the long search queries (on the horizontal axis) reached 38.3%, and the average Rouge-L score was 27.4%.

For each search query, the *MEP* score is calculated for the overall explanations generated for all retrieved results, as given in Eq. (23). An average value of 99% is achieved for the *MEP* score. This result is traced back to the inherently open-box nature of the KG, which provides the IR and explanation algorithms with sufficient information about the retrieved results. Evaluation scores are shown in Table 7.

Table 7. Evaluation results of the proposed framework on use case 1.

	Graph evaluation measure	IR evaluation measure	Explainability evaluation measure			
	<i>Class Richness</i>	<i>Avg. Relevance</i>	<i>Information Availability</i>	<i>MEP</i>	<i>Avg. BLEU</i>	<i>Avg. Rouge-L</i>
Short search query	88%	87.3%	78.7%	82.5%	37.2%	25.1%
Long search query		99%	91.8%	99.6%	38.3%	27.4%

Use case2: CV-Job recommendations

In the second use case, a different domain to test the KG is selected in the field of job recommendation system. The field of matching job-postings to job-seekers features a

different set of requirements, which is a result of the varying structures and content of job postings and job-seekers' CVs (Upadhyay et al., 2021).

CV documents include information that summarize the user profile. Recommending a certain job to the CV holder requires finding the similarities between their skills and qualifications to those required by the job posting. This, in turn, requires the information retrieval algorithm to focus on parts of the CV that include the information about the user's qualifications and skills, more than other sections, such as the hobbies and interests. This also applies to the job postings themselves, where one can find several parts about the company portfolio, which are repeated in every job posting and thus does not participate in generating an effective matching to the specific job seeker.

To accomplish this matching effectively, a customized named entity classifier (NER) is built to support the information retrieval, and later the recommendation and explanation systems, in separating the different textual parts of the job posting and the user's CV to better model them in the knowledge graph.

Dataset Preparation

For this use case, an experimental dataset that includes 50 user profiles and 1606 job postings is created and utilized. Job descriptions in the fields of IT and Engineering are collected from the job-hunting website "Monster"¹, in compliance with the European General Data Protection Regulations (GDPR)². All user profiles have been anonymized before including them in the dataset.

In the experiment dataset, personal and demographic information about the user are removed, without compromising the performance of the job recommendation system. This is because the recommender is designed as a graph-based collaborative-filtering (CF) system, following the shortest path algorithm of Dijkstra (Dijkstra, 1959).

Knowledge Graph Construction

The knowledge graph is constructed from the two selected types of documents. It models them in two node types: *userProfile* (CV) nodes and *jobPosting* (JOB) ones. Node properties in this case include the textual content of multiple parts of the document, as well as the classified entities in the text, as predicted by the customized entity classifier.

The graph includes three types of relations:

1. *userProfile-jobPosting* which relates the user to potential jobs directly, see Figure 47,

¹ <https://www.monsterindia.com/>

² <https://gdpr-info.eu/art-5-gdpr/>

2. *userProfile-userProfile* which finds the relevancies amongst users, and
3. *job-Posting-jobPosting* relations.

The last two relation types are extracted to support the CF-based recommender system. This happens through finding job posting recommendations based on the transitive relations in the knowledge graph. These create paths in the graph connecting a userProfile node to a jobPosting one, through an intermediate node, which is either a userProfile or a jobPosting, as shown in Figure 48.

Figure 43 shows the calculated similarities for two user profiles with the job postings in the database. From the similarity distribution analysis, the threshold value of >40% (Slightly above average distribution) is selected, and the top 4 similar job postings are recommended to the respective user profile.

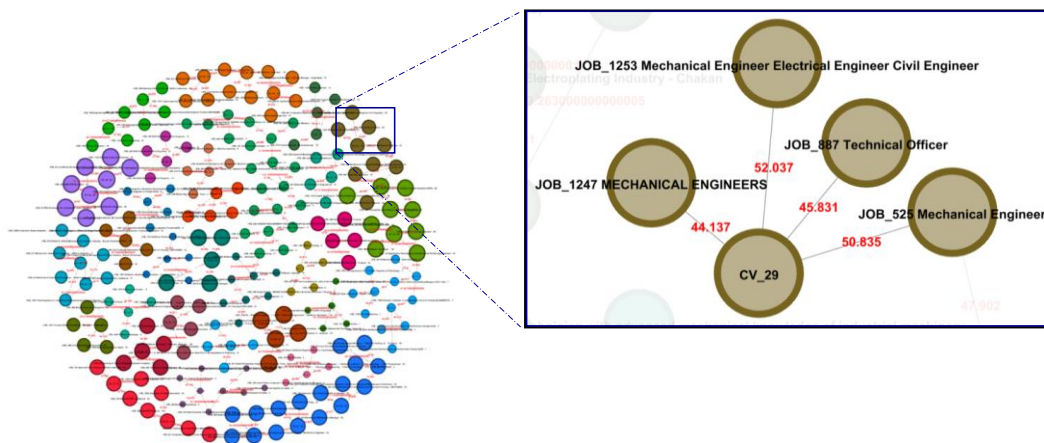


Figure 47. Direct relations between user profiles and relevant job postings. Each relation is defined based on the similarity score between the two nodes.

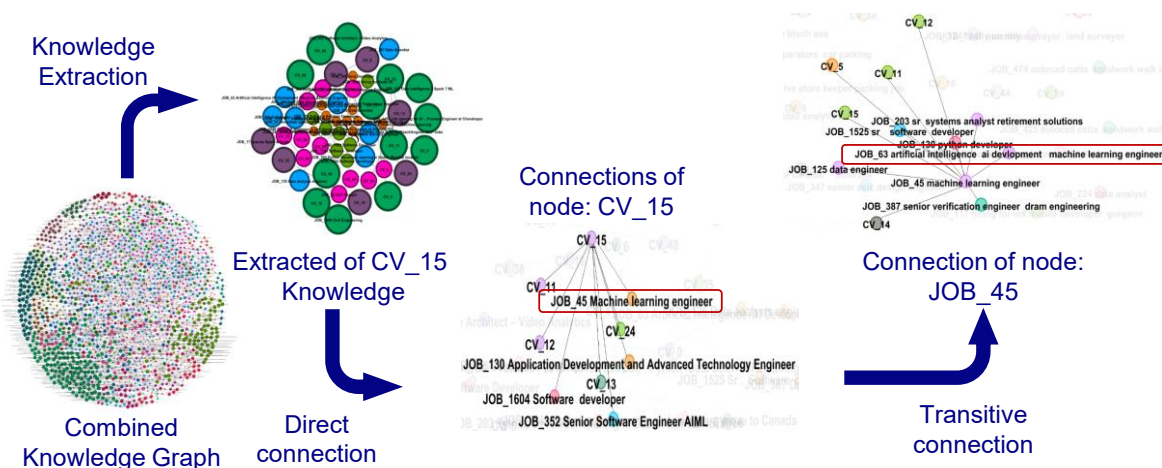


Figure 48. The use of direct and transitive relations in the knowledge graph. Here, jobPosting 63 is recommended to the userProfile_15, since it is relevant to the directly connected jobPosting_45.

The requirements of the job recommendation domain were very different from those in the chip-design one. The structure of user profiles and job descriptions, as well as the textual content in each one, revealed the need to classify the different sections of the document, to enhance the recommendation result. This was a domain-specific requirement that is modeled through the “domain-requirements” component in the framework. It took the form of a list of document classes that created section-based relations in the knowledge graph, and influenced the recommendations of a job postings.

Experts also provided rules on how different sections of the user profile have different levels of importance when generating the relation between a user profile node and a job posting one. Expert defined rules in this use case also influenced the textual content of document nodes. This, in turn, affected the documents vectors and the corresponding similarities. An example of such rules would be the handling of numerical content of user profiles, where values of months and years were irrelevant to the relation extraction algorithm.

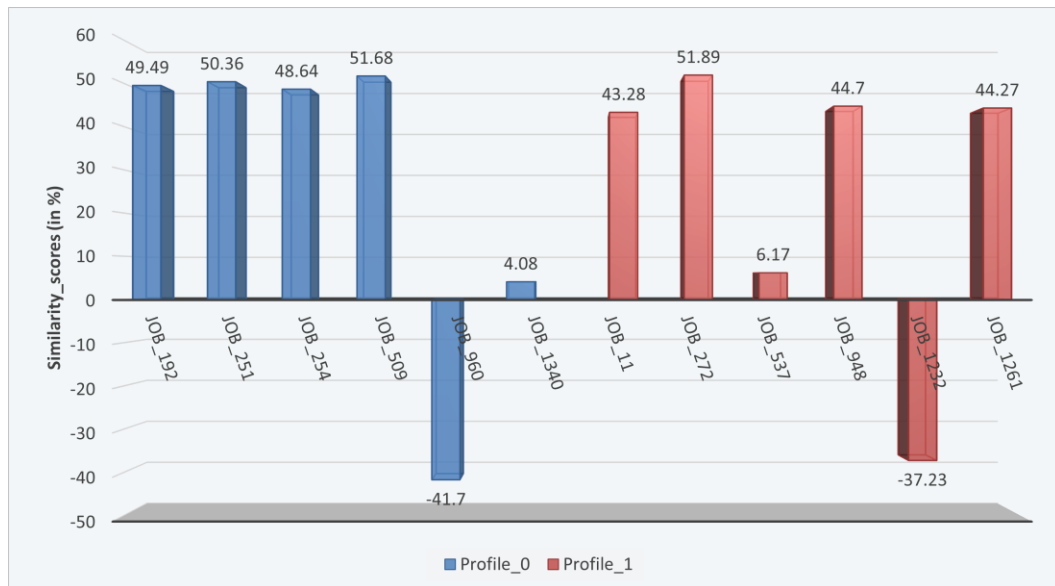


Figure 49. A sample of the overall calculated similarity scores between user profiles and job postings. Similarities above a predefined threshold are used to create corresponding relations in the knowledge graph. Negative similarity scores are caused by the angle value between document vectors, which results in a negative cosine value of that angle when calculating the cosine similarity.

The recommendations of relevant jobs are achieved on two levels:

1. the direct level, where the user is recommended the most similar job postings that are directly connected to their corresponding node in the graph, and
2. CF-based recommendations that are recommended to the user based on the intermediate relevancy to another user or job posting. This means that a userProfile (A) can get a recommendation of a jobPosting (B) because userProfile (A) is directly connected to userProfile (B), which is directly connected to a jobPosting (B).

For example, if User (A) has a skill as a data scientist in their profile, while User (B) has a skill as a data analyst, then the similarity between those skills will link both users. In this case, User (A) will also receive a job recommendation for a data analyst position if it was recommended to User (B).

Similarly, another transitive path can go through an intermediate jobPosting node, as follows: userProfile (A) → jobPosting (A) → jobPosting (B). An example of this case would be: If User A, receives a job recommendation for a data scientist position, and this position is already related to a data analyst position, then User (A) will also receive the recommendation for the data analyst position. An example of the path generation for the CF-based recommender is shown in Figure 48.

Figure 49 illustrates the cosine similarity scores calculated between two userProfile nodes and multiple jobPosting nodes in the KG. The horizontal axis represents the jobPosting in question, while the vertical axis represents the calculated similarity score. Each userProfile is illustrated in a different color. Negative scores that appear in Figure 49 are a result of the angle between the document vectors (here 90°). Those scores are considered in their absolute value in the graph relation definition.

Explanation Algorithm

The explanations generated for this use case used the same algorithm than in use case 1. Textual explanation templates were used with information slots, which are filled based on the knowledge graph. In this use case, however, the explanation requirements were different. This is because explaining the job recommendation needed to include specific information about the parts of the user profile that matched the job description. Those parts were categorized based on the custom entity classifier.

The NER is designed as a support vector machine (SVM) and trained it on a custom dataset of 60000+ entities. The dataset was created from documents in the technical and engineering domain. Those documents mainly represented job postings. The data was labeled manually by three independent technical personnel. Dataset labels have been chosen to correspond to the domain of job matching and recommendation. The classifier was therefore trained to predict 7 classes of entities that a user profile may include. These classes are *programming_framework*, *skill*, *software*, *study*, *job_title*, *location*, and *programming_language*. Based on the predicted classes, the similarity scores between a user profile and a job description, as well as the graph transitive relations, the explanation template was constructed as shown in Figure 50. The first explanation template uses syntactically similar words to fill the information slots with similar terminology between the user profile and the job posting. The second template is equipped with more information through implementing a customized name entity classifier on the language model. An example of the final explanation is:

“I recommend you Job number 509 as a ”Software Engineer”, which is a match to your profile by 51.682%, through the following classes:

Study: ['science', 'analysis', 'control', 'bachelor', 'data'],

Software: ['oracle'],

Programming language : ['java', 'python'],

Skills: ['agile', 'research', 'development', 'software']

Other similar terms to this job are: ['architecture', 'server', 'product', 'framework', 'design', 'technology']

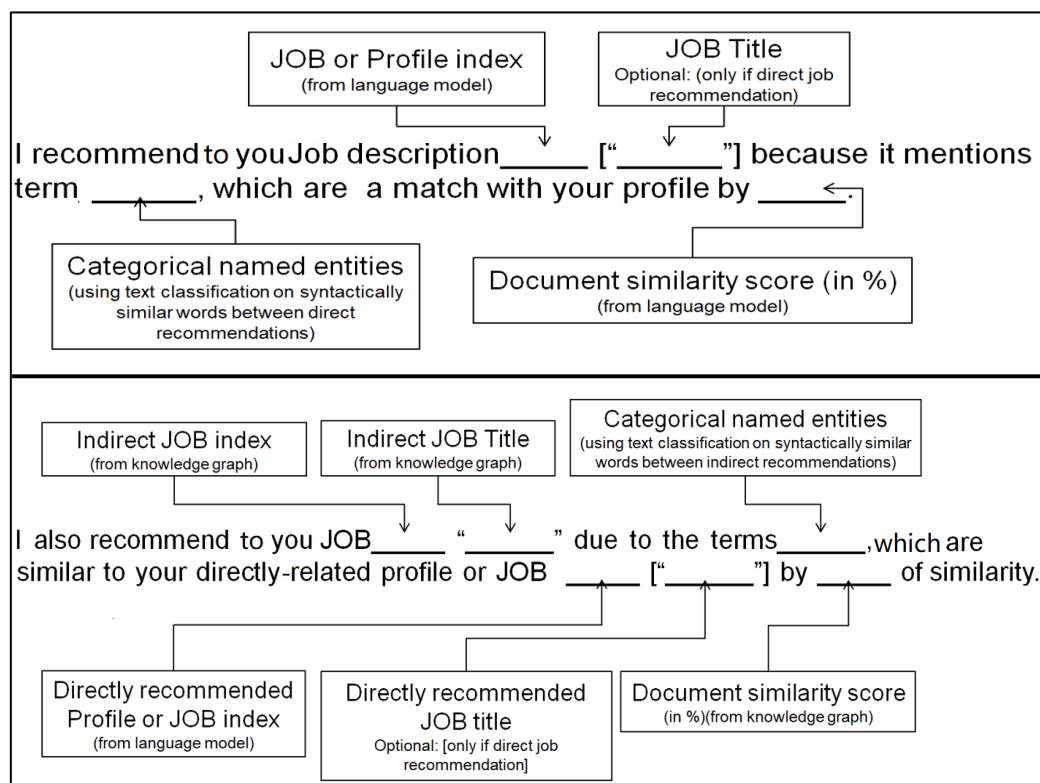


Figure 50. Two templates that are used to generate the textual explanations. The first template is designed for direct recommendations, where only the link between the user profile and the job posting is explained. The second template includes information about the in

To evaluate the different parts of the framework in this use case, the KPI metrics defined were calculated and shown in Table 8, where the accuracy, precision and recall measures for the customized NER are documented. An overall F1 score of 88% is achieved in this experiment based on the training dataset in the domain of job-matching.

To evaluate the knowledge graph structure, *Class Richness* measure is used. A percentage of 89% coverage of the graph is achieved for the instances of CVs and job descriptions.

Table 8. Precision, Recall and F1 scores of the customized NER, with regard to the predefined classes

	Precision	Recall	F1-score	Support
Programming_framework	0.84	0.85	0.84	241
skill	0.794	0.88	0.83	2610
software	0.79	0.90	0.85	902
study	0.87	0.86	0.86	4481
job_title	0.89	0.99	0.94	480
location	0.99	0.96	0.98	485
other_word	0.93	0.88	0.90	9190
Programming_language	0.92	0.98	0.895	642
accuracy			0.88	19031

Evaluating the explainability was conducted through calculating the *MEP*, *BLEU* and *Rouge-L* scores for the explanations generated for recommended job postings. An *MEP* score of 85.6% is achieved for the recommended job postings. To calculate *BLEU* and *Rouge-L* scores, the user profile and its recommendations' content are combined for reference, and the generated explanation is fed as a test sentence. The results are shown below in Figure 51 for multiple profiles and job postings. According to (Gambhir & Gupta, 2017), the *BLEU* score and *Rouge-L* precision score are in a good range if they are above 20%, which shows that the generated explanations are related to the recommended job descriptions. *BLEU* score ranged around an average of 23.5%, and the *Rouge-L* scores ranged around an average of 26%.

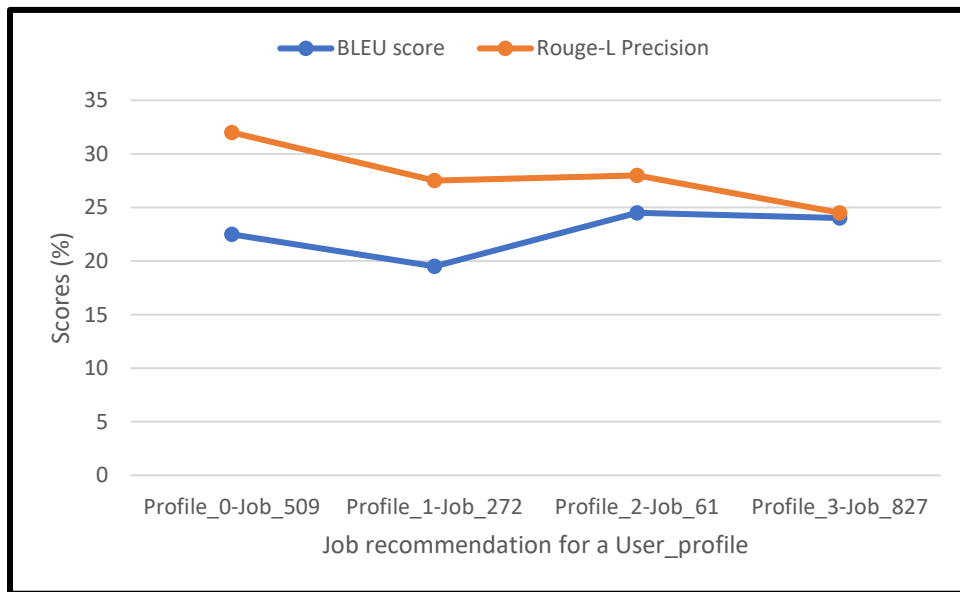


Figure 51. BLEU and Rouge-L scores for a sample of recommended job postings to two user profiles.

Discussion

The implementation of the framework in the presented use cases shows its adaptability to multiple domains while generating explanations for the retrieved search results and recommendations. In the evaluation experiments, a focus was put on the main contributions of the framework, namely the ability of the framework to embed domain-specific requirements, and the transferability of the framework to multiple domains with minimum to no changes in the intelligent algorithms.

Embedding domain requirements in the two use cases was a direct result of adopting the domain-specific components in the proposed framework. Those requirements are captured from multiple parts of the system, including the rules that domain experts define, the results of the EDA process, and the content of databases. To the best of our knowledge, no other solution has been suggested in the literature to and how to include domain requirements from all of those parts. The importance of including other sources of domain requirements can be seen in the ability of expert defined rules in contrast to existing solutions, for example, to represent differences between the documents, which cannot be included in the database itself. This hinders other solutions from reaching the level of domain-adaptability that the presented framework achieves.

Results from the experiments show that adopting the IR and explainability algorithms to requirements from a certain domain did not limit its ability to be transferred to other domains. The domains, within which the framework was implemented, revealed different features and requirements. Despite that fact, the use of the same IR and explainability algorithms was possible without compromising their accuracy. Numerical evaluation showed that the IR algorithm maintained an accuracy level higher than 95% in both use cases. The comparable literature focuses on the accuracy of the IR or recommendation algorithms, namely (Moon et al., 2019; Song et al., 2019; Wang, Wang, et al., 2019; Xie et al., 2021). Within this context, Tables 8 and 9 show that the framework was able to achieve a high performance of the IR algorithm, ranging between 87% and 99% in both use cases, depending on the length of the user query and database size. This result is not only comparable to the 88.1% of accuracy, achieved by Wang et al. (Wang, Wang, et al., 2019) using their Knowledge-aware Path Recurrent Network (KPRN), but adds to it the ability of considering domain requirements in the intelligent IR algorithm. Unlike the domain-intrinsic nature of the KPRN, the proposed framework provides the same or higher accuracy in domain restricted environments, such as the ones demonstrated in the previous two use case.

The explainability algorithm reached 85.6% in one use case and ranged between 82.5% and 99.6% depending on the length of the search query. In both cases the *MEP* score is comparable or surpasses the results of Chen and Miyazaki (Y. Chen & Miyazaki, 2020), who achieve 87.25% *MEP* score from a similar explainability algorithm implemented on a Slack-Item-KG of 100k items. It should be also noted in this comparison, that the framework’s main contribution is not achieving the highest accuracy

levels of the IR and explainability algorithms, but rather achieving accurate results in a domain-specific environment, with a higher level of restrictions and data complexity than the one in the general-purpose datasets used in (Y. Chen & Miyazaki, 2020; Song et al., 2019; Wang, Wang, et al., 2019; Xie et al., 2021). The ability of the solution to consider those requirements and achieve high accuracy scores of the IR and explainability at the same time, is the value-added that the framework offers in comparison to the current state-of-the-art. In Table 9, the comparative *MEP* scores of the proposed framework and other similar solutions are summarized.

It can be noticed from Table 9 that the *MEP* score increases with the increase of data size. This increase can be in the dataset itself, such as the case of using Slack-Item-KG 1M instead of 100K, or in the search query, such as the difference in the first use case between long and short queries. The enhanced potential for generating explanations from larger sizes of data is due to the increased potential to find relations and paths in the KG, which correspond to the user query. With more information in the search query, or larger knowledge graphs, more paths can be extracted and used for constructing the result explanation. Here, the explainability algorithm in the proposed approach is mainly compared to the work of Chen and Miyazaki, since the authors explicitly use the *MEP* measure to evaluate their work. The evaluation is then extended to include *BLEU* and *Rouge-L* scores, as illustrated in Figures 45, 46, and 51.

Table 9. Explainability comparison based on the *MEP* scores.

Approach	MEP score %
Chen and Miyazaki - Slack-Item-KG 100k	87.25
Chen and Miyazaki - Slack-Item-KG 1M	91.67
Our framework - Use case1 (short query)	82.5
Our framework - Use case1 (long query)	99.6
Our framework - Use case2	85.6

The evaluation of the proposed framework was implemented on two use cases that have different domain requirements, to demonstrate the transferability of the solution. In general, the proposed framework is designed in a generic way that adopts to any domain of interest. The implementation of the framework in other domains, such as the medical (Naiseh, 2020), energy (H. Li et al., 2021) or educational (Abu-Rasheed et al., 2018), can be similarly accomplished, as long as the domain requirements are defined by the experts and data sources. The IR and explainability algorithms are then transferrable amongst those domains, due to their dependence on the KG structure. This fact implies the sustainability of the proposed framework, when used in a certain domain and then transferred to another one. Another sustainability aspect of the framework comes from the potential to develop and extend the knowledge graph itself, during the lifetime and therefore ongoing extension of the industrial process that is supported by the framework and its underlying data. Since the proposed framework includes the necessary elements to construct the knowledge graph, it inherently enables its expansion and development. This

allows for a long-term use of the framework, which can take into account new data, new domain requirements, and new user queries.

Our framework has been developed to handle textual data sources. This fact influenced the nature of domain-specific and domain-agnostic components used in the framework. That being said, the limitation of the framework to include image data sources has to be addressed, since it lacks the image-processing components on the input part of the framework. Although the current framework excels in handling pure textual data sources, and imagery data that is well annotated, it can still be extended to include specific elements that handle pure visual data sources, to embed them in the knowledge graph construction process. Once the image data is embedded in the KG, the IR and explainability algorithms will be able to query this data types directly, corresponding the same user queries.

8.4 Conversational explainability platform

8.4.1 Experiment 6: User acceptance in nursing VET

The inclusion of explainability in the VET user study played a role in improving learner engagement, reducing resistance, and fostering reflective learning. Explainability was designed to provide learners with clear and transparent insights into why specific learning recommendations were made. By offering contextual guides, explanatory prompts, and clear reasoning behind algorithmic decisions, this feature served as a supportive mechanism, empowering learners both emotionally and cognitively.

Reduction in resistance

One of the most noticeable effects of explainability was its capacity to lower resistance among learners. In groups receiving correct recommendations, explainability enhanced confidence in the learning process, as participants felt more secure in the system's decision-making. This transparency encouraged them to trust the learning pathways and alleviated any defensiveness or hesitation.

Even in cases where learners were subjected to reversed recommendations, intentionally misaligned sequences designed to test adaptability, explainability minimized frustration by offering clarity. Learners were able to interpret the misalignment as an intentional and meaningful exploration rather than a technical error. This reframing helped maintain constructive engagement despite the disruptive nature of the reversed sequencing.

For groups that did not receive algorithmic recommendations at all, explainability played also a role. The explanatory features allowed learners to independently interpret and navigate the material, creating a sense of autonomy that significantly reduced resistance.

Increased motivation and engagement

Explainability also had an effect on motivating learners to actively engage with the educational material. By providing clear reasoning for the algorithm's choices, it fostered trust in the system. Learners became more confident in their ability to progress meaningfully, regardless of whether their recommendations were aligned or reversed. This trust transformed into a willingness to explore even challenging or unfamiliar topics, demonstrating the motivational potential of transparent communication in algorithm-assisted systems.

In groups without recommendations, explainability became a source of empowerment. Instead of relying on pre-determined pathways, learners were guided by explanatory prompts that allowed them to make informed, independent decisions. This sense of control and autonomy further enhanced their engagement and interest in the learning material.

Enhanced cognitive reflection and processing

Explainability did not only influence engagement, but also contributed to deeper cognitive processing and reflective ability. Across all groups, those with explainability demonstrated a stronger ability to engage in reflective learning. By understanding the rationale behind their path recommendations, participants were better equipped to connect the content to their own experiences and draw meaningful insights from the material.

In the reversed recommendation groups, explainability showed a unique role. While the misaligned sequencing could have caused confusion and disengagement, the explanation was able to clearly reduce learners' resistance towards the recommendation. Instead of becoming frustrated, they treated these disruptions as experimental exercises and were able to make better decisions about following the recommendation.

Explainability also showed an effect for groups without recommendations. These learners, who lacked the scaffold of algorithmic guidance, relied heavily on the contextual input provided by explanation to direct their progress. While their reflective capacities did not reach the same levels as those with correct recommendations, explainability ensured that they still achieved meaningful engagement and avoided surface-level learning. Figure 52 shows the levels of reflection and resistance measured in the test.

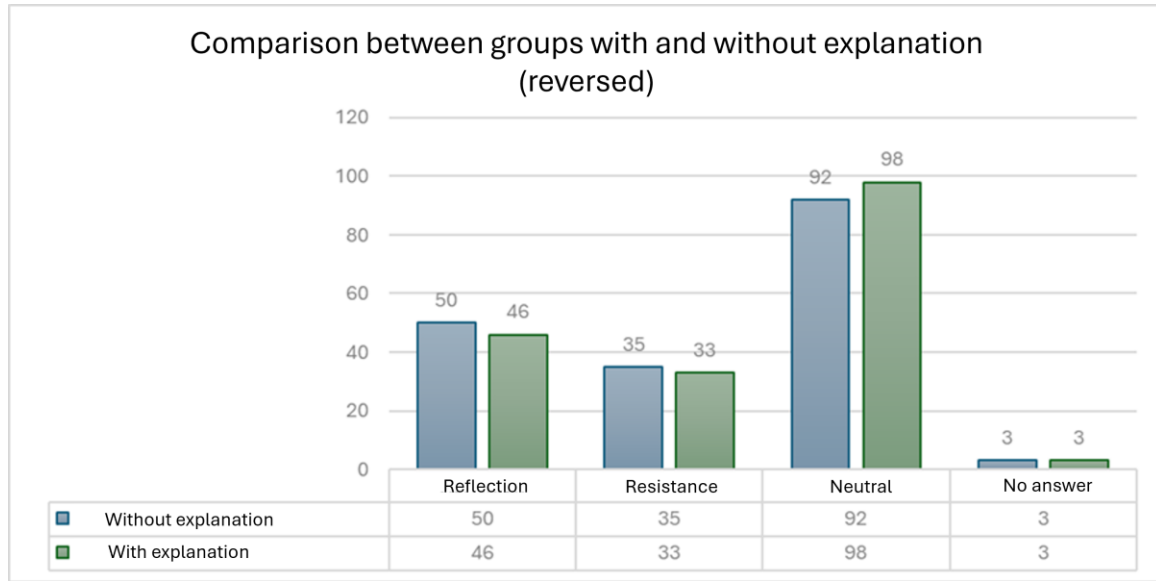


Figure 52. Groups 2 and 4 (reversed) with recommendation, comparison with and without explanations (Abu-Rasheed et al., 2024).

8.4.2 Experiment 7: higher education user studies

To evaluate the conversational explainability approach through the chatbot introduced in chapter 7, this experiment is designed to test the performance of the proposed intent classifier. Then, a user study is done to evaluate the features of the explainability chatbot. For the intent classification test, a set of 182 user requests is collected and labeled. They were spread equally over the 7 supported intent categories. Figure 53 shows the confusion matrix of the classifier's performance. The classifier has reached an accuracy level of 88% over all classes. Table 10 **Error! Reference source not found.** shows the values for the precision, recall, and F1 measure calculated per class. From the performance scores, one can notice that the classifier mostly confuses the questions about the recommendation reasoning with the questions from the third and fourth classes. By analyzing user queries in the test, it was noticed that several queries about the benefit of a learning material (class 3) and the relations to other learning materials (class 4) were phrased in a way that demands a "justification" of the benefit or the material connections. This may explain the reason why the classifier considered these as questions about the reasoning that justifies the whole recommendation and classified them under class 1.

To evaluate the chatbot features, a preliminary user study is conducted, which is intended as a proof-of-concept, before running a larger-scale evaluation. Participants in the experiment were surveyed to measure their:

1. satisfaction with the chatbot design.
2. satisfaction with the quality of its answers.
3. perception of the correctness of the chatbot answers based on the user intent.
4. satisfaction with the speed of responses.

5. perception of the chatbot’s responses to out-of-scope questions, i.e., in the “other” intent class

For points 2-5, a set of eight scenarios is designed, depicted by short stories. Six scenarios require the user to validate each of the intent classes, except class 6 that covers the relation to the user’s daily work. This is because the same recommendation was offered to all users, to avoid introducing a new dependent variable to the experiment through personalization. The seventh scenario validates the chatbot answers in cases it is not able to generate a response. The eighth scenario is meant to validate the mentor-chat session. Likert scales (1-5) and (1-10) were used to record the user answers.

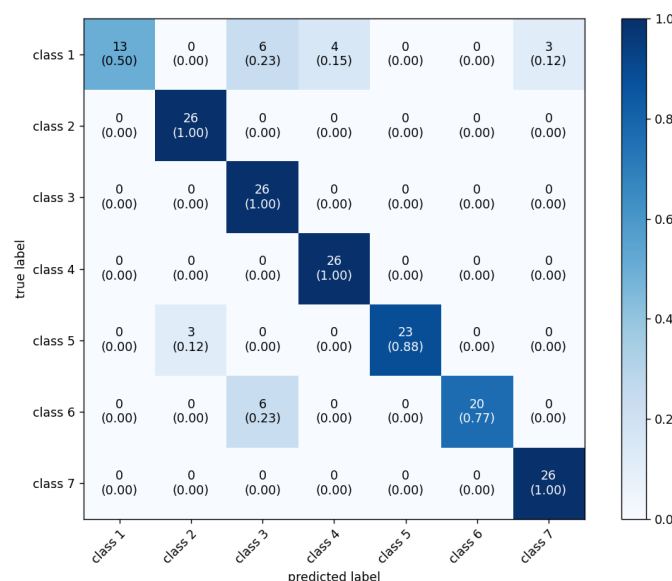


Figure 53. Confusion matrix for the intent classifier (Abu-Rasheed et al., 2024).

This qualitative experiment is conducted with a small sample of nine participants from the target group. The participants were from the academic field and included one post-doctorate, two PhD candidates, and six graduate students. The test duration ranged between 60 and 90 minutes, during which one of the authors took the role of the mentor, due to their knowledge about the recommended path.

User evaluation of the chatbot’s design and menus, see Figure 54, got an average of 4.7/5. The quality of answers was evaluated with an average of 4.4/5. The speed of response scored an average of 4.6/5. It is important to mention that the speed of responses is dependent on different factors, including the specifications of the servers hosting the chatbot, the speed of the KG search and information retrieval algorithm, and the response time from the API. While there is no control over the API’s response time, the speed of the KG search algorithm is measured, showing a response time that ranges between 0.72 and 0.9 seconds. The server used for the experiment was equipped with a CPU 13th Gen Intel[®] Core[™] i7-13700KF, 3.40 GHz, and 32.0 GB of RAM. The response time recorded for the complete process, from the user query to the final answer, ranged between four and eight seconds.

Figure 55 shows the results of the eight test scenarios. It can be noticed that the scenarios that cover cases in which the chatbot was allowed to provide an answer received a higher score than the two scenarios (6 and 7), for which the chatbot should avoid providing an answer, but rather re-direct the user to another conversation state, e.g., the fallback state of contacting a mentor.

Table 10. Precision, Recall, and F1 measure for each intent class

Class ID	Class description	P	R	F1
1	Query about the reason behind the recommendation.	0.50	1	0.67
2	Query about the content on the recommendation page in the web application.	0.89	1	0.94
3	Query about the benefits that will be gained from learning a certain learning material in the recommended path.	0.68	1	0.81
4	Query about the relations and similarities between recommended materials and those in the KG.	0.86	1	0.92
5	Query for additional information about the recommended materials.	1	0.88	0.94
6	Query about the relation of recommended materials to the student's context (e.g., their daily work).	1	0.77	0.87
7	Other queries	0.89	1	0.94

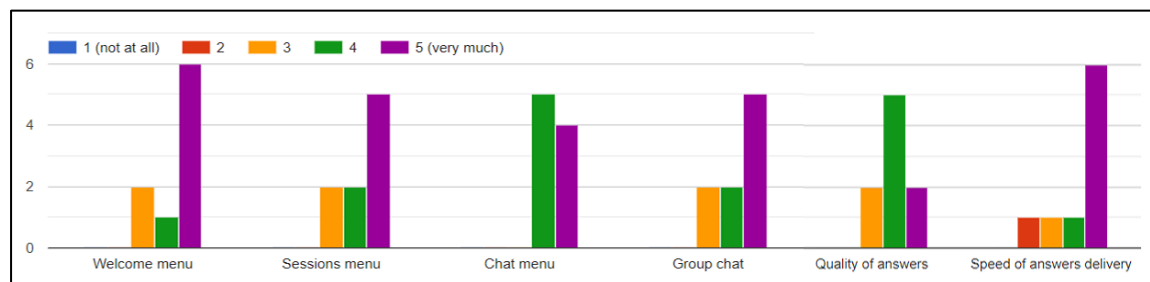


Figure 54. User evaluation of the chatbot menus, group chat, answer quality, and speed (Abu-Rasheed et al., 2024).

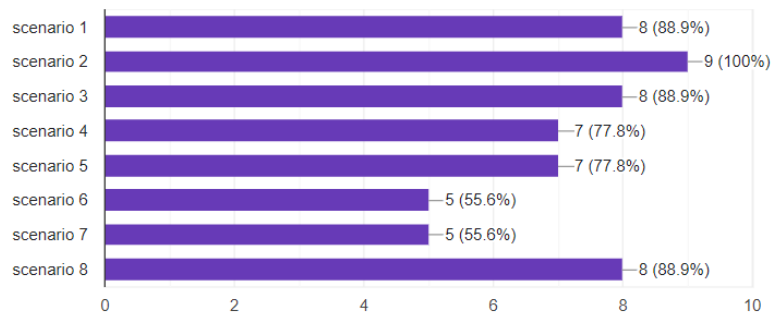


Figure 55. Average user evaluation of the chatbot responses for each of the eight scenarios (Abu-Rasheed et al., 2024).

Discussion and limitations

While the evaluation above provides a preliminary proof-of-concept for the proposed approach, one cannot argue that it reflects statistically sufficient evidence that the chatbot has an influence on the learning performance of a student. A larger and more thorough test process should therefore be designed within to test the effect of the chatbot on students in a longer-term learning setting. This follows the growing concerns, which the authors share, that the rapid development of Gen-AI and the approaches based on it are not accompanied by the same level of real-world testing of the learning impact that those approaches have in real educational settings. From the user study and the involvement of domain experts in the design and evaluation of the chatbot, evidence for several lessons learned is found, which is summarized as follows:

- The student utilization of the chatbot is greatly influenced by the way they phrase their questions. Even when intent classifiers are highly accurate, different students may still use very similar sentences and mean different things. LLMs are one solution to support intent classification. However, they require thorough contextualization to understand the sentence's meaning.
- LLM outputs are mostly used exactly as the models generate them. However, using parts of the output, or arranging several partial outputs in the pre-defined slots of a larger explanation template offers educators more flexibility for controlling the final explanation content.
- Quantitative evaluation of LLM responses does not necessarily reflect its quality for an education use case. Pedagogy experts pointed out in some of the interviews that several LLM-generated texts were not wrong, but they did not offer a high value-added to students. One expert expressed this idea as: *"[the chatbot response] is not wrong. It is quite fine by me. But I wouldn't give this answer to my students because it doesn't enable them to reflect. [...] Reflection needs additional information, which is not simply an answer to the question."*
- The context added to an LLM query may be phrased in different ways, even when the same information is included. The evaluation of the phrasing's effect is important, but it presents a complex challenge. In this system, the context phrasing is

adjusted for each intent, to keep the overall prompt objective -to the best of our ability- and to keep the LLM's focus on the task itself, without being distracted by the contextual information that may not represent the user intent.

The latter point is a limitation of this study since it requires another test for the effect of context phrasing. Other limitations that are addressed for the study include:

- The limited sample size of the user study. Therefore, this evaluation is mainly considered qualitative, which is meant to prove the proposed concept. A larger scale test is required to acquire statistically significant results, and to focus closer on the real effect of using the chatbot explanations in learning. For that purpose, an A/B test needs to be designed to compare the effect of conversation explainability against single-step textual explanation modality.
- This experiment depended mainly on GPT-4 model. A comparison is needed to evaluate the results and their potential differences when using other LLMs.
- The context in this experiment does not include user-profile data, which is meant to comply with the GDPR, since an API from a third party is being used. This presented a limitation for the information used in the evaluation. To solve this issue, local LLMs can be used to enrich the context with user-specific data from the profiles and study the role of this type of information on the LLM's response.

Despite those limitations, the concept proposed for the conversational explainability showed a promising potential for enhancing the learner's agency over the learning recommendations. This approach also highlighted the role that Gen-AI can play even when its tasks are restricted to comply with the pedagogical requirements. As mediators with concretely defined tasks and boundaries, intelligent conversational agents are able to provide learners and educators with engaging, communicative, and highly-available explanations and further input on the learning recommendations.

9 Conclusion

Human learning is inherently complex and highly context-dependent, shaped by individual learner characteristics, diverse learning goals, varying learning environments, and multiple stakeholder demands. The core research problem in this thesis addressed, therefore, the intersection of personalized education, context-aware recommendation systems (CARS), and explainable artificial intelligence (XAI), focusing on vocational education and training (VET) settings. This research orientation is motivated by the increasing demand for personalization in learning, particularly in vocational education and training (VET), where learners' needs and contexts are highly heterogeneous. While existing context-aware recommender systems (CARS) have made significant progress, they have not adequately captured the structural complexity of learning contexts nor addressed the pressing need for transparent decision-making mechanisms that are interpretable and pedagogically relevant. This thesis sought to bridge these gaps through a proposing CARExKG: a method for integrative generation of context-aware recommendations and explainability through knowledge graphs. This knowledge graph-based method is designed not only to capture contextual variables and their relationships, but also to support dynamic reasoning and explanation generation from pedagogical and user-centric perspectives.

The research was guided by three central research questions, each aimed at resolving a key component of the broader problem. The first research question (RQ1) concerned the structural and algorithmic integration of contextual learning factors to enhance the personalization and adaptability of learning path recommendations. In response, the thesis investigated the definitions of learning context and its factors thoroughly, and proposed a knowledge graph representation of the learning context, capturing the interrelations among diverse contextual factors, such as learner attributes, goals, learning materials, and learning-environment variables. Algorithmically, a Markov decision process (MDP) was implemented over the graph, with a novel context-aware reward function. This function allowed the system to dynamically assess and weigh different learning paths based on the context-sensitive value of state transitions, thereby enabling personalized recommendations that adapt to the learner's current situation and objectives.

The second research question (RQ2) focused on the effectiveness of tailoring learning path recommendations through graph-based algorithms in complex educational contexts. The proposed MDP-based algorithm demonstrated its ability to account for multi-dimensional contextual data in the recommendation process. By utilizing the semantic richness of the knowledge graph and adjusting recommendation trajectories through the context-sensitive reward function, the system successfully generated recommendations that reflected the specific constraints, goals, and situational parameters of individual learners. The results from the evaluation, which was conducted in a real-world scenario involving nursing staff in elderly care facilities, showed that the recommendations led to improved learner performance and alignment with professional development needs, validating the system's capacity for context-sensitive personalization.

The third research question (RQ3) addressed the design of an explainability framework that reduces learner resistance and enhances their engagement with the recommended learning paths. This aspect of the research recognized that technically accurate recommendations may still be disregarded if they are not relevant or understood by learners. In response, the thesis introduced an explainability layer that utilized large language models and domain expertise to generate human-readable, pedagogically meaningful explanations for each recommended path. These explanations clarified the rationale behind the selection of particular learning items in the learning path, and the relevance of their sequencing, grounded in both the learner's prior knowledge and contextual circumstances. In this sense, the explanations were not only technical or generic but rather contextualized, reflecting the pedagogical logic as shaped by expert input and the domain requirements. The evaluation results revealed that learners perceived the explanations as clear, relevant, and supportive of their decision-making process. Moreover, the inclusion of explanations contributed to increased acceptance of the recommendations and a stronger sense of agency in managing one's learning experience.

When considered holistically, the thesis provides a unifying method that answers the three research questions not only in isolation but in a mutually reinforcing manner. The structural representation of context in the knowledge graph (RQ1) supports the algorithmic reasoning required for contextual recommendation (RQ2), and both elements supply the necessary semantic foundation for generating tailored explanations (RQ3). The integration of these components into a single method (CARExKG) demonstrates the feasibility and efficacy of designing context-aware, explainable learning recommender systems capable of addressing real-world educational challenges. The method's deployment in a complex vocational learning setting further underscores the relevance and scalability of the proposed approach.

The CARExKG method offers a promising model for future developments in personalized and explainable learning systems. As educational settings increasingly adopt AI-based technologies, the insights, frameworks, data structure, and algorithms presented

in this work provide a robust foundation for designing systems that are not only intelligent but also transparent, ethical, pedagogy-oriented, and learner-centered.

9.1 Future work

Future research may extend this work by exploring its applicability and potential in new learning environments, such as self-regulated learning, and distance learning. Considering the pedagogical requirements in such learning settings, and the important role of the learning environment itself in such settings, the proposed CARExKG method can be examined and further fine-tuned to account for unique their requirements.

Additionally, it is important to consider long-term evaluation for assessing the effects of educational technologies, especially the ones integrating new technologies, such as Gen-AI and LLMs. While the experiments, quantitative and qualitative, in the is thesis where limited by the time- and personal-resources for conducting long-term evaluation, it is therefore an important following step to run experiments with systems based on the proposed method for several study semesters, or several training months. This is to ensure that the long-term effect of the method aligns with the promising results from the evaluation rounds and experiments conducted in the timeframe of this thesis.

Moreover, as the method focuses on contextualizing learning path recommendations and their explainability, the integration of the methos with other educational-process steps, including the learning assessment and automated feedback generation, is highly recommended to extend the feature and functions of CARExKG with new ones related to the assessment and feedback tasks. This, in turn, aims for covering more education functions and processes comprehensively and based on the solid structural and algorithmic foundations that CARExKG provides.

List of Publications

Journal Publications

- Abu-Rasheed, H., Weber, C., & Fathi, M. (2023). Context based learning: A survey of contextual indicators for personalized and adaptive learning recommendations – a pedagogical and technical perspective. *Frontiers in Education*, 8, 1210968. <https://doi.org/10.3389/feduc.2023.1210968>
- Abu-Rasheed, H., Weber, C., Zenkert, J., Dornhöfer, M., & Fathi, M. (2022). Transferrable Framework Based on Knowledge Graphs for Generating Explainable Results in Domain-Specific, Intelligent Information Retrieval. *Informatics*, 9(1), Article 1. <https://doi.org/10.3390/informatics9010006>
- Zenkert, J., Weber, C., Dornhöfer, M., Abu-Rasheed, H., & Fathi, M. (2021). Knowledge Integration in Smart Factories. *Encyclopedia*, 1(3), Article 3. <https://doi.org/10.3390/encyclopedia1030061>
- Ilkou, E., Abu-Rasheed, H., Chaves-Fraga, D., Engelbrecht, E., Jiménez-Ruiz, E., & Labra-Gayo, J. E. (2025). Teaching knowledge graph for knowledge graphs education. *Semantic Web* (Submitted - Under Review). <https://www.semantic-web-journal.net/content/teaching-knowledge-graph-knowledge-graphs-education>

Conference Publications

- Abu-Rasheed, H., Jumbo, C., Amin, R. A., Weber, C., Wiese, V., Obermaisser, R., & Fathi, M. (2025). LLM-Assisted Knowledge Graph Completion for Curriculum and Domain Modelling in Personalized Higher Education Recommendations (No. arXiv:2501.12300). *arXiv*. <https://doi.org/10.48550/arXiv.2501.12300>
- Abu-Rasheed, H., Weber, C., & Fathi, M. (2025). Tell Me Why I Should “Not” Follow Your Recommendation: On the Role of Explainable AI in Collaborative Human-AI Decision Making. *Open Science Framework*. <https://doi.org/10.31219/osf.io/x867z>
- Abu-Rasheed, H., Abdulsalam, M. H., Weber, C., & Fathi, M. (2024). Supporting Student Decisions on Learning Recommendations: An LLM-Based Chatbot with Knowledge Graph Contextualization for Conversational Explainability and Mentoring. In the First international workshop on Generative AI for Learning Analytics (GenAI-LA), in the 14th International Conference on Learning Analytics and Knowledge (LAK24), *CEUR Workshop Proc.* (Vol. 3667, pp. 230–239).
- Abu-Rasheed, H., Nadeem, M., Dornhöfer, M., Zenkert, J., Weber, C., & Fathi, M. (2024). TExKG in Health Domain: The Application of Knowledge Graph Based

- Framework for Explainable Recommendations in the Contexts of Elderly Care, Mental Health, and Emergency Responses. In M.-R. Alam & M. Fathi (Eds.), *Integrated Systems: Data Driven Engineering* (pp. 265–285). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-53652-6_16
- Abu-Rasheed, H., Ikeda, R., Ferriyan, A., Weber, C., Fathi, M., Okawa, K., & Thamrin, A. H. (2024). Problem-Based Learning-Path Recommendations Through Integrating Knowledge Graphs and Large Language Models. <https://doi.org/10.35542/osf.io/w8kdt>
- Abu-Rasheed, H., Weber, C., & Fathi, M. (2024). Experimental Interface for Multimodal and Large Language Model Based Explanations of Educational Recommender Systems (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2402.07910>
- Abu-Rasheed, H., Weber, C., & Fathi, M. (2024). Knowledge Graphs as Context Sources for LLM-Based Explanations of Learning Recommendations. 2024 IEEE Global Engineering Education Conference (EDUCON), 1–5. <https://doi.org/10.1109/EDUCON60312.2024.10578654>
- Buchmann, U., Dornhöfer, M., Fathi, M., Kismihók, G., Köhler, S., Abu-Rasheed, H., Schröder, M., & Weber, C. (2024). KI-basierter digitaler Weiterbildungsraum WBsmart. *Digitalisierung in Den Gesundheitsberufen*, 214.
- Abu-Rasheed, H., Efthymiou, Y., Fathi, M., Ghadamighalandari, P., Medina, J. L., García, C. O., Tsardanidis, G., Zenkert, J., & Zgeras, G. (2023). Supporting Remote Students Through Utilizing Web-Based Exercise-Templates and a Mobile Learning Chatbot for Creating and Interacting with Learning Materials. In O. Viberg, I. Jivet, P. J. Muñoz-Merino, M. Perifanou, & T. Papathoma (Eds.), *Responsive and Sustainable Educational Futures* (Vol. 14200, pp. 668–673). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-42682-7_59
- Abu-Rasheed, H., Weber, C., Dornhöfer, M., & Fathi, M. (2023). Pedagogically-Informed Implementation of Reinforcement Learning on Knowledge Graphs for Context-Aware Learning Recommendations. In O. Viberg, I. Jivet, P. J. Muñoz-Merino, M. Perifanou, & T. Papathoma (Eds.), *Responsive and Sustainable Educational Futures* (Vol. 14200, pp. 518–523). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-42682-7_35
- Abu-Rasheed, H., Weber, C., Zenkert, J., Krumm, R., & Fathi, M. (2022). Explainable Graph-Based Search for Lessons-Learned Documents in the Semiconductor Industry. In K. Arai (Ed.), *Intelligent Computing* (Vol. 283, pp. 1097–1106). Springer International Publishing. https://doi.org/10.1007/978-3-030-80119-9_73
- Reichow, I., Buntins, K., Paaßen, B., Abu-Rasheed, H., Weber, C., & Dornhöfer, M. (2022). Recommendersysteme in der beruflichen Weiterbildung. Grundlagen, Herausforderungen und Handlungsempfehlungen. Ein Dossier im Rahmen des

INVITE-Wettbewerbs. pedocs. <http://nbn-resolving.de/urn:nbn:de:0111-pedocs-245172>

- Weber, C., Abu-Rasheed, H., & Fathi, M. (2022). Adding Context to Industry 4.0 Analytics: A New Document Driven Knowledge Graph Construction and Contextualization Approach. 2022 IEEE International Conference on Electro Information Technology (eIT), 550–555. <https://doi.org/10.1109/eIT53891.2022.9813992>
- Abu Rasheed, H., Weber, C., Zenkert, J., Czerner, P., Krumm, R., & Fathi, M. (2021). A Text Extraction-Based Smart Knowledge Graph Composition for Integrating Lessons Learned During the Microchip Design. In K. Arai, S. Kapoor, & R. Bhatia (Eds.), *Intelligent Systems and Applications* (pp. 594–610). Springer International Publishing. https://doi.org/10.1007/978-3-030-55187-2_43
- Upadhyay, C., Abu-Rasheed, H., Weber, C., & Fathi, M. (2021). Explainable Job-Posting Recommendations Using Knowledge Graphs and Named Entity Recognition. 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 3291–3296. <https://doi.org/10.1109/SMC52423.2021.9658757>
- Ilkou, E., Abu-Rasheed, H., Tavakoli, M., Hakimov, S., Kismihók, G., Auer, S., & Nejd, W. (2021). EduCOR: An Educational and Career-Oriented Recommendation Ontology. In A. Hotho, E. Blomqvist, S. Dietze, A. Fokoue, Y. Ding, P. Barnaghi, A. Haller, M. Dragoni, & H. Alani (Eds.), *The Semantic Web – ISWC 2021* (Vol. 12922, pp. 546–562). Springer International Publishing. https://doi.org/10.1007/978-3-030-88361-4_32
- Tomar, S., Abu Rasheed, H., & Fathi, M. (2020). Educational Multimodal Data Mining And Fusion Through Knowledge Graphs For Topic-Relation Extraction In Study Recommendations. 3696–3705. <https://doi.org/10.21125/edulearn.2020.1020>
- Abu-Rasheed, H., Zenkert, J., Weber, C., Dornhöfer, M., Klahold, A., & Fathi, M. (2019). Language Learning Tool Based On Augmented Reality And The Concept For Imitating Mental Ability Of Word Association (Cimawa). 4118–4125. <https://doi.org/10.21125/edulearn.2019.1041>
- Abu-Rasheed, H., Zenkert, J., Weber, C., & Fathi, M. (2019). Conversational Chatbot System For Student Support In Administrative Exam Information. 8294–8301. <https://doi.org/10.21125/iceri.2019.1974>
- Ishaq, I., Jayousi, R., Odeh, S., Edwan, E., Shaheen, A., Elnaggar, M., Elagha, A., Salamah, S., Weber, C., Abu Rasheed, H., Kirner, R., Doolan, M., Ahmadian, H., Obermaisser, R., Abou-Tair, D. el D. I., Khalifeh, A., Alounch, S., Alhalhouli, Z., Alemerien, K., ... Saleh, M. (2019). Work in Progress – Establishing a Master Program in Cyber Physical Systems: Basic Findings and Future Perspectives.

2019 International Conference on Promising Electronic Technologies (ICPET), 4–9. <https://doi.org/10.1109/ICPET.2019.00009>

Abu Rasheed, H., Weber, C., Harrison, S., Zenkert, J., & Fathi, M. (2018). What To Learn Next: Incorporating Student, Teacher And Domain Preferences For A Comparative Educational Recommender System. 6790–6800. <https://doi.org/10.21125/edulearn.2018.1610>

Abu-Rasheed, H. (2018). Hybrid Educational Recommender System, Utilizing Student, Teacher and Domain Preferences and Amazon Alexa Conversation Engine. <https://doi.org/10.13140/RG.2.2.22897.63847/1>

Bibliography

- Abdi, S., Khosravi, H., Sadiq, S., & Darvishi, A. (2021). Open Learner Models for Multi-activity Educational Systems. In I. Roll, D. McNamara, S. Sosnovsky, R. Luckin, & V. Dimitrova (Eds.), *Artificial Intelligence in Education* (pp. 11–17). Springer International Publishing. https://doi.org/10.1007/978-3-030-78270-2_2
- Abdollahi, B., & Nasraoui, O. (2016). Explainable Matrix Factorization for Collaborative Filtering. *Proceedings of the 25th International Conference Companion on World Wide Web - WWW '16 Companion*, 5–6. <https://doi.org/10.1145/2872518.2889405>
- Abu Rasheed, H., Weber, C., Zenkert, J., Czerner, P., Krumm, R., & Fathi, M. (2021). A Text Extraction-Based Smart Knowledge Graph Composition for Integrating Lessons Learned During the Microchip Design. In K. Arai, S. Kapoor, & R. Bhatia (Eds.), *Intelligent Systems and Applications* (pp. 594–610). Springer International Publishing. https://doi.org/10.1007/978-3-030-55187-2_43
- Abu-Rasheed, H., Abdulsalam, M. H., Weber, C., & Fathi, M. (2024). Supporting Student Decisions on Learning Recommendations: An LLM-Based Chatbot with Knowledge Graph Contextualization for Conversational Explainability and Mentoring. In the First international workshop on Generative AI for Learning Analytics (GenAI-LA), in the 14th International Conference on Learning Analytics and Knowledge (LAK24), CEUR Workshop Proc. (Vol. 3667, pp. 230–239).
- Abu-Rasheed, H., Dornhöfer, M., Weber, C., Kismihók, G., Buchmann, U., & Fathi, M. (2023). Building Contextual Knowledge Graphs for Personalized Learning Recommendations Using Text Mining and Semantic Graph Completion. *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)*, 36–40. <https://doi.org/10.1109/ICALT58122.2023.00016>
- Abu-Rasheed, H., Ikeda, R., Ferriyan, A., Weber, C., Fathi, M., Okawa, K., & Thamrin, A. H. (2024). *Problem-Based Learning-Path Recommendations Through Integrating Knowledge Graphs and Large Language Models*. <https://doi.org/10.35542/osf.io/w8kdt>
- Abu-Rasheed, H., Köhler, S., Schröder, M., Weber, C., Buchmann, U., Kismihók, G., & Fathi, M. (2024). *Bildungswissenschaftliche Grundlegung eines smarten KI-basierten digitalen Weiterbildungsraums für die Altenhilfe mit-tels personalisierter Empfehlungssysteme* (BMBF-Innovationswettbewerb INVITE (Digitale Plattform berufliche Weiterbildung)). University of Siegen.
- Abu-Rasheed, H., Nadeem, M., Dornhöfer, M., Zenkert, J., Weber, C., & Fathi, M. (2024). TExKG in Health Domain: The Application of Knowledge Graph Based Framework for Explainable Recommendations in the Contexts of Elderly Care, Mental Health, and Emergency Responses. In M.-R. Alam & M. Fathi (Eds.), *Integrated Systems: Data Driven Engineering* (pp. 265–285). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-53652-6_16
- Abu-Rasheed, H., Weber, C., Dornhöfer, M., & Fathi, M. (2023). Pedagogically-Informed Implementation of Reinforcement Learning on Knowledge Graphs for Context-Aware Learning Recommendations. In O. Viberg, I. Jivet, P. J. Muñoz-Merino, M. Perifanou, & T. Papathoma (Eds.), *Responsive and Sustainable*

- Educational Futures* (Vol. 14200, pp. 518–523). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-42682-7_35
- Abu-Rasheed, H., Weber, C., & Fathi, M. (2023). Context based learning: A survey of contextual indicators for personalized and adaptive learning recommendations—a pedagogical and technical perspective. *Frontiers in Education*, 8, 1210968. <https://www.frontiersin.org/articles/10.3389/feduc.2023.1210968/full>
- Abu-Rasheed, H., Weber, C., & Fathi, M. (2024a). Experimental Interface for Multimodal and Large Language Model Based Explanations of Educational Recommender Systems. (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2402.07910>
- Abu-Rasheed, H., Weber, C., & Fathi, M. (2024b). Knowledge Graphs as Context Sources for LLM-Based Explanations of Learning Recommendations. *2024 IEEE Global Engineering Education Conference (EDUCON)*, 1–5. <https://doi.org/10.1109/EDUCON60312.2024.10578654>
- Abu-Rasheed, H., Weber, C., Harrison, S., Zenkert, J., & Fathi, M. (2018). What to Learn Next: Incorporating Student, Teacher and Domain Preferences for a Comparative Educational Recommender System. *EduLEARN19 Proceedings*, 6790–6800. <https://doi.org/10.21125/edulearn.2018.1610>
- Abu-Rasheed, H., Weber, C., Zenkert, J., Dornhöfer, M., & Fathi, M. (2022). Transferrable framework based on knowledge graphs for generating explainable results in domain-specific, intelligent information retrieval. *Informatics*, 9(1), 6. <https://www.mdpi.com/2227-9709/9/1/6>
- Abu-Rasheed, H., Weber, C., Zenkert, J., Krumm, R., & Fathi, M. (2022). Explainable Graph-Based Search for Lessons-Learned Documents in the Semiconductor Industry. In K. Arai (Ed.), *Intelligent Computing* (Vol. 283, pp. 1097–1106). Springer International Publishing. https://doi.org/10.1007/978-3-030-80119-9_73
- Abu-Salih, B., & Alotaibi, S. (2024). A systematic literature review of knowledge graph construction and application in education. *Heliyon*, 10(3), e25383. <https://doi.org/10.1016/j.heliyon.2024.e25383>
- Adomavicius, G., Bauman, K., Mobasher, B., Ricci, F., Tuzhilin, A., & Unger, M. (2020). Workshop on Context-Aware Recommender Systems. *Fourteenth ACM Conference on Recommender Systems*, 635–637. <https://doi.org/10.1145/3383313.3411533>
- Adomavicius, G., Bauman, K., Tuzhilin, A., & Unger, M. (2022). Context-Aware Recommender Systems: From Foundations to Recent DevelopmentsContext-aware recommender systems. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender Systems Handbook* (pp. 211–250). Springer US. https://doi.org/10.1007/978-1-0716-2197-4_6
- Adomavicius, G., & Tuzhilin, A. (2015). Context-Aware Recommender Systems. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender Systems Handbook* (pp. 191–226). Springer US. https://doi.org/10.1007/978-1-4899-7637-6_6
- Afchar, D., Melchiorre, A. B., Schedl, M., Hennequin, R., Epure, E. V., & Moussallam, M. (2022). Explainability in music recommender systems. *AI Magazine*, 43(2), 190–208. <https://doi.org/10.1002/aaai.12056>

- Aguiar, E., Ambrose, G. A. A., Chawla, N. V., Goodrich, V., & Brockman, J. (2014). Engagement vs Performance: Using Electronic Portfolios to Predict First Semester Engineering Student Persistence. *Journal of Learning Analytics*, 1(3), Article 3. <https://doi.org/10.18608/jla.2014.13.3>
- Ahmad, A., Schneider, J., & Drachsler, H. (2020). OpenLAIR an Open Learning Analytics Indicator Repository Dashboard. In C. Alario-Hoyos, M. J. Rodríguez-Triana, M. Scheffel, I. Arnedillo-Sánchez, & S. M. Dennerlein (Eds.), *Addressing Global Challenges and Quality Education* (pp. 467–471). Springer International Publishing. https://doi.org/10.1007/978-3-030-57717-9_46
- Aldowah, H., Al-Samarraie, H., & Ghazal, S. (2019). How Course, Contextual, and Technological Challenges are Associated With Instructors' Individual Challenges to Successfully Implement E-Learning: A Developing Country Perspective. *IEEE Access*, 7, 48792–48806. <https://doi.org/10.1109/ACCESS.2019.2910148>
- Alnuaim, A., Caleb-Solly, P., & Perry, C. (2016). Enhancing student learning of human-computer interaction using a contextual mobile application. *2016 SAI Computing Conference (SAI)*, 952–959. <https://doi.org/10.1109/SAI.2016.7556095>
- Alwarthan, S., Aslam, N., & Khan, I. U. (2022). An Explainable Model for Identifying At-Risk Student at Higher Education. *IEEE Access*, 10, 107649–107668. <https://doi.org/10.1109/ACCESS.2022.3211070>
- Amasha, M. A., Areed, M. F., Alkhalaf, S., Abougalala, R. A., Elatawy, S. M., & Khairy, D. (2020). The future of using Internet of Things (IoT) and Context-Aware Technology in E-learning. *Proceedings of the 2020 9th International Conference on Educational and Information Technology*, 114–123. <https://doi.org/10.1145/3383923.3383970>
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84, 261–271. <https://doi.org/10.1037/0022-0663.84.3.261>
- Anderson, J. R. (2000). *Learning and memory: An integrated approach, 2nd ed* (pp. xviii, 487). John Wiley & Sons Inc.
- Askarbekuly, N., & Luković, I. (2024). *Learning Outcomes, Assessment, and Evaluation in Educational Recommender Systems: A Systematic Review* (No. arXiv:2407.09500). arXiv. <https://doi.org/10.48550/arXiv.2407.09500>
- Aytekin, M. C., & Saygın, Y. (2024). ACE: AI-Assisted Construction of Educational Knowledge Graphs with Prerequisite Relations. *Journal of Educational Data Mining*, 16(2), Article 2. <https://doi.org/10.5281/zenodo.14250896>
- Baccari, S., & Neji, M. (2016). Design for a context-aware and collaborative mobile learning system. *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, 1–6. <https://doi.org/10.1109/ICCIC.2016.7919578>
- Barlovits, S., Caldeira, A., Fesakis, G., Jablonski, S., Koutsomanoli Filippaki, D., Lázaro, C., Ludwig, M., Mammana, M. F., Moura, A., Oehler, D.-X. K., Recio, T., Taranto, E., & Volika, S. (2022). Adaptive, Synchronous, and Mobile Online Education: Developing the ASYMPTOTE Learning Environment. *Mathematics*, 10(10), Article 10. <https://doi.org/10.3390/math10101628>

- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Barria-Pineda, J., Akhuseyinoglu, K., & Brusilovsky, P. (2019). Explaining Need-based Educational Recommendations Using Interactive Open Learner Models. *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, 273–277. <https://doi.org/10.1145/3314183.3323463>
- Barria-Pineda, J., & Brusilovsky, P. (2019). Explaining educational recommendations through a concept-level knowledge visualization. *Companion Proceedings of the 24th International Conference on Intelligent User Interfaces*, 103–104. <https://doi.org/10.1145/3308557.3308690>
- Bhardwaj, P. (2022). *Adversarial Robustness of Representation Learning for Knowledge Graphs* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2210.00122>
- Bicans, J. (2016). *Context data learner model for classroom and intelligent tutoring systems*.
- Bicans, J., & Grundspenkis, J. (2017). Student Learning Style Extraction from On-Campus Learning Context Data. *Procedia Computer Science*, 104, 272–278. <https://doi.org/10.1016/j.procs.2017.01.135>
- Bidarra, J., & Rusman, E. (2017). Towards a pedagogical model for science education: Bridging educational contexts through a blended learning approach. *Open Learning: The Journal of Open, Distance and e-Learning*, 32(1), 6–20. <https://doi.org/10.1080/02680513.2016.1265442>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability(Formerly: Journal of Personnel Evaluation in Education)*, 21(1), 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Blundell, R., Dearden, L., Meghir, C., & Sianesi, B. (1999). Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy. *Fiscal Studies*, 20(1), 1–23.
- Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems*, 46, 109–132. <https://doi.org/10.1016/j.knosys.2013.03.012>
- Borgnakke, K. (2015). Coming Back to Basic Concepts of the Context. *Seminar.Net*, 11(1). <https://doi.org/10.7577/seminar.2363>
- Bougsiaa, H. (2016). Teaching and Learning Context in Augmented Reality Environment. *Ars Educandi*, 13, 23–31. <https://doi.org/10.26881/ae.2016.13.02>
- Bouihi, B., & Bahaj, M. (2017). An ontology-based architecture for context recommendation system in E-learning and mobile-learning applications. *2017 International Conference on Electrical and Information Technologies (ICEIT)*, 1–6. <https://doi.org/10.1109/EITech.2017.8255278>

- Brandes, U., & Pich, C. (2007). Centrality estimation in large networks. *International Journal of Bifurcation and Chaos*, 17(07), 2303–2318. <https://doi.org/10.1142/S0218127407018403>
- Bredo, E. (1994). Reconstructing educational psychology: Situated cognition and Deweyian pragmatism. *Educational Psychologist*, 29(1), 23–35. https://doi.org/10.1207/s15326985ep2901_3
- Brézillon, P. (2014). Representing Experience-Based Reasoning by Contextual Graphs. In P. Brézillon & A. J. Gonzalez (Eds.), *Context in Computing: A Cross-Disciplinary Approach for Modeling the Real World* (pp. 509–521). Springer. https://doi.org/10.1007/978-1-4939-1887-4_31
- Brik, M., & Touahria, M. (2020). Contextual Information Retrieval within Recommender System: Case Study “E-learning System.” *TEM Journal*, 1150–1162. <https://doi.org/10.18421/TEM93-41>
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher*, 18(1), 32–42. <https://doi.org/10.3102/0013189X018001032>
- Brusilovsky, P. (2001). Adaptive educational hypermedia. *International PEG Conference*, 8–12.
- Buchmann, U. (2022). Vocational-Scientific Education in Qualification Design – Conditional Factors for Quality Assurance and Development of Qualifications Across Educational Sectors. In V. Rein & J. Wildt (Eds.), *Professional-scientific education: Discourses, perspectives, implications and options for science and practice* (pp. 341–371). Verlag Barbara Budrich.
- Buchmann, U., Dornhöfer, M., Fathi, M., Kismihók, G., Köhler, S., Rasheed, H. A., Schröder, M., & Weber, C. (2024). KI-basierter digitaler Weiterbildungsraum WBsmart. *Digitalisierung in Den Gesundheitsberufen*, 214.
- Bull, S. (1997). Promoting Effective Learning Strategy Use in Call. *Computer Assisted Language Learning*, 10(1), 3–39. <https://doi.org/10.1080/0958822970100102>
- Bull, S., & Kay, J. (2016). SMILI©: A Framework for Interfaces to Learning Data in Open Learner Models, Learning Analytics and Related Fields. *International Journal of Artificial Intelligence in Education*, 26(1), 293–331. <https://doi.org/10.1007/s40593-015-0090-8>
- Calero López, I., & Rodríguez-López, B. (2020). The relevance of transversal competences in vocational education and training: A bibliometric analysis. *Empirical Research in Vocational Education and Training*, 12(1), 12. <https://doi.org/10.1186/s40461-020-00100-0>
- Canbaloglu, G., & Treur, J. (2022). Using Boolean Functions of Context Factors for Adaptive Mental Model Aggregation in Organisational Learning. In V. V. Klimov & D. J. Kelley (Eds.), *Biologically Inspired Cognitive Architectures 2021* (Vol. 1032, pp. 54–68). Springer International Publishing. https://doi.org/10.1007/978-3-030-96993-6_5
- Caponera, E., & Losito, B. (2016). Context factors and student achievement in the IEA studies: Evidence from TIMSS. *Large-Scale Assessments in Education*, 4(1), 12. <https://doi.org/10.1186/s40536-016-0030-6>

- Carbone, M., Colace, F., Lombardi, M., Marongiu, F., Santaniello, D., & Valentino, C. (2021). An Adaptive Learning Path Builder based on a Context Aware Recommender System. *2021 IEEE Frontiers in Education Conference (FIE)*, 1–5. <https://doi.org/10.1109/FIE49875.2021.9637465>
- Cerinšek, G., Petersen, S. A., & Heikura, T. (2013). Contextually enriched competence model in the field of sustainable manufacturing for simulation style technology enhanced learning environments. *Journal of Intelligent Manufacturing*, 24(3), 441–455. <https://doi.org/10.1007/s10845-011-0554-0>
- Chanaa, A., & El Faddouli, N. (2022). An Analysis of learners' affective and cognitive traits in Context-Aware Recommender Systems (CARS) using feature interactions and Factorization Machines (FMs). *Journal of King Saud University - Computer and Information Sciences*, 34(8, Part A), 4796–4809. <https://doi.org/10.1016/j.jksuci.2021.06.008>
- Chau, H., Barria-Pineda, J., & Brusilovsky, P. (2018). Learning Content Recommender System for Instructors of Programming Courses. In C. Penstein Rosé, R. Martínez-Maldonado, H. U. Hoppe, R. Luckin, M. Mavrikis, K. Porayska-Pomsta, B. McLaren, & B. du Boulay (Eds.), *Artificial Intelligence in Education* (pp. 47–51). Springer International Publishing. https://doi.org/10.1007/978-3-319-93846-2_9
- Chaudhry, M. A., Cukurova, M., & Luckin, R. (2022). A Transparency Index Framework for AI in Education. In M. M. Rodrigo, N. Matsuda, A. I. Cristea, & V. Dimitrova (Eds.), *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners' and Doctoral Consortium* (Vol. 13356, pp. 195–198). Springer International Publishing. https://doi.org/10.1007/978-3-031-11647-6_33
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access*, 8, 75264–75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Chen, Y., & Miyazaki, J. (2020). A Model-Agnostic Recommendation Explanation System Based on Knowledge Graph. In S. Hartmann, J. Küng, G. Kotsis, A. M. Tjoa, & I. Khalil (Eds.), *Database and Expert Systems Applications* (Vol. 12392, pp. 149–163). Springer International Publishing. https://doi.org/10.1007/978-3-030-59051-2_10
- Cheng, Y., Chen, K., Sun, H., Zhang, Y., & Tao, F. (2018). Data and knowledge mining with big data towards smart production. *Journal of Industrial Information Integration*, 9, 1–13. <https://doi.org/10.1016/j.jii.2017.08.001>
- Cheng, Z., Yu, J., & Xing, X. (2025). *A Survey on Explainable Deep Reinforcement Learning* (No. arXiv:2502.06869). arXiv. <https://doi.org/10.48550/arXiv.2502.06869>
- Chicaiza, J., & Valdiviezo-Diaz, P. (2021). A Comprehensive Survey of Knowledge Graph-Based Recommender Systems: Technologies, Development, and Contributions. *Information*, 12(6), Article 6. <https://doi.org/10.3390/info12060232>
- Chimalakonda, S., & Nori, K. V. (2020). An ontology based modeling framework for design of educational technologies. *Smart Learning Environments*, 7(1), 28. <https://doi.org/10.1186/s40561-020-00135-6>

- Chiou, C.-K., & Tseng, J. C. R. (2012). Design of a personalized navigation support system for context-aware ubiquitous learning environment. *Proceedings of the 2012 RecSys Workshop on Personalizing the Local Mobile Experience*, 1–6. <https://doi.org/10.1145/2365946.2365948>
- Chorfi, H. O., Sevkli, A. Z., & Bousbahi, F. (2012). From e-Learning to m-Learning: Context-Aware CBR System. In T. Huang, Z. Zeng, C. Li, & C. S. Leung (Eds.), *Neural Information Processing* (Vol. 7666, pp. 528–534). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-34478-7_64
- Chung, K. S. (2020). The Learning Context Analysis System for Digital Textbook Service on Learning Cloud. *Journal of Physics: Conference Series*, 1584(1), 012002. <https://doi.org/10.1088/1742-6596/1584/1/012002>
- Ciloglugil, B., & Inceoglu, M. M. (2018). A Learner Ontology Based on Learning Style Models for Adaptive E-Learning. In O. Gervasi, B. Murgante, S. Misra, E. Stankova, C. M. Torre, A. M. A. C. Rocha, D. Tanar, B. O. Apduhan, E. Tarantino, & Y. Ryu (Eds.), *Computational Science and Its Applications – ICCSA 2018* (pp. 199–212). Springer International Publishing. https://doi.org/10.1007/978-3-319-95165-2_14
- Ciordas-Hertel, G.-P. (2020). How to Complement Learning Analytics with Smartwatches?: Fusing Physical Activities, Environmental Context, and Learning Activities. *Proceedings of the 2020 International Conference on Multimodal Interaction*, 708–712. <https://doi.org/10.1145/3382507.3421151>
- Ciordas-Hertel, G.-P., Biedermann, D., Winter, M., Mordel, J., & Drachsler, H. (2022). How can Interaction Data be Contextualized with Mobile Sensing to Enhance Learning Engagement Assessment in Distance Learning? *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION*, 105–112. <https://doi.org/10.1145/3536220.3558037>
- Clarkes-Nias, J., Mutahi, J., Kinai, A., Bent, O., Weldemariam, K., & Srivastava, S. (2015). Towards Capturing Learners Sentiment and Context. *Proceedings of the Second (2015) ACM Conference on Learning @ Scale*, 217–222. <https://doi.org/10.1145/2724660.2728662>
- Colman, A., Hussein, M., Han, J., & Kapuruge, M. (2014). Context Aware and Adaptive Systems. In P. Brézillon & A. J. Gonzalez (Eds.), *Context in Computing: A Cross-Disciplinary Approach for Modeling the Real World* (pp. 63–82). Springer. https://doi.org/10.1007/978-1-4939-1887-4_5
- Conati, C., Barral, O., Putnam, V., & Rieger, L. (2021). Toward personalized XAI: A case study in intelligent tutoring systems. *Artificial Intelligence*, 298, 103503. <https://doi.org/10.1016/j.artint.2021.103503>
- Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society*, 1(1), 104. <https://doi.org/10.1007/BF03177550>
- Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology*, 88(4), 715–730. <https://doi.org/10.1037/0022-0663.88.4.715>

- Danes, S. M., & Brewton, K. E. (2014). The Role of Learning Context in High School Students' Financial Knowledge and Behavior Acquisition. *Journal of Family and Economic Issues*, 35(1), 81–94. <https://doi.org/10.1007/s10834-013-9351-6>
- Daoudi, M., Lebdiri, N., & Oumaira, I. (2020). Determining the Learner's Profile and Context Profile in Order to Propose Adaptive Mobile Interfaces Based on Machine Learning. *2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, 1–6. <https://doi.org/10.1109/ICECOCS50124.2020.9314518>
- Definition of CONTEXT.* (2023, March 25). <https://www.merriam-webster.com/dictionary/context>
- del Carmen Rodríguez-Hernández, M., & Ilarri, S. (2021). AI-based mobile context-aware recommender systems from an information management perspective: Progress and directions. *Knowledge-Based Systems*, 215, 106740. <https://doi.org/10.1016/j.knosys.2021.106740>
- Dennen, V. P., Burner, K. J., & Cates, M. L. (2018). Information and Communication Technologies, and Learning Theories: Putting Pedagogy into Practice. In J. Voogt, G. Knezek, R. Christensen, & K.-W. Lai (Eds.), *Second Handbook of Information Technology in Primary and Secondary Education* (pp. 143–160). Springer International Publishing. https://doi.org/10.1007/978-3-319-71054-9_9
- Dewey, J. (1916). *Democracy and Education (1916)*. Macmillan.
- Dey, A. K. (2001). Understanding and Using Context. *Personal and Ubiquitous Computing*, 5(1), 4–7. <https://doi.org/10.1007/s007790170019>
- di Mascio, T., Gennari, R., & Vittorini, P. (2013). Blending Evidence and Users for TEL: An Overture. In D. Hernández-Leo, T. Ley, R. Klamka, & A. Harrer (Eds.), *Scaling up Learning for Sustained Impact* (Vol. 8095, pp. 544–547). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40814-4_54
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1), Article 1. <https://doi.org/10.1007/BF01386390>
- Dikaya, L. A., Avanesian, G., Dikiy, I. S., Kirik, V. A., & Egorova, V. A. (2021). How Personality Traits Are Related to the Attitudes Toward Forced Remote Learning During COVID-19: Predictive Analysis Using Generalized Additive Modeling. *Frontiers in Education*, 6. <https://doi.org/10.3389/educ.2021.629213>
- Drachsler, H., Verbert, K., Santos, O., & Manouselis, N. (2015). Panorama of Recommender Systems to Support Learning. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender Systems Handbook* (pp. 421–451). Springer.
- D'Souza, J., & Auer, S. (2022). *Computer Science Named Entity Recognition in the Open Research Knowledge Graph* (No. arXiv:2203.14579; Version 2). arXiv. <https://doi.org/10.48550/arXiv.2203.14579>
- Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior*, 52, 338–348. <https://doi.org/10.1016/j.chb.2015.05.041>
- Durand, G., Laplante, F., & Kop, R. (2011). *A Learning Design Recommendation System Based on Markov Decision Processes*.

- Edmonds, B. (2014). Contextual Cognition in Social Simulation. In P. Brézillon & A. J. Gonzalez (Eds.), *Context in Computing: A Cross-Disciplinary Approach for Modeling the Real World* (pp. 273–290). Springer. https://doi.org/10.1007/978-1-4939-1887-4_18
- Engin, G., Aksoyer, B., Avdagic, M., Bozanlı, D., Hanay, U., Maden, D., & Ertek, G. (2014). Rule-based Expert Systems for Supporting University Students. *Procedia Computer Science*, 31, 22–31. <https://doi.org/10.1016/j.procs.2014.05.241>
- Eraut, M. (2000). Non-formal learning and tacit knowledge in professional work. *British Journal of Educational Psychology*, 70(1), 113–136. <https://doi.org/10.1348/000709900158001>
- Eraut, M. (2004). Informal learning in the workplace. *Studies in Continuing Education*, 26(2), 247–273. <https://doi.org/10.1080/158037042000225245>
- Ertmer, P. A., & Newby, T. J. (2013). Behaviorism, Cognitivism, Constructivism: Comparing Critical Features From an Instructional Design Perspective. *Performance Improvement Quarterly*, 26(2), 43–71. <https://doi.org/10.1002/piq.21143>
- European Centre for the Development of Vocational Training. (2017). *The changing nature and role of vocational education and training in Europe. Volume 1, Conceptions of vocational education and training: An analytical framework*. Publications Office. <https://data.europa.eu/doi/10.2801/532605>
- Evans, P. (2015). Open online spaces of professional learning: Context, personalisation and facilitation. *TechTrends*, 59(1), 31–36. <https://doi.org/10.1007/s11528-014-0817-7>
- Fancsali, S. E., & Ritter, S. (2014). Context personalization, preferences, and performance in an intelligent tutoring system for middle school mathematics. *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, 73–77. <https://doi.org/10.1145/2567574.2567615>
- Foltýnek, T., Dlabolová, D., Anohina-Naumeca, A., Razi, S., Kravjar, J., Kamzola, L., Guerrero-Dib, J., Çelik, Ö., & Weber-Wulff, D. (2020). Testing of support tools for plagiarism detection. *International Journal of Educational Technology in Higher Education*, 17(1), 46. <https://doi.org/10.1186/s41239-020-00192-4>
- Forissier, T., Bourdeau, J., Mazabraud, Y., & Nkambou, R. (2014). Computing the Context Effect for Science Learning. In P. Brézillon & A. J. Gonzalez (Eds.), *Context in Computing: A Cross-Disciplinary Approach for Modeling the Real World* (pp. 255–269). Springer. https://doi.org/10.1007/978-1-4939-1887-4_17
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215–239. [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7)
- Gallego, D., Barra, E., Gordillo, A., & Huecas, G. (2013). Enhanced recommendations for e-Learning authoring tools based on a proactive context-aware recommender. *2013 IEEE Frontiers in Education Conference (FIE)*, 1393–1395. <https://doi.org/10.1109/FIE.2013.6685060>
- Gambhir, M., & Gupta, V. (2017). Recent automatic text summarization techniques: A survey. *Artificial Intelligence Review*, 47(1), 1–66. <https://doi.org/10.1007/s10462-016-9475-9>

- García, J., García-Peñalvo, F. J., & Therón, R. (2010). A Survey on Ontology Metrics. In M. D. Lytras, P. Ordonez De Pablos, A. Ziderman, A. Roulstone, H. Maurer, & J. B. Imber (Eds.), *Knowledge Management, Information Systems, E-Learning, and Sustainability Research* (Vol. 111, pp. 22–27). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-16318-0_4
- Giabbanelli, P. J., Tawfik, A. A., & Gupta, V. K. (2019). Learning Analytics to Support Teachers' Assessment of Problem Solving: A Novel Application for Machine Learning and Graph Algorithms. In D. Ifenthaler, D.-K. Mah, & J. Y.-K. Yau (Eds.), *Utilizing Learning Analytics to Support Study Success* (pp. 175–199). Springer International Publishing. https://doi.org/10.1007/978-3-319-64792-0_11
- Gianotti, R. C., Cazella, S. C., & Behar, P. A. (2019). A Model for Integrating Personality Traits into an Educational Recommender System. *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*, 2161-377X, 383–385. <https://doi.org/10.1109/ICALT.2019.00119>
- Glaserfeld, E. von. (1995). *Radical Constructivism: A Way of Knowing and Learning*. Taylor and Francis.
- Gogus, A. (2012). Constructivist Learning. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 783–786). Springer US. https://doi.org/10.1007/978-1-4419-1428-6_142
- Gómez, S., Zervas, P., Sampson, D. G., & Fabregat, R. (2014). Context-aware adaptive and personalized mobile learning delivery supported by UoLmP. *Journal of King Saud University - Computer and Information Sciences*, 26(1), 47–61. <https://doi.org/10.1016/j.jksuci.2013.10.008>
- Guéret, C., Groth, P., Stadler, C., & Lehmann, J. (2012). Assessing Linked Data Mappings Using Network Measures. In E. Simperl, P. Cimiano, A. Polleres, O. Corcho, & V. Presutti (Eds.), *The Semantic Web: Research and Applications* (Vol. 7295, pp. 87–102). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-30284-8_13
- Guo, Q., Zhuang, F., Qin, C., Zhu, H., Xie, X., Xiong, H., & He, Q. (2020). A Survey on Knowledge Graph-Based Recommender Systems. *IEEE Transactions on Knowledge and Data Engineering*, 1–1. <https://doi.org/10.1109/TKDE.2020.3028705>
- Hampel, R., & de los Arcos, B. (2013). Interacting at a distance: A critical review of the role of ICT in developing the learner–context interface in a university language programme. *Innovation in Language Learning and Teaching*, 7(2), 158–178. <https://doi.org/10.1080/17501229.2013.776051>
- Haruna, K., Akmar Ismail, M., Suhendroyono, S., Damiasih, D., Pierewan, A., Chiroma, H., & Herawan, T. (2017). Context-Aware Recommender System: A Review of Recent Developmental Process and Future Research Direction. *Applied Sciences*, 7(12), 1211. <https://doi.org/10.3390/app7121211>
- Hemmler, Y. M., & Ifenthaler, D. (2022). Indicators of the Learning Context for Supporting Personalized and Adaptive Learning Environments. *2022 International Conference on Advanced Learning Technologies (ICALT)*, 61–65. <https://doi.org/10.1109/ICALT55010.2022.00026>

- Herrero-Martín, J., Canaleta, X., del Valle, J., & Torres, R. (2022). Teaching Learning Interactions in Secondary School: Towards a New Narrative Learning Context. In P. Zaphiris & A. Ioannou (Eds.), *Learning and Collaboration Technologies. Designing the Learner and Teacher Experience* (Vol. 13328, pp. 433–444). Springer International Publishing. https://doi.org/10.1007/978-3-031-05657-4_31
- Hilkenmeier, F., Goller, M., & Schaper, N. (2021). The Differential Influence of Learner Factors and Learning Context on Different Professional Learning Activities. *Vocations and Learning*, 14(3), 411–438. <https://doi.org/10.1007/s12186-021-09266-4>
- Hogan, A., Blomqvist, E., Cochez, M., D'amato, C., Melo, G. D., Gutierrez, C., Kirrane, S., Gayo, J. E. L., Navigli, R., Neumaier, S., Ngomo, A.-C. N., Polleres, A., Rashid, S. M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., & Zimmermann, A. (2021). Knowledge Graphs. *ACM Comput. Surv.*, 54(4), 71:1-71:37. <https://doi.org/10.1145/3447772>
- Huang, S., Yin, B., & Liu, M. (2017). Research on Individualized Learner Model Based on Context-Awareness. *2017 International Symposium on Educational Technology (ISET)*, 163–167. <https://doi.org/10.1109/ISET.2017.45>
- Hubert, N., Brun, A., & Monticolo, D. (2022, October). New Ontology and Knowledge Graph for University Curriculum Recommendation. *ISWC 2022 - The 21st International Semantic Web Conference*. <https://hal.archives-ouvertes.fr/hal-03768154>
- Hung, D. W. L., & Chen, D.-T. (2001). Situated Cognition, Vygotskian Thought and Learning from the Communities of Practice Perspective: Implications for the Design of Web-Based E-Learning. *Educational Media International*, 38(1), 3–12. <https://doi.org/10.1080/09523980121818>
- Hwang, G.-J. (2013). Paradigm Shifts in E-Learning: From Web-Based Learning to Context-Aware Ubiquitous Learning. In R. Huang, Kinshuk, & J. M. Spector (Eds.), *Reshaping Learning* (pp. 253–271). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-32301-0_11
- Hwang, G.-J., & Chang, C.-Y. (2020). Facilitating decision-making performances in nursing treatments: A contextual digital game-based flipped learning approach. *Interactive Learning Environments*, 1–16. <https://doi.org/10.1080/10494820.2020.1765391>
- Ilkou, E., Abu-Rasheed, H., Tavakoli, M., Hakimov, S., Kismihók, G., Auer, S., & Nejd, W. (2021). EduCOR: An Educational and Career-Oriented Recommendation Ontology. In A. Hotho, E. Blomqvist, S. Dietze, A. Fokoue, Y. Ding, P. Barnaghi, A. Haller, M. Dragoni, & H. Alani (Eds.), *The Semantic Web – ISWC 2021* (Vol. 12922, pp. 546–562). Springer International Publishing. https://doi.org/10.1007/978-3-030-88361-4_32
- Illeris. (2017). *How We Learn: Learning and non-learning in school and beyond*. Routledge & CRC Press. <https://www.routledge.com/How-We-Learn-Learning-and-non-learning-in-school-and-beyond/Illeris/p/book/9781138689817>
- Intayoad, W., Becker, T., & Temdee, P. (2017). Social Context-Aware Recommendation for Personalized Online Learning. *Wireless Personal Communications*, 97(1), 163–179. <https://doi.org/10.1007/s11277-017-4499-2>

- Isaias, P., Miranda, P., & Pifano, S. (2022). Are Learning Theories Being Used to Support the Implementation of Learning Technologies Within Higher Education? A Systematic Review. In P. Zaphiris & A. Ioannou (Eds.), *Learning and Collaboration Technologies. Designing the Learner and Teacher Experience* (Vol. 13328, pp. 445–460). Springer International Publishing. https://doi.org/10.1007/978-3-031-05657-4_32
- Jiang, X., Xu, C., Shen, Y., Sun, X., Tang, L., Wang, S., Chen, Z., Wang, Y., & Guo, J. (2025). *On the Evolution of Knowledge Graphs: A Survey and Perspective* (No. arXiv:2310.04835; Version 3). arXiv. <https://doi.org/10.48550/arXiv.2310.04835>
- Junianto, & Wutsqa, D. U. (2019). Comparison of Effectiveness between Contextual Teaching and Learning (CTL) and Problem-Based Learning (PBL) Approach on the Interest of Junior High School Students. *Proceedings of the 2019 International Conference on Mathematics, Science and Technology Teaching and Learning*, 39–43. <https://doi.org/10.1145/3348400.3348407>
- Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable Artificial Intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074. <https://doi.org/10.1016/j.caeai.2022.100074>
- Kim, M. K., & McCarthy, K. S. (2021). Using graph centrality as a global index to assess students' mental model structure development during summary writing. *Educational Technology Research and Development*, 69(2), 971–1002. <https://doi.org/10.1007/s11423-021-09942-1>
- Knobelsdorf, M., & Tenenbergs, J. (2013). The Context-Based Approach IniK in Light of Situated and Constructive Learning Theories. In I. Diethelm & R. T. Mittermeir (Eds.), *Informatics in Schools. Sustainable Informatics Education for Pupils of all Ages* (Vol. 7780, pp. 103–114). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-36617-8_9
- Kolb, D. (1984). Experiential Learning: Experience As The Source Of Learning And Development. In *Journal of Business Ethics* (Vol. 1).
- Koubek, J., Schulte, C., Schulze, P., & Witten, H. (2009). *Informatik im Kontext (IniK) – Ein integratives Unterrichtskonzept für den Informatikunterricht*. Gesellschaft für Informatik e.V. <http://dl.gi.de/handle/20.500.12116/20333>
- Ku, J.-H. (2014). Design and Implementation of a Context Aware Contents Service Model in Mobile LMS. *International Journal of Multimedia and Ubiquitous Engineering*, 9(10), 307–316. <https://doi.org/10.14257/ijmue.2014.9.10.30>
- Kuger, S., & Klieme, E. (2016). Dimensions of Context Assessment. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds.), *Assessing Contexts of Learning: An International Perspective* (pp. 3–37). Springer International Publishing. https://doi.org/10.1007/978-3-319-45357-6_1
- Kuger, S., Klieme, E., Jude, N., & Kaplan, D. (Eds.). (2016). *Assessing Contexts of Learning*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-45357-6>
- Kulkarni, S., & Rodd, S. F. (2020). Context Aware Recommendation Systems: A review of the state of the art techniques. *Computer Science Review*, 37, 100255. <https://doi.org/10.1016/j.cosrev.2020.100255>

- Lakkaraju, H., Aguiar, E., Shan, C., Miller, D., Bhanpuri, N., Ghani, R., & Addison, K. L. (2015). A Machine Learning Framework to Identify Students at Risk of Adverse Academic Outcomes. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1909–1918. <https://doi.org/10.1145/2783258.2788620>
- Lallemant, C., & Koenig, V. (2020). Measuring the Contextual Dimension of User Experience: Development of the User Experience Context Scale (UXCS). *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, 1–13. <https://doi.org/10.1145/3419249.3420156>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press. <https://doi.org/10.1017/CBO9780511815355>
- Lee, H., & Kizilcec, R. F. (2020). *Evaluation of Fairness Trade-offs in Predicting Student Success* (No. arXiv:2007.00088). arXiv. <https://doi.org/10.48550/arXiv.2007.00088>
- Li, H., Deng, J., Yuan, S., Feng, P., & Arachchige, D. D. K. (2021). Monitoring and Identifying Wind Turbine Generator Bearing Faults Using Deep Belief Network and EWMA Control Charts. *Frontiers in Energy Research*, 9, 799039. <https://doi.org/10.3389/fenrg.2021.799039>
- Li, H., Wang, L., Du, X., & Zhang, M. (2017). Research on the Strategy of E-Learning Resources Recommendation Based on Learning Context. *2017 International Conference of Educational Innovation through Technology (EITT)*, 209–213. <https://doi.org/10.1109/EITT.2017.58>
- Li, K. (2019). MOOC learners' demographics, self-regulated learning strategy, perceived learning and satisfaction: A structural equation modeling approach. *Computers & Education*, 132, 16–30. <https://doi.org/10.1016/j.compedu.2019.01.003>
- Li, Y., Liang, Y., Yang, R., Qiu, J., Zhang, C., & Zhang, X. (2024). CourseKG: An Educational Knowledge Graph Based on Course Information for Precision Teaching. *Applied Sciences*, 14(7), Article 7. <https://doi.org/10.3390/app14072710>
- Li, Y., Zhou, J., Verma, S., & Chen, F. (2022). *A Survey of Explainable Graph Neural Networks: Taxonomy and Evaluation Metrics* (No. arXiv:2207.12599). arXiv. <https://doi.org/10.48550/arXiv.2207.12599>
- Li, Z. (2016). Rethinking the relationship between learner, learning contexts, and technology: A critique and exploration of Archer's morphogenetic approach. *Learning, Media and Technology*, 41(3), 501–520. <https://doi.org/10.1080/17439884.2014.978336>
- Lin, C.-Y. (2004). ROUGE: A Package for Automatic Evaluation of Summaries. *ACL 2004*.
- Liu, D., He, M., Luo, J., Lin, J., Wang, M., Zhang, X., Pan, W., & Ming, Z. (2022). User-Event Graph Embedding Learning for Context-Aware Recommendation. *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 1051–1059. <https://doi.org/10.1145/3534678.3539458>
- Liu, Q., Pinto, J. D., & Paquette, L. (2024). Applications of Explainable AI (XAI) in Education. In D. Kourkoulou, A.-O. (Olnancy) Tzirides, B. Cope, & M. Kalantzis (Eds.), *Trust and Inclusion in AI-Mediated Education: Where Human Learning*

- Meets Learning Machines* (pp. 93–109). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-64487-0_5
- Long, Y., & Aleven, V. (2017). Enhancing learning outcomes through self-regulated learning support with an Open Learner Model. *User Modeling and User-Adapted Interaction*, 27(1), 55–88. <https://doi.org/10.1007/s11257-016-9186-6>
- Luckin, R. (2010). *Re-designing learning contexts technology-rich, learner-centred ecologies*. Routledge.
- Macchia, G., & Brézillon, P. (2021). Automated generation of pedagogical activities adapted to the learning context. *Modélisation et Utilisation Du Contexte*, 4(Special Issue). <https://doi.org/10.21494/ISTE.OP.2021.0689>
- Madhu Sudhana, K. (2015). Contextual Diversity and Rule-based Adaptive E-learning System Scheme. *International Journal of Emerging Technologies in Learning (iJET)*, 10(5), 72. <https://doi.org/10.3991/ijet.v10i5.4799>
- Manouselis, N., Drachsler, H., Verbert, K., & Duval, E. (2013). TEL as a Recommendation Context. In N. Manouselis, H. Drachsler, K. Verbert, & E. Duval, *Recommender Systems for Learning* (pp. 21–36). Springer New York. https://doi.org/10.1007/978-1-4614-4361-2_2
- Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., & Koper, R. (2011). Recommender Systems in Technology Enhanced Learning. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender Systems Handbook* (pp. 387–415). Springer US. https://doi.org/10.1007/978-0-387-85820-3_12
- Marton, F., & Saljo, R. (1976). On qualitative differences in learning: I. Outcome and process. *British Journal of Educational Psychology*, 46(1), 4–11. <https://doi.org/10.1111/j.2044-8279.1976.tb02980.x>
- Mausz, J., & Tavares, W. (2017). Learning in professionally ‘distant’ contexts: Opportunities and challenges. *Advances in Health Sciences Education*, 22(3), 581–600. <https://doi.org/10.1007/s10459-016-9693-6>
- Mayeku, B., & Hogrefe, D. (2017). A Framework for Integrating Pedagogy and Technology in Context-Aware Learning Environment. In H. Xie, E. Popescu, G. Hancke, & B. Fernández Manjón (Eds.), *Advances in Web-Based Learning – ICWL 2017* (pp. 215–220). Springer International Publishing. https://doi.org/10.1007/978-3-319-66733-1_24
- McGivney, V. (1999). *Informal Learning in the Community: A Trigger for Change and Development*. National Institute for Adult Continuing Education, 21 De Montfort Street, Leicester, LE1 7GE, United Kingdom (13).
- Moebert, T., Zender, R., & Lucke, U. (2016). A Generalized Approach for Context-Aware Adaptation in Mobile E-Learning Settings. In A. Peña-Ayala (Ed.), *Mobile, Ubiquitous, and Pervasive Learning: Fundaments, Applications, and Trends* (pp. 23–53). Springer International Publishing. https://doi.org/10.1007/978-3-319-26518-6_2
- Mohammad, M., Alsaig, A., & Alsaig, A. (2015). Introducing Context-Awareness to MOOC Systems. *Proceedings of the 3rd International Conference on Context-Aware Systems and Applications*. 3rd International Conference on Context-Aware Systems and Applications, Dubai, United Arab Emirates. <https://doi.org/10.4108/icst.iccasa.2014.257285>

- Monge, A. E., & Elkan, C. (1997). An Efficient Domain-Independent Algorithm for Detecting Approximately Duplicate Database Records. *Workshop on Research Issues on Data Mining and Knowledge Discovery, DMKD 1997 in Cooperation with ACM SIGMOD'97, Tucson, Arizona, USA, May 11, 1997*.
- Moon, S., Shah, P., Kumar, A., & Subba, R. (2019). OpenDialKG: Explainable Conversational Reasoning with Attention-based Walks over Knowledge Graphs. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 845–854. <https://doi.org/10.18653/v1/P19-1081>
- Mowafi, Y., Alazrai, R., & Zmily, A. (2014). User-Centered Approaches to Context Awareness: Prospects and Challenges. In P. Brézillon & A. J. Gonzalez (Eds.), *Context in Computing: A Cross-Disciplinary Approach for Modeling the Real World* (pp. 221–236). Springer. https://doi.org/10.1007/978-1-4939-1887-4_15
- Musumba, G. W., & Wario, R. D. (2019). Towards a Context-Aware Adaptive e-Learning Architecture. In S. Kabanda, H. Suleman, & S. Gruner (Eds.), *ICT Education* (pp. 191–206). Springer International Publishing. https://doi.org/10.1007/978-3-030-05813-5_13
- Mutambik, I., Lee, J., & Almuqrin, A. (2020). Role of gender and social context in readiness for e-learning in Saudi high schools. *Distance Education*, 41(4), 515–539. <https://doi.org/10.1080/01587919.2020.1821602>
- Naiseh, M. (2020). Explainability Design Patterns in Clinical Decision Support Systems. In F. Dalpiaz, J. Zdravkovic, & P. Loucopoulos (Eds.), *Research Challenges in Information Science* (Vol. 385, pp. 613–620). Springer International Publishing. https://doi.org/10.1007/978-3-030-50316-1_45
- Negovan, V., & Bogdan, C. (2013). Learning Context and Undergraduate Students' Needs for Autonomy and Competence, Achievement Motivation and Personal Growth Initiative. *Procedia - Social and Behavioral Sciences*, 78, 300–304. <https://doi.org/10.1016/j.sbspro.2013.04.299>
- Niemann, K., & Wolpers, M. (2013). Usage Context-Boosted Filtering for Recommender Systems in TEL. In D. Hernández-Leo, T. Ley, R. Klamma, & A. Harrer (Eds.), *Scaling up Learning for Sustained Impact* (Vol. 8095, pp. 246–259). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40814-4_20
- Nijenhuis-Voogt, J., Meijer, P. C., & Barendsen, E. (2018). Context-based teaching and learning of fundamental computer science concepts: Exploring teachers' ideas. *Proceedings of the 13th Workshop in Primary and Secondary Computing Education*, 1–4. <https://doi.org/10.1145/3265757.3265772>
- Nuankaew, W., & Nuankaew, P. (2019). The Study of the Factors and Development of Educational Model: The Relationship between the Learner Context and the Curriculum Context in Higher Education. *International Journal of Emerging Technologies in Learning (iJET)*, 14(21), 205–226.
- Ogata, H., Flanagan, B., Takami, K., Dai, Y., Nakamoto, R., & Takii, K. (2023). EXAIT: Educational eXplainable Artificial Intelligent Tools for personalized learning. *Research and Practice in Technology Enhanced Learning*, 19, 019. <https://doi.org/10.58459/rptel.2024.19019>
- Ogata, H., Flanagan, B., Takami, K., Dai, Y., Nakamoto, R., & Takii, K. (2024). EXAIT: Educational eXplainable Artificial Intelligent Tools for personalized learning.

- Research and Practice in Technology Enhanced Learning*, 19, 019–019. <https://doi.org/10.58459/rptel.2024.19019>
- Ooge, J., Kato, S., & Verbert, K. (2022). Explaining Recommendations in E-Learning: Effects on Adolescents' Trust. *27th International Conference on Intelligent User Interfaces*, 93–105. <https://doi.org/10.1145/3490099.3511140>
- Opel, S., & Brinda, T. (2013). *Arguments for Contextual Teaching with Learning Fields in Vocational IT Schools – Results of an Interview Study among IT and CS Training Companies*.
- Ortegón Romero, O. Y., & Krug Wives, L. (2021). Systematic Review of Context-Aware Systems that use Item Response Theory in Learning Environments. *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, 100–104. <https://doi.org/10.1145/3450614.3464482>
- Páez, A. (2019). The Pragmatic Turn in Explainable Artificial Intelligence (XAI). *Minds and Machines*, 29(3), 441–459. <https://doi.org/10.1007/s11023-019-09502-w>
- Pal, S., Pramanik, P. K. D., & Choudhury, P. (2021). Enhanced metadata modelling and extraction methods to acquire contextual pedagogical information from e-learning contents for personalised learning systems. *Multimedia Tools and Applications*, 80(16), 25309–25366. <https://doi.org/10.1007/s11042-020-10380-z>
- Panniello, U., Tuzhilin, A., & Gorgoglione, M. (2014). Comparing context-aware recommender systems in terms of accuracy and diversity. *User Modeling and User-Adapted Interaction*, 24(1), 35–65. <https://doi.org/10.1007/s11257-012-9135-y>
- Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2001). BLEU: A method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*, 311. <https://doi.org/10.3115/1073083.1073135>
- Pardo, A., Han, F., & Ellis, R. A. (2017). Combining University Student Self-Regulated Learning Indicators and Engagement with Online Learning Events to Predict Academic Performance. *IEEE Transactions on Learning Technologies*, 10(1), 82–92. <https://doi.org/10.1109/TLT.2016.2639508>
- Parker, J. E., & Hollister, D. L. (2014). The Cognitive Science Basis for Context. In P. Brézillon & A. J. Gonzalez (Eds.), *Context in Computing: A Cross-Disciplinary Approach for Modeling the Real World* (pp. 205–219). Springer. https://doi.org/10.1007/978-1-4939-1887-4_14
- Pecune, F., Murali, S., Tsai, V., Matsuyama, Y., & Cassell, J. (2019). A Model of Social Explanations for a Conversational Movie Recommendation System. *Proceedings of the 7th International Conference on Human-Agent Interaction*, 135–143. <https://doi.org/10.1145/3349537.3351899>
- Peña de Carrillo, C. I., & Choquet, C. (2013). Adaptive e-tutorship using students context in technology enhanced learning environments. *Entramado*, 9(1).
- Pilz, M. (2016). Typologies in Comparative Vocational Education: Existing Models and a New Approach. *Vocations and Learning*, 9(3), 295–314. <https://doi.org/10.1007/s12186-016-9154-7>

- Pimmer, C., Pachler, N., & Genewein, U. (2013). Contextual dynamics in clinical workplaces: Learning from doctor-doctor consultations: Learning in the dynamics of clinical workplaces. *Medical Education*, 47(5), 463–475. <https://doi.org/10.1111/medu.12130>
- Polley, S., Koparde, R. R., Gowri, A. B., Perera, M., & Nuernberger, A. (2021). Towards Trustworthiness in the Context of Explainable Search. *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2580–2584. <https://doi.org/10.1145/3404835.3462799>
- Prain, V., Cox, P., Deed, C., Dorman, J., Edwards, D., Farrelly, C., Keeffe, M., Lovejoy, V., Mow, L., Sellings, P., Waldrip, B., & Yager, Z. (2013). Personalised learning: Lessons to be learnt. *British Educational Research Journal*, 39(4), 654–676. <https://doi.org/10.1080/01411926.2012.669747>
- Qiuyan, Z., Xiaodong, L., & Shaobo, J. (2013). *Research of Context-oriented Adaptive Content Framework in Seamless Learning*.
- Rachha, A., & Seyam, M. (2023). Explainable AI In Education: Current Trends, Challenges, And Opportunities. *SoutheastCon 2023*, 232–239. <https://doi.org/10.1109/SoutheastCon51012.2023.10115140>
- Rajabalee, Y. B., & Santally, M. I. (2021). Learner satisfaction, engagement and performances in an online module: Implications for institutional e-learning policy. *Education and Information Technologies*, 26(3), 2623–2656. <https://doi.org/10.1007/s10639-020-10375-1>
- Ramakrishnan, A. K., Preuveneers, D., & Berbers, Y. (2014). A Bayesian Framework for Life-Long Learning in Context-Aware Mobile Applications. In P. Brézillon & A. J. Gonzalez (Eds.), *Context in Computing: A Cross-Disciplinary Approach for Modeling the Real World* (pp. 127–141). Springer. https://doi.org/10.1007/978-1-4939-1887-4_9
- Ramírez Luelmo, S. I., El Mawas, N., & Heutte, J. (2021). Learner Models for MOOC in a Lifelong Learning Context: A Systematic Literature Review. In H. C. Lane, S. Zvacek, & J. Uhomoibhi (Eds.), *Computer Supported Education* (Vol. 1473, pp. 392–415). Springer International Publishing. https://doi.org/10.1007/978-3-030-86439-2_20
- Reynolds, R., & Chiu, M. M. (2013). Formal and informal context factors as contributors to student engagement in a guided discovery-based program of game design learning. *Learning, Media and Technology*, 38(4), 429–462. <https://doi.org/10.1080/17439884.2013.779585>
- Ricci, F., Rokach, L., & Shapira, B. (Eds.). (2022). *Recommender Systems Handbook*. Springer US. <https://doi.org/10.1007/978-1-0716-2197-4>
- Ryu, H., & Parsons, D. (2008). A learner-centred design of a location-aware learning reminder. *International Journal of Mobile Learning and Organisation*, 2(2), 187–200. <https://doi.org/10.1504/IJMLLO.2008.019768>
- Samek, W., Wiegand, T., & Müller, K.-R. (2017). *Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models* (No. arXiv:1708.08296). arXiv. <https://doi.org/10.48550/arXiv.1708.08296>
- Sancenon, V., Wijaya, K., Wen, X. Y. S., Utama, D. A., Ashworth, M., Ng, K. H., Cheong, A., & Neo, Z. (2022). A New Web-Based Personalized Learning System

- Improves Student Learning Outcomes: *International Journal of Virtual and Personal Learning Environments*, 12(1), 1–21. <https://doi.org/10.4018/IJVPLE.295306>
- Sawyer, R. K. (Ed.). (2014). *The Cambridge Handbook of the Learning Sciences* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139519526>
- Schunk, D. H. (2020). *Learning theories: An educational perspective* (Eighth Edition). Pearson.
- Selwyn, N. (2019). *Should robots replace teachers?: AI and the Future of Education*. Polity Press. <https://research.monash.edu/en/publications/should-robots-replace-teachers-ai-and-the-future-of-education>
- Sequeda, J., Allemang, D., & Jacob, B. (2023). *A Benchmark to Understand the Role of Knowledge Graphs on Large Language Model's Accuracy for Question Answering on Enterprise SQL Databases* (No. arXiv:2311.07509). arXiv. <http://arxiv.org/abs/2311.07509>
- Sevkli, A., & Abdulkarem, H. (2015). A Context-aware Mobile Application for Cultural Learning. *Proceedings of the 3rd International Conference on Context-Aware Systems and Applications*. 3rd International Conference on Context-Aware Systems and Applications, Dubai, United Arab Emirates. <https://doi.org/10.4108/icst.iccasa.2014.257520>
- Shalaby, W., AlAila, B., Korayem, M., Pournajaf, L., AlJadda, K., Quinn, S., & Zadrozny, W. (2017). Help me find a job: A graph-based approach for job recommendation at scale. *2017 IEEE International Conference on Big Data (Big Data)*, 1544–1553. <https://doi.org/10.1109/BigData.2017.8258088>
- Shcherbachenko, L., & Nowakowski, S. (2018). Analysis of Means for building Context-Aware Recommendation System for Mobile Learning. In D. Kergel, B. Heidkamp, P. K. Tellés, T. Rachwal, & S. Nowakowski (Eds.), *The Digital Turn in Higher Education* (pp. 235–245). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-19925-8_16
- Shi, C., Hu, B., Zhao, W. X., & Yu, P. S. (2019). Heterogeneous Information Network Embedding for Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 31(2), 357–370. <https://doi.org/10.1109/TKDE.2018.2833443>
- Shi, D., Wang, T., Xing, H., & Xu, H. (2020). A learning path recommendation model based on a multidimensional knowledge graph framework for e-learning. *Knowledge-Based Systems*, 195, 105618. <https://doi.org/10.1016/j.knosys.2020.105618>
- Shi, L., Cristea, A. I., Hadzidedic, S., & Dervishalidovic, N. (2014). Contextual Gamification of Social Interaction – Towards Increasing Motivation in Social E-learning. In E. Popescu, R. W. H. Lau, K. Pata, H. Leung, & M. Laanpere (Eds.), *Advances in Web-Based Learning – ICWL 2014* (Vol. 8613, pp. 116–122). Springer International Publishing. https://doi.org/10.1007/978-3-319-09635-3_12
- Shin, H. S., & Jeong, A. (2021). Modeling the relationship between students' prior knowledge, causal reasoning processes, and quality of causal maps. *Computers & Education*, 163, 104113. <https://doi.org/10.1016/j.compedu.2020.104113>

- Siraj, F. I., Elgohary, R., Moawad, I., & Abdel Aziz, M. H. (2018). A Context-aware Approach for Adapting Scientific E-Learning Objects. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3389810>
- Sonboli, N., Smith, J. J., Cabral Berenfus, F., Burke, R., & Fiesler, C. (2021). Fairness and Transparency in Recommendation: The Users' Perspective. *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, 274–279. <https://doi.org/10.1145/3450613.3456835>
- Song, W., Duan, Z., Yang, Z., Zhu, H., Zhang, M., & Tang, J. (2019). Explainable Knowledge Graph-based Recommendation via Deep Reinforcement Learning. *arXiv:1906.09506 [Cs]*. <http://arxiv.org/abs/1906.09506>
- Sotsenko, A., Zbick, J., Jansen, M., & Milrad, M. (2016). Flexible and Contextualized Cloud Applications for Mobile Learning Scenarios. In A. Peña-Ayala (Ed.), *Mobile, Ubiquitous, and Pervasive Learning: Fundamentals, Applications, and Trends* (pp. 167–192). Springer International Publishing. https://doi.org/10.1007/978-3-319-26518-6_7
- Soualah-Alila, F., Mendes, F., & Nicolle, C. (2013). *A Context-Based Adaptation In Mobile Learning*.
- Starr, L. M. (2020). *Leadership, Contexts, and Learning—Part 2. Theories of Learning, Channels, and Curricula*.
- Sudhana, Kalla. M., Raj, V. C., & Suresh, R. M. (2013). An ontology-based framework for context-aware adaptive e-learning system. *2013 International Conference on Computer Communication and Informatics*, 1–6. <https://doi.org/10.1109/ICCCI.2013.6466162>
- Supic, H. (2016). A model of a case-based approach to context-aware content sequencing in Mobile learning environments. *2016 International Conference on Information and Digital Technologies (IDT)*, 261–265. <https://doi.org/10.1109/DT.2016.7557183>
- Takami, K., Flanagan, B., & Dai, Y. (2021). Toward Educational Explainable Recommender System: Explanation Generation based on Bayesian Knowledge Tracing Parameters. *International Conference on Computers in Education*. <https://library.apsce.net/index.php/ICCE/article/view/4288>
- Tartir, S., & Arpinar, I. B. (2007). Ontology Evaluation and Ranking using OntoQA. *International Conference on Semantic Computing (ICSC 2007)*, 185–192. <https://doi.org/10.1109/ICSC.2007.19>
- Tarus, J. K., Niu, Z., & Kalui, D. (2018). A hybrid recommender system for e-learning based on context awareness and sequential pattern mining. *Soft Computing*, 22(8), 2449–2461. <https://doi.org/10.1007/s00500-017-2720-6>
- Tavakoli, M., Faraji, A., Molavi, M., T. Mol, S., & Kismihók, G. (2022). Hybrid Human-AI Curriculum Development for Personalised Informal Learning Environments. *LAK22: 12th International Learning Analytics and Knowledge Conference*, 563–569. <https://doi.org/10.1145/3506860.3506917>
- Taylor, D. L., Yeung, M., & Bashet, A. Z. (2021). Personalized and Adaptive Learning. In J. Ryoo & K. Winkelmann (Eds.), *Innovative Learning Environments in STEM Higher Education: Opportunities, Challenges, and Looking Forward* (pp. 17–34). Springer International Publishing. https://doi.org/10.1007/978-3-030-58948-6_2

- Tempelaar, D. T., Wosnitza, M., Volet, S., Rienties, B., Giesbers, B., & Gijssels, W. H. (2013). The role of self- and social directed goals in a problem-based, collaborative learning context. *Higher Education*, 66(2), 253–267. <https://doi.org/10.1007/s10734-012-9602-8>
- Thüs, H., Chatti, M. A., Brandt, R., & Schroeder, U. (2015). Evolution of Interests in the Learning Context Data Model. In G. Conole, T. Klobučar, C. Rensing, J. Konert, & E. Lavoué (Eds.), *Design for Teaching and Learning in a Networked World* (Vol. 9307, pp. 479–484). Springer International Publishing. https://doi.org/10.1007/978-3-319-24258-3_43
- Tiddi, I., Lécué, F., & Hitzler, P. (Eds.). (2020). *Knowledge graphs for explainable artificial intelligence: Foundations, applications and challenges*. IOS Press.
- Tiwari, A., Rajesh, K., & Srujana, P. (2015). Semantically Enriched Knowledge Extraction With Data Mining. *International Journal of Computer Applications Technology and Research*, 4(1), 7–10. <https://doi.org/10.7753/IJCATR0401.1002>
- Tomlinson, C. A. (1999). *The differentiated classroom: Responding to the needs of all learners*. Association for Supervision and Curriculum Development.
- Troussas, C., & Krouska, A. (2023). Path-Based Recommender System for Learning Activities Using Knowledge Graphs. *Information*, 14(1), Article 1. <https://doi.org/10.3390/info14010009>
- Trumpower, D. L., Filiz, M., & Sarwar, G. S. (2014). Assessment for Learning Using Digital Knowledge Maps. In D. Ifenthaler & R. Hanewald (Eds.), *Digital Knowledge Maps in Education* (pp. 221–237). Springer New York. https://doi.org/10.1007/978-1-4614-3178-7_12
- Tsai, C.-H., & Brusilovsky, P. (2019). Explaining recommendations in an interactive hybrid social recommender. *Proceedings of the 24th International Conference on Intelligent User Interfaces*, 391–396. <https://doi.org/10.1145/3301275.3302318>
- Unger, M., Tuzhilin, A., & Livne, A. (2020). Context-Aware Recommendations Based on Deep Learning Frameworks. *ACM Transactions on Management Information Systems*, 11(2), 1–15. <https://doi.org/10.1145/3386243>
- Upadhyay, C., Abu-Rasheed, H., Weber, C., & Fathi, M. (2021). Explainable Job-Posting Recommendations Using Knowledge Graphs and Named Entity Recognition. *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 3291–3296. <https://doi.org/10.1109/SMC52423.2021.9658757>
- Urdaneta-Ponte, M. C., Mendez-Zorrilla, A., & Oleagordia-Ruiz, I. (2021). Recommendation Systems for Education: Systematic Review. *Electronics*, 10(14), Article 14. <https://doi.org/10.3390/electronics10141611>
- US Department of Education. (2017). *Reimagining the Role of Technology in Education: 2017 National Education Technology Plan Update*. Office of Educational Technology, US Department of Education. <https://eric.ed.gov/?id=ED577592>
- VanLehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational Psychologist*, 46(4), 197–221. <https://doi.org/10.1080/00461520.2011.611369>
- Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachsler, H., Bosnic, I., & Duval, E. (2012). Context-Aware Recommender Systems for Learning: A Survey and

- Future Challenges. *IEEE Transactions on Learning Technologies*, 5(4), 318–335. <https://doi.org/10.1109/TLT.2012.11>
- Verma, G., Sengupta, S., Simanta, S., Chen, H., Perge, J. A., Pillai, D., McCrae, J. P., & Buitelaar, P. (2025). *Empowering recommender systems using automatically generated Knowledge Graphs and Reinforcement Learning* (No. arXiv:2307.04996). arXiv. <https://doi.org/10.48550/arXiv.2307.04996>
- vom Brocke, J., Simons, A., Niehaves, B., Niehaves, B., Reimer, K., Plattfaut, R., & Cleven, A. (2009). RECONSTRUCTING THE GIANT: ON THE IMPORTANCE OF RIGOUR IN DOCUMENTING THE LITERATURE SEARCH PROCESS. *ECIS 2009 Proceedings*. <https://aisel.aisnet.org/ecis2009/161>
- Vygotsky, L. S. (1978). *Mind in Society: Development of Higher Psychological Processes* (M. Cole, V. Jolm-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press. <https://doi.org/10.2307/j.ctvjf9vz4>
- Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard university press.
- Wang, X., He, X., Cao, Y., Liu, M., & Chua, T.-S. (2019). KGAT: Knowledge Graph Attention Network for Recommendation. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 950–958. <https://doi.org/10.1145/3292500.3330989>
- Wang, X., Wang, D., Xu, C., He, X., Cao, Y., & Chua, T.-S. (2019). Explainable Reasoning over Knowledge Graphs for Recommendation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33, 5329–5336. <https://doi.org/10.1609/aaai.v33i01.33015329>
- Weber, C. (2016). STUDIO: A Solution on Adaptive Testing. In A. Gábor & A. Kő (Eds.), *Corporate Knowledge Discovery and Organizational Learning* (Vol. 2, pp. 131–153). Springer International Publishing. https://doi.org/10.1007/978-3-319-28917-5_6
- Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), xiii–xxiii.
- Wertsch, J. V. (1985). *Vygotsky and the social formation of mind* (pp. xiv, 262). Harvard University Press.
- Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are We There Yet? - A Systematic Literature Review on Chatbots in Education. *Frontiers in Artificial Intelligence*, 4. <https://www.frontiersin.org/articles/10.3389/frai.2021.654924>
- Wongchiranut, S., Samat, C., Kanjug, I., & Wattanachai, S. (2020). The Study of Learner Context for the Development of Constructivist Learning Environment Model Combined with Mixed Reality Flipped Classroom to Enhance Creative Thinking in Product Design for the High School Students. In T.-C. Huang, T.-T. Wu, J. Barroso, F. E. Sandnes, P. Martins, & Y.-M. Huang (Eds.), *Innovative Technologies and Learning* (Vol. 12555, pp. 338–348). Springer International Publishing. https://doi.org/10.1007/978-3-030-63885-6_38

- Wood, D., Bruner, J. S., & Ross, G. (1976). The Role of Tutoring in Problem Solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- Wu, L., Quan, C., Li, C., Wang, Q., Zheng, B., & Luo, X. (2019). A Context-Aware User-Item Representation Learning for Item Recommendation. *ACM Trans. Inf. Syst.*, 37(2), 22:1-22:29. <https://doi.org/10.1145/3298988>
- Xian, Y., Fu, Z., Muthukrishnan, S., de Melo, G., & Zhang, Y. (2019). Reinforcement Knowledge Graph Reasoning for Explainable Recommendation. *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 285–294. <https://doi.org/10.1145/3331184.3331203>
- Xie, L., Hu, Z., Cai, X., Zhang, W., & Chen, J. (2021). Explainable recommendation based on knowledge graph and multi-objective optimization. *Complex & Intelligent Systems*, 7(3), Article 3. <https://doi.org/10.1007/s40747-021-00315-y>
- Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2023). Practical and Ethical Challenges of Large Language Models in Education: A Systematic Scoping Review. *British Journal of Educational Technology*, bjet.13370. <https://doi.org/10.1111/bjet.13370>
- Yu, R., Pardos, Z., Chau, H., & Brusilovsky, P. (2021). Orienting Students to Course Recommendations Using Three Types of Explanation. *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, 238–245. <https://doi.org/10.1145/3450614.3464483>
- Zarrad, A., & Zaguia, A. (2015). Building a dynamic context aware approach for smart e-learning system. *2015 Second International Conference on Computing Technology and Information Management (ICCTIM)*, 144–149. <https://doi.org/10.1109/ICCTIM.2015.7224609>
- Zaveri, A., Rula, A., Maurino, A., Pietrobon, R., Lehmann, J., & Auer, S. (2016). Quality assessment for Linked Data: A Survey: A systematic literature review and conceptual framework. *Semantic Web*, 7(1), 63–93. <https://doi.org/10.3233/SW-150175>
- Zbick, J., Vogel, B., Spikol, D., Jansen, M., & Milrad, M. (2016). Toward an Adaptive and Adaptable Architecture to Support Ubiquitous Learning Activities. In A. Peña-Ayala (Ed.), *Mobile, Ubiquitous, and Pervasive Learning: Fundamentals, Applications, and Trends* (pp. 193–222). Springer International Publishing. https://doi.org/10.1007/978-3-319-26518-6_8
- Zenkert, J., Weber, C., Klahold, A., Fathi, M., & Hahn, K. (2018). Knowledge-based production documentation analysis: An integrated text mining architecture. *2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS)*, 717–720. <https://ieeexplore.ieee.org/abstract/document/8623836/>
- Zhang, F., Yuan, N. J., Lian, D., Xie, X., & Ma, W.-Y. (2016). Collaborative Knowledge Base Embedding for Recommender Systems. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 353–362. <https://doi.org/10.1145/2939672.2939673>
- Zhang, J., Hao, B., Chen, B., Li, C., Chen, H., & Sun, J. (2019). Hierarchical Reinforcement Learning for Course Recommendation in MOOCs. *Proceedings of*

- the AAAI Conference on Artificial Intelligence*, 33(01), Article 01. <https://doi.org/10.1609/aaai.v33i01.3301435>
- Zhang, J., & Li, J. (2019). Enhanced Knowledge Graph Embedding by Jointly Learning Soft Rules and Facts. *Algorithms*, 12(12), Article 12. <https://doi.org/10.3390/a12120265>
- Zheng, Y. (2017). *Interpreting Contextual Effects By Contextual Modeling In Recommender Systems* (No. arXiv:1710.08516). arXiv. <https://doi.org/10.48550/arXiv.1710.08516>
- Zheng, Y., & Jose, A. A. (2019). Context-aware recommendations via sequential predictions. *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*, 2525–2528. <https://doi.org/10.1145/3297280.3297639>
- Zheng, Y., Shekhar, S., Jose, A. A., & Rai, S. K. (2019). Integrating context-awareness and multi-criteria decision making in educational learning. *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*, 2453–2460. <https://doi.org/10.1145/3297280.3297522>
- Zhong, J., & Negre, E. (2022). Context-Aware Explanations in Recommender Systems. In L. Troiano, A. Vaccaro, N. Kesswani, I. Díaz Rodriguez, & I. Brigui (Eds.), *Progresses in Artificial Intelligence & Robotics: Algorithms & Applications* (pp. 76–85). Springer International Publishing. https://doi.org/10.1007/978-3-030-98531-8_8
- Zhou, Y., Liu, Q., Wu, J., Wang, F., Huang, Z., Tong, W., Xiong, H., Chen, E., & Ma, J. (2021). Modeling Context-aware Features for Cognitive Diagnosis in Student Learning. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 2420–2428. <https://doi.org/10.1145/3447548.3467264>
- Zimmerman, B. J. (2001). Theories of self-regulated learning and academic achievement: An overview and analysis. In *Self-regulated learning and academic achievement: Theoretical perspectives, 2nd ed* (pp. 1–37). Lawrence Erlbaum Associates Publishers.
- Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2
- Zimmerman, H. T., Land, S. M., & Jung, Y. J. (2016). Using Augmented Reality to Support Children's Situational Interest and Science Learning During Context-Sensitive Informal Mobile Learning. In A. Peña-Ayala (Ed.), *Mobile, Ubiquitous, and Pervasive Learning: Fundamentals, Applications, and Trends* (pp. 101–119). Springer International Publishing. https://doi.org/10.1007/978-3-319-26518-6_4

Appendix A: Post test questions of the user study (conducted in German language)

Bitte überlegen Sie, ob sich nach der Lernreise ihre Einschätzung zu den bereits bekannten Statements verändert hat. Diese zweite Einschätzung dient uns nicht als „Lernkontrolle“, sondern hilft uns bei der Einschätzung der Qualität unserer Inhalte und unserer individualisierten Lernpfade.

Itemblock 1: Beruflicher Datenschutz

- Item 1.1: Datenschutz ist etwas, womit ich mich beruflich befasse.
- Item 1.2: Datenschutz erschwert meine Arbeitsabläufe.
- Item 1.3: Datenschutz kostet unnötig viel Zeit.
- Item 1.4: Datenschutz ist anstrengend.
- Item 1.5: Datenschutz erleichtert meine Arbeitsabläufe.
- Item 1.6: Datenschutz ist für mich beruflich wichtig.

Itemblock 2: Privater Datenschutz

- Item 2.1: Datenschutz ist etwas womit ich mich privat befasse.
- Item 2.2: Datenschutz ist für mich im Alltag wichtig.
- Item 2.3: Datenschutz kostet unnötig viel Zeit.
- Item 2.4: Ich mache mir privat nicht so viele Gedanken um Datenschutz.
- Item 2.5: Ich bin bereit, meine Daten zur Verfügung zu stellen, um mir das Leben zu erleichtern.
- Item 2.6: Ich finde es wichtig, eine Kamera-Abdeckung an meinen technischen Geräten zu haben.
- Item 2.7: Ich bevorzuge regelmäßig eine Smart-Watch statt einer analogen Uhr zu tragen.
- Item 2.8: Es ist mir wichtig, mit meinen privaten Daten umzugehen.
- Item 2.9: Ich finde es angenehm, mein Zuhause (teilweise) smart (zum Beispiel Sicherheit, Beleuchtung, Heizung, Geräte, etc.) zu steuern.
- Item 2.10: Mir ist es wichtig, alle AGBs aufmerksam zu lesen.
- Item 2.11: Mir ist es zu zeitraubend alle AGB zu lesen.
- Item 2.12: Für mich ist es in Ordnung, privat ein Amazon-Echo oder ein ähnliches sprachgesteuertes Gerät zu benutzen.
- Item 2.13: Manchmal denke ich: „Big Brother is watching you“.
- Item 2.14: Ich finde es wichtig sich die Zeit zu nehmen, optionale Cookies abzulehnen.

Itemblock 3: Tracking

- Item 3.1: Es ist wichtig für mich zu wissen, wo meine privaten Daten landen.
- Item 3.2: Es ist wichtig für mich zu wissen, wer Zugriff auf meine privaten Daten hat.
- Item 3.3: Ich vertraue denen, die Zugriff auf meine privaten Daten haben.

- Item 3.4: Es ist mir wichtig, die Datenschutzeinstellungen meiner Accounts regelmäßig zu prüfen.
- Item 3.5: Ich bin sicher, dass niemand Zugriff auf meine Daten hat, der darauf keinen Zugriff haben sollte.
- Item 3.6: Für mich ist es in Ordnung überwacht zu werden, solange ich nichts davon mitbekomme.
- Item 3.7: Es ist mir wichtig, regelmäßig die Sprachaufnahmen, die Alexa (oder ein ähnliches Gerät) von mir speichert zu prüfen.
- Item 3.8: Ich interessiere mich dafür, welche Daten meine Geräte und mein Zuhause über mich sammeln und speichern.
- Item 3.9: Die Daten meiner Smart-Watch teile ich mit anderen.
- Item 3.10: Für mich ist es in Ordnung überwacht zu werden, wenn es meiner Sicherheit dient.

Itemblock 4: Bewertung der Lernreise

- Item 4.1: Waren die für Sie empfohlenen Inhalte aus Ihrer Sicht für die Lernreise Datenschutz und Demenz von Bedeutung? (1-5 Likert)
- Item 4.2: Warum waren die für Sie empfohlenen Inhalte aus Ihrer Sicht für die Lernreise Datenschutz und Demenz von Bedeutung oder warum nicht? (offen – keine Pflicht)
- Item 4.3: Waren die für Sie empfohlenen Inhalte der Lernreise für Sie neu? (1-5 Likert)
- Item 4.4: Welche Inhalte waren für Sie neu? (offen – keine Pflicht)
- Item 4.5: Die Reihenfolge der Lernempfehlung hat für mich gepasst. (1-5 Likert)
- Item 4.6: Waren Sie von der Auswahl der Inhalte überrascht? (1-5 Likert)
- Item 4.7: Welche der für Sie empfohlenen Inhalte haben Sie überrascht? (offen – keine Pflicht)
- Item 4.8: War die Lernreise für Ihre Situation / Ihr Vorwissen angemessen? (1 – 5 Likert)

Fragen zur Erklärbarkeit nur für Personen mit Erklärbarkeit.

- Item 4.9: Haben Sie die zusätzlichen Informationen zu den Lernelementen zum Nachdenken angeregt? (Likert 1-5)
- Item 4.10: Die textliche Erklärung macht deutlich, warum mir ein Lerninhalt empfohlen wird. (1 - 5 Likert)
- Item 4.11: Die textliche Erklärung war verständlich. (1 - 5 Likert)
- Item 4.12: Die Erklärung verdeutlicht, warum Kurse und Lerninhalte im Lernpfad miteinander verbunden sind. (1 - 5 Likert)
- Item 4.13: Die visuelle Erklärung gibt einen guten Überblick über den empfohlenen Lernpfad. (1 - 5 Likert)
- Item 4.14: Die visuelle Erklärung hat mir meine Position im Lernpfad gezeigt. (1 - 5 Likert)
- Item 4.15: Ich bevorzuge die rein textliche Erklärung.
- Item 4.16: Ich bevorzuge eine Kombination aus visueller und per Mauszeiger zusätzlich abrufbarer textlicher Erklärung.

Zum Abschluss möchten wir Sie um Ihre Einschätzung einer schwierig zu handhabenden und dennoch alltäglichen Datenschutzfrage bitten:

Item 5: Bitte stellen Sie sich vor, dass Sie für die Betreuung / Pflege eines demenziell erkrankten Menschen verantwortlich sind. Es handelt sich um einen Menschen, der starken Bewegungsdrang hat und diesen konsequent auslebt, sich jedoch nicht ohne Unterstützung orientieren kann, sehr oft stolpert, schnell stürzt und meistens (insbesondere beim raschen Aufstehen) unter starkem Schwindel leidet. Ist es angemessen, diesen Menschen mit mobile Devices (Smart-Watches, Tracker, AirTags, Kameras, Sturzerkennungssysteme oder ähnliches) zu überwachen um Unfälle zu vermeiden? Welche Konsequenzen hat eine solche Überwachung für Sie als pflegende / betreuende Person?

Fragen zum Abschluss

- Item 6.1: War der für Sie empfohlene Inhalt für die Lösung der themenbezogenen Frage aus der Praxis nachvollziehbar? (1-5 Likert)
- Item 6.2: Welche der empfohlenen Inhalte, die zur Beantwortung der themenbezogenen Frage aus der Praxis ausgewählt wurden, konnten Sie nicht nachvollziehen? (offen – keine Pflicht)
- Item 6.3: Wie bewerten Sie den WBsmart-Prototyp insgesamt? (offen – keine Pflicht)

