information

MDPI

Review

Ontology Learning in Educational Systems

Tatyana Ivanova [1,*] and Valentina Terzieva [2,*]

[1] Technical College of Sofia, Technical University of Sofia, 1756 Sofia, Bulgaria
[2] Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria
* Correspondence: t.ivanova@tu-sofia.bg (T.I.); valentina.terzieva@iict.bas.bg (V.T.)

Abstract

E-learning content and participants in the learning process are usually annotated with metadata. Complicated metadata models are necessary for organizing personalized learning, so an ontological metadata representation is used. Since ontologies represent static knowledge, changes in e-learning systems and related description metadata require frequent changes to corresponding ontologies. Only a few professionals in the educational domain have some expertise in ontology development. So, maximal possible automation is of great importance for the development and maintenance of knowledge models, needed for intelligent e-learning environments. Ontology learning is an approach for automatic ontology development and evolution, affected significantly by recent advances in Artificial Intelligence and Language Models. The main objective of this study is to explore and analyze ontology learning approaches and techniques and the specifics of their use in an intelligent e-learning environment. It examines and summarizes recent scientific research to reveal the degree of development and the extent to which ontology learning is applied to support personalized tutoring. The paper outlines trends and challenges of ontology learning from textual e-learning content and comprehensively discusses ontology learning and its applications in intelligent e-learning. It also describes a use case concerning the implementation and practical usage of ontology learning.

Keywords: ontology learning; e-learning; intelligent e-learning environment; personalization

1. Introduction

In recent years, with the penetration of computer-related information and communication technologies (ICT) and because of the COVID-19 pandemic, e-learning has become an essential form of education for a broad audience of students. Thus, e-learning is widely applied in schools, universities, and even business companies where lifelong learning is necessary, so it has already proven effective. With vast amounts of learning resources, such as online courses, videos, articles, podcasts, and e-books, learners can quickly become overwhelmed by the options available to them. Clear guidance is needed to indicate which sources are credible or relevant to their learning objectives. Many digital innovations have been involved in the educational context to do so, changing the educational paradigm to smart education in this manner [1,2]. The implementation of intelligent technologies and approaches has the potential to significantly improve the quality of educational services [3–5]. Intelligent education takes place in a smart e-learning environment that integrates emerging technologies (e.g., learning analytics, Artificial Intelligence, Big Data, ontologies, data mining, and cloud computing) and various innovative ICT tools [6–8].

Check for updates

Academic Editor: Yu-Cheng Wang

Received: 30 December 2025
Revised: 22 January 2026
Accepted: 28 January 2026
Published: 2 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

In contemporary intelligent e-learning environments, it is essential to support personalized learning and, thus, enhance learning performance. The adoption of intelligent educational technologies and specialized software for processing learning-related data enables personalized learning [9–11]. The approaches to personalization include finding and recommending particular learning resources that meet students' characteristics [12] or using criteria to select the most appropriate educational methods for a specific learning context [13] to improve knowledge acquisition. Another approach is the creation and provision of learning resources personalized according to students' profiles, including prior knowledge and other learning-related features [14]. Ontologies can significantly help personalization and enhance e-learning.

An ontology is a formal, explicit description of a shared conceptualization of a domain. Usually, ontologies are built as a hierarchy of concepts and relationships and constraints, expressed by axioms in a machine-interpretable language to support semantic consistency, knowledge sharing, interoperability, and automated mathematical logic-based reasoning.

In education, ontologies can provide a structured model for organizing, representing, and sharing knowledge, defining the relationships among various concepts, disciplines, learning resources, and learner profiles, enabling a more organized, personalized, and efficient learning experience. Learning domain ontologies also have great potential to contribute to the content development of personalized educational resources [11]. In addition, ontologies can support semantic search and the discovery of learning content, as well as interoperability across educational systems. Thus, integrating ontology techniques can promote resource reuse and enhance their usage in intelligent e-learning systems.

Although innovative technologies enable improved educational services, updating the usual e-learning systems is challenging. Integrating ontology approaches into e-learning systems helps overcome information overhead and enables an easy-to-adapt, easily modified e-learning process, including intelligent functionality such as scalability, content reusability, and personalized learning [15]. Such information overload often requires students to exert extra cognitive load or mental effort while processing and navigating learning resources. In this context, ontologies can significantly reduce information overload by providing a structured semantic model for organizing, representing, and retrieving knowledge. Ontologies can help learners and educators manage complex information more efficiently and meaningfully by simplifying the organization of the adaptive and personalized learning and tutoring process. Ontologies are also valuable for supporting collaboration between resource developers, teachers in collaborative resource development, and learners in cooperative learning. They also help improve information retrieval for learners, including searching for external learning content or scientific news in the area of the learning domain. Teachers can also use ontologies to stimulate reasoning of available knowledge and, thus, a deep understanding of learning content.

Ontology development is difficult, time-consuming, and requires knowledge and engineering skills. Considerable efforts have been made to develop ontologies in many domains, such as Linguistics, Mathematics, Medicine, Chemistry, Agriculture, Geosciences, Education, and Computer science [16]. The ontologies created in almost all scientific fields can serve as initial domain ontologies in e-learning courses. Most of these ontologies are available online, but they are not suitable for direct usage in personalized learning. Some of these ontologies have been successfully used in e-learning for curriculum modeling and management, modeling tutoring strategies, learning domains, learners' metadata, and e-learning services, but only after significant modifications. Well-developed general domain ontologies are not aligned with specific curricula, educational standards, or the specific needs of learners. They also lack information about difficulty, prerequisites, and pedagogical dependencies. And most of them contain a large amount of knowledge,

unnecessary for every concrete course, that can lead to very big computational complexity, making them unusable in real-time applications. All these ontologies represent static knowledge, and real-time adaptation is needed before their usage. So, ontologies have great potential to improve learning and to dynamically adapt to frequent changes and dynamics in the e-learning domain, which is crucial for their use in practical intelligent educational environments.

Ontology development is difficult, time-consuming, and requires educational domain modeling, which is performed mainly by teachers or learning content developers. These specialists often are required to possess significant expertise in the knowledge engineering field, and they consider ontology development very challenging. Direct reuse of previously developed ontologies is not applicable, as each learning course has its own specifics. Moreover, some of the metadata used in personalized learning change dynamically during tutoring, and corresponding ontologies need to be updated accordingly. Therefore, in this field, it is essential to use ontology learning, which is one of the main factors in supporting the widespread adoption of ontologies in e-learning.

There is a wide variety of ontology learning approaches, methods, and techniques depending on the sources used and the requirements of the learned ontologies. This review analyses ontology learning approaches and techniques from the perspective of their usability in e-learning environments. The main aim is to outline the applicability of ontology learning approaches and techniques in e-learning and, thereby, facilitate the development and maintenance of educational ontologies. It presents the strengths and drawbacks of ontology learning approaches and techniques in the context of their usage for learning educational ontologies. Since the usual type of learning content is unstructured text, this article mainly discusses ontology learning from plain text. Thus, the research questions are as follows:

- RQ1: For the development of the types of ontologies that are used in e-learning, which ontology learning techniques are valuable and useful?
- RQ2: What are the significant specifics of ontology learning for usage in e-learning tasks in an educational context?
- RQ3: How can a combination of Large Language Models (LLMs) and ontology learning techniques make ontology development and ontology evolution in e-learning easier and effective?
- RQ4: For what types of tasks in e-learning are the (semi) automatically developed ontologies the most applicable or valuable?
- RQ5: What are the main trends and challenges of ontology learning for the e-learning domain?

The current review explores the role of ontology learning techniques. It analyzes their specifics and uses in e-learning, both for supporting the ontology maintenance process and directly applying knowledge extraction results in the learning and tutoring process. The article sheds light on the methods for automating ontology development, including the impact of recent advances of Artificial Intelligence (AI), as LLMs on ontology learning, which can improve the effectiveness and quality of e-learning services within an intelligent educational environment. Ontology learning will make ontology development and evolution more effortless and more efficient, which is of great importance for the educational domain. Dynamic (semi-) automated ontology development and maintenance can help provide personalized education by adapting the resource authoring and delivery to match individual student skills and preferences. This review outlines and discusses trends and challenges of ontology learning from textual e-learning content, web content, and databases, and its possible impact on knowledge modeling for personalized e-learning.

2. Methodology

E-learning as an interdisciplinary field comprises four main subfields of knowledge that can be modeled ontologically: learning content, scientific domain, pedagogical domain, learner description domain (including personal data, psychology, etc.), and technology domain. Since e-learning content is mainly stored in textual sources and web documents, and learner data is stored in databases, the focus of research is primarily on methods and approaches for ontology learning from text, databases, and web documents. Ontology learning refers to the automation of ontology development. The field of ontology learning encompasses various technologies, depending on the resources utilized and the types of ontologies, such as unstructured, semi-structured, and non-structured resources, where the range of applicable techniques is extensive [17].

The research focuses on the possibilities for automating the development of tutoring-domain ontologies and the automated generation and maintenance of learner profile ontologies. Thus, this review aims to identify key trends and developments in ontology learning and to discuss the applicability of the most useful ontology learning approaches and methods aligning with the specifics of the learning content used and the targeted type of ontologies needed for personalization in education. The research methodology was designed to correspond to the review goal and the research questions. The research process consists of three steps:

1. Formulating a search query;
2. Data retrieval, selection, and pre-processing;
3. Descriptive analysis and classification.

Firstly, the search was performed with the keyword "ontology learning" to gain an idea of scientific publications in this area. As a source of information, reputable scientific databases such as Web of Science (WoS), Scopus, and IEEE are used. Also, recent, scientifically valuable surveys and research papers found on Google Scholar are included. The search was performed in the titles, keywords, and abstracts of scientific papers published over the last 20 years, from 2006 to 2025, in English only.

The number of publications on ontology learning is significant (Table 1); however, the goal of this research is not to provide an exhaustive review of all these works. Considering the specific knowledge stored in e-learning systems that can serve as resources for ontology learning, and the knowledge about ontology learning methods that work well with these resources, the authors initially aimed to analyze ontology learning approaches specialized for these resources and discuss possibilities to adapt these general methods to the educational context.

Then, to gain a more detailed picture of ontology learning in the educational context, the authors formulated three search queries to find original research on this topic: SQ1: "ontology learning" and e-learning, SQ2: "ontology learning" and education, and SQ3: "ontology learning" and education and LLMs. Although all results retrieved by SQ3 theoretically are included within the broader result set of SQ2, some SQ3-related results are challenging to locate or may be inaccessible when searching directly using SQ2. Therefore, SQ3 enables more reliable retrieval of relevant publications on ontology learning for education that include LLMs.

It was challenging to find relevant research on ontology learning in the educational domain due to two factors: most research in ontology learning is not related to this domain, and most research on ontologies in education focuses on ontology usage or manual ontology development. Ontology learning in the e-learning domain is a specific research area. For many research papers, it is difficult to determine whether they are truly relevant, as the abstracts often do not clearly explain whether any automation is used in the ontology development process. Therefore, in many cases, a scan reading of the entire publication is

essential to classify a study precisely as ontology learning. This specificity makes the free availability of papers (i.e., open access) a significant factor for current research. For this reason, open access publications are noted in Table 1, which presents search results over scientific databases and digital libraries.

Table 1. Number of results of search queries by years and scientific databases.

Sources by Period / Search Query		"Ontology Learning"	SQ1 "Ontology Learning" and e-Learning	SQ2 "Ontology Learning" and Education	SQ3 "Ontology Learning" and Education and LLM
2006–2009	Scopus	270	8	41	0
	WoS	179	4	4	0
	IEEE	1197	196	207	0
	G. Scholar	3440	372	748	0
2010–2013	Scopus	315	11	12	0
	WoS	172	4	5	0
	IEEE	1240	210	237	0
	G. Scholar	4170	478	1060	0
2014–2017	Scopus	312	6	27	0
	WoS	225	3	8	0
	IEEE	1193	139	192	0
	G. Scholar	3810	448	1180	0
2018–2021	Scopus	265	3	8	0
	WoS	171	1	5	0
	IEEE	1268	91	185	0
	G. Scholar	3600	340	1280	4
2022–2025	Scopus	196	4	8	3
	WoS	89	0	4	2
	IEEE	1807	46	223	11
	G. Scholar	3080	307	1240	236
20 years 2006–2025	Scopus	1358/241 *	32/6 *	96/7 *	3/1 *
	WoS	836/169 *	12/0 *	26/2 *	2/1 *
	IEEE	6705	682/13 *	1044/48 *	11/0 *
	G. Scholar	18,100	1945	5508	240

* Number of publications with open access.

The search results in reputable scientific databases were few, and analysis of their titles and abstracts classified most of them as thematically irrelevant. Only 12 results seemed suitable. After a brief analysis of the publication's content, the authors selected four papers from Scopus, three from WoS as relevant to SQ1 and SQ2, and one paper relevant to SQ3. Then, a comprehensive search was conducted across the digital library of Google Scholar, which returns many more results (Table 1). Titles, snippets, and abstracts of the first 1000 returned results were carefully analyzed. A selection technique based on citation counts was used to identify the highest-quality, thematically related papers. It included browsing and exploring most citations of scientifically valuable (with an impact factor score, impact rank, or a considerable number of citations) and thematically relevant papers.

From Google Scholar results, 42 were selected based on the analysis of titles and abstracts. After reviewing these publications, only eight papers that comprehensively discussed ontology learning applied to the development of ontologies in the e-learning

field were found. The query SQ3 returned 236 results. Among them were very significant research papers and reviews on the application of LLMs to ontology learning, but none of them were directly related to the educational field. From these results, it can be inferred that the automation of ontology development in the e-learning area lacks sufficient research attention. The impact of advances in generative AI models for ontology learning, which support personalized education, was also not explored. As there were only a few studies on ontology learning in the e-learning field, the analysis performed and classification of ontology learning methods based on the types of resources used in education, and exploration of the impact of LLM advances on ontology learning methods using these resource types were the most significant parts of the research. The authors explored the possibilities of ontology learning techniques to support the development of educational ontologies and, specifically, the impact of LLM-based technologies on ontology learning methods suitable for automating ontology development in education.

3. Ontology Learning in the Context of the Educational Field

The usage of ontologies in e-learning and the applicability of ontology learning methods are closely related to the desired ontology properties. It is essential to classify ontologies based on the properties that are valuable for e-learning and ontology learning. Ontologies can be classified according to several dimensions, e.g., purpose, modeling domain, structure, and logical complexity.

3.1. Classifications of Ontologies in the e-Learning Domain

The classification of ontologies according to their modeling domain and purpose can be as follows: *domain ontologies, application ontologies, global (upper) ontologies, and hybrid ontologies*. Global ontologies usually contain well-known general terminology classification. Typically, developers use some of the available standardized upper ontologies (e.g., UFO DOLCE or WordNet) or parts of them when needed for general concept classification. Application ontologies that represent knowledge models specific to some applications require manual development by professionals. Ontology learning in the e-learning field is most frequently used for domain ontologies.

According to their structure and logical complexity, ontologies can be classified into the following categories: *taxonomies, ontologies with specific type relations or properties, logically rich ontologies, and populated ontologies*.

Particular methods exist for concept extraction, learning relations specific to the learning content type, or ontology population [18]. Therefore, suitable ontology learning techniques should be selected and used according to the target type of developed ontologies. Usage of ontologies in e-learning is task-specific. For example, domain-related taxonomies are frequently used for resource searching and personalized recommendations, whereas logically rich ontologies are most useful in learning and assessment.

In [19], Rahayu et al. classified ontologies based on their use in e-learning environments into five categories:

- Ontologies modeling the curriculum (e.g., modeling the relationships among learning objects, learning goals, and the objectives of the study program);
- Ontologies intended for data integration (e.g., to integrate knowledge in closely related domains);
- Ontologies describing domains and learning activities;
- Ontologies describing student profiles (as PAPI, LIP, or FOAF ontology);
- Ontologies that are developed for usage in tasks related to resource or tool recommendations for personalized learning.

Based on the previously discussed classifications, we can classify ontologies, used in education, thematically as follows:

- Tutoring domain ontologies (including simple taxonomies, modeling subdomains, complex relations, inter-domain relations, etc.);
- Task-specific ontologies (for resource recommendation, for personalization, based on learning performance, learning disabilities, etc.);
- Learner profile modeling ontologies (IMS LIP, IEEE PAPI, including behavioral data, learning preferences, disabilities, competences, motivations, etc.);
- Learning content structure modeling ontologies (related to e-learning standards, as IMS LD, SCORM, IMS CP, or organizing other specific metadata categories for learning objects);
- Pedagogical ontologies (for modeling teaching knowledge, including instructional methods, learning strategies and theories, pedagogical goals, teaching activities and roles, sequencing and instructional design logic, pedagogical constraints and dependencies, didactic models, learning theories, and educational principles, etc.).

This classification is essential for the current research, as some of the ontologies (as those modeling standards) represent clearly specified and rarely changed and standardized knowledge and should be well developed manually, but others, such as specific domain ontologies, include dynamically changed knowledge (most learning courses need periodical updates to their content, including terminology), and automating ontology maintenance is essential for course domain ontologies. It is also challenging to extract relationships among learning objects and learning goals automatically, storing them in task-specific ontologies. The best way to guarantee their correctness is to evaluate them manually. As learning domain knowledge and its sources are usually changed frequently, ontology learning will be the most useful for developing or evolving corresponding learning domain ontologies.

3.2. Analysis of Ontology Learning Techniques and Approaches and Their Possible Applications in the Educational Domain

Ontology learning is the automatic or semi-automatic process of developing an ontology by extracting information from various sources, including plain text, diagrams, databases, and social networks. Ontology learning research proposes techniques for constructing different types of ontologies and automating many ontology development steps to simplify the process of ontology building for experts or educators. Diverse classes of approaches are used, including statistical, data mining, logic-based, linguistic, machine learning, and deep learning approaches, among others. As the applicability of specific ontology learning methods closely relates to the available information sources, this review will analyze ontology learning methods from various types of learning content and learner information sources (such as textual e-learning content, diagrams, concept maps, databases, linked data), classify ontology learning techniques for essential types of information sources in the e-learning domain, and discuss its applicability for automating ontology development to support intelligent learning and tutoring.

After examining the scientific literature on the topic, the authors developed a classification of ontology learning from text techniques, which can be useful in the e-learning domain. Figure 1 presents the authors' two-dimensional classification of techniques for ontology learning from textual educational content (including pre-processing) based on ontology learning goals and the technologies used. Most of these techniques will be discussed in the following subsections.

This paper first systematizes ontology learning techniques, useful for the development of ontologies for intelligent tutoring, in three main categories according to the used sources: learning from unstructured sources (as plain text), learning from semi-structured sources

(as web documents, databases), and learning using structured sources (as UML diagrams, thesauruses, and other ontologies). As most learning content has a plain textual format, ontology learning from text is the most frequently used approach and is analyzed first, with learning objects' structural and grammatical specifics in mind.

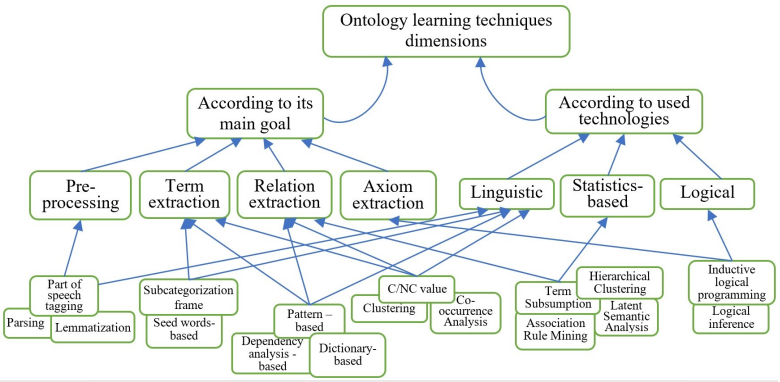


Figure 1. Classification of classical ontology learning from text techniques.

3.2.1. Techniques for Ontology Learning from Unstructured Text

Ontology learning [20] and enrichment [21] techniques from natural language text can be categorized into the following five main categories:

- Natural language processing (NLP);
- Machine learning;
- Statistical techniques;
- Data mining and information retrieval;
- Logic-based.

The authors visualize the step-by-step process of ontology learning from textual sources based on traditional approaches in Figure 2. The main steps are pre-processing, term extraction, concept formation, extraction of taxonomic and non-taxonomic relations, axiom extraction, and evaluation. Learning content is pre-processed using linguistic techniques such as parsing, lemmatization, and part-of-speech tagging. Syntactic parsing techniques are also used. Using linguistics or statistical methods helps to extract terms. Then, concepts are extracted using linguistic or statistical techniques, such as pattern-based techniques, C-/NC-value, and co-occurrence analysis. Subcategorization frames, latent semantic analysis (LSA), and clustering, or semantic lexicons, serve to extract taxonomic and non-taxonomic relations among concepts.

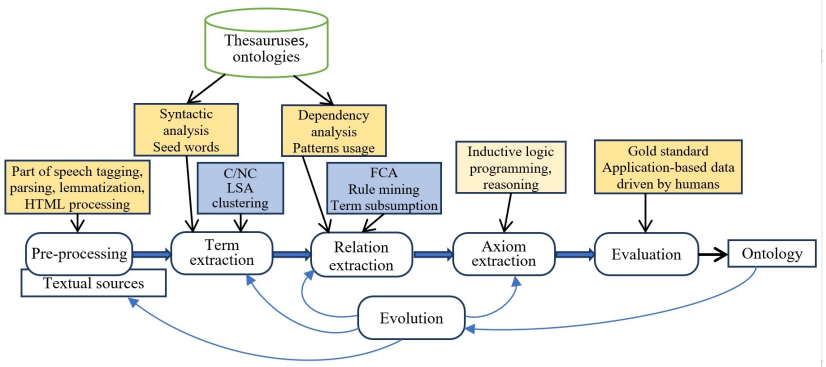


Figure 2. The process of ontology learning from text.

NLP techniques and statistical approaches, including Dependency analysis, lexico-syntactic analysis, term subsumption, formal concept analysis (FCA), hierarchical clustering,

and association rule mining (ARM), can be used at both the term/concept extraction and relationship extraction stages. Extracting axioms is mainly performed by reasoning or using Inductive Logical Programming (ILP).

Automatically developed ontologies usually contain terminological, linguistic, or logical errors or inconsistencies, and evaluation is an essential step in creating high-quality ontologies. Four main approaches (gold standard-based evaluation, application-based evaluation, data-driven evaluation, and human evaluation) and several evaluation measures (including precision, recall, and F-measure) are used to evaluate the integrity and quality of the developed ontology. The following subsections briefly discuss ontology learning technologies in the context of their application in the educational domain.

LLMs can perform most classical NLP tasks directly but without task-specific training. But LLMs do not replace all NLP techniques. They may generate plausible but incorrect facts, or LLM-based terminology extraction may lack deterministic precision. A discussion of the benefits and drawbacks of using LLMs for each of the following subtasks of ontology learning in the educational context is given below.

3.2.2. Ontology Learning from Textual e-Learning Content

There are two different tasks related to ontology learning: learning for developing new ontologies and learning for enriching and evolving previously developed ontologies.

Developing ontology from scratch is a complex, time-consuming, and labor-intensive task. Pre-existing seed ontologies are used frequently in the ontology learning process. There are also cases when dynamic changes in the existing ontology are needed. In e-learning, when the main aim is ontology enrichment, some of the above-mentioned steps can be applied independently of others. For example, if the targeted task is to compare the definition of a particular concept in the ontology and external textual resources, full pre-processing and lexical analysis are unnecessary. It is sufficient to identify the defined concept and its pattern, then extract properties and relations used in this definition and compare them with the definition in the ontological representation of the course.

When the main task is the development of a new ontology that describes learning content, the entire step-by-step ontology learning process should be followed (see Figure 2). However, in this case, ontology learning can be semiautomatic or interactive. In the e-learning domain, two contexts of interactive ontology learning exist: interacting with resource developers (teachers or experts) and interacting with learners. When the primary aim is to develop an ontology describing learning content, the developer can use ontology learning methods for a long time. In the case of ontology enrichment, some of the steps in Figure 2 can be applied independently of the others. However, it is necessary to manually verify the final results of ontology learning, as the quality of the domain ontology is essential in the learning domain. In particular, it is useful for some learners to develop small ontologies. This can stimulate comprehension and logical thinking. Ontology learning methods can provide guidance and suggestions to the learners during ontology development.

Pre-Processing of Textual e-Learning Content

Pre-processing textual sources using robust linguistic techniques is a prerequisite for ontology learning tasks [15]. Three main linguistic techniques are applied for pre-processing: part-of-speech tagging, parsing, and lemmatization. Preparing these steps is crucial for applying knowledge extraction techniques to natural language text. Usually, e-learning texts have a good structure, follow correct grammatical rules, and use restricted dictionaries. These facts can make pre-processing simpler and more efficient.

Considering a simple sentence structure of learning content, rule-based part-of-speech taggers, such as Brill Tagger [22], are effective pre-processing tools due to their better performance. Probabilistic taggers such as Tree Tagger [23] are rarely used. Parsing is a syntactic analysis of text aimed at identifying various dependencies between words in a sentence and representing them in a tree-based data structure called a parsing tree. GATE [24] and OpenNLP [25] are also good tools for pre-processing the learning content for ontology learning. Lemmatization is a linguistic pre-processing technique that brings words into their base form. The Stanford CoreNLP API [26] or WordNet-based Java Library [27] contribute to the lemmatization of textual data for ontology learning. LLMs can perform context-aware, semantic preprocessing that goes far beyond traditional rule-based methods.

Because they understand *meaning*, *syntax*, and *world knowledge*, LLMs can preprocess text based on context and language understanding, rather than rigid rules. LLMs can perform preprocessing tasks at high accuracy and, according to some researchers, often achieve higher F1 scores than those using classical pre-processing [28]. LLMs can generate shorter, clearer versions of text while preserving meaning. This is extremely useful before sentiment classification, topic modeling, semantic search, or information extraction. Also, LLMs can be used to extract topic-relevant nouns/phrases.

Traditional pre-processing techniques generally work well for e-learning content because of its precise semantic structure and correct grammar. Nevertheless, e-learning content can contain domain-specific elements such as formulas and definitions. Domain- or task-specific pre-processing is often better performed using LLMs. Another specific aspect of learning texts is the frequent usage of terminology in two or more natural languages. Extraction and recognition of multilingual terminology are essential for future steps in ontology learning, and, in this context, LLM-based techniques also are more useful. LLM-based pre-processing also can be used for text simplification and the removal of grammatical or terminological errors.

Linguistic Techniques for Knowledge Extraction

Linguistic techniques are used in many research projects to extract terms, concepts, and relations [16,17,19]. Syntactic structure analysis and subcategorization frames are useful for term extraction from well-structured texts, such as e-learning content. Dependency analysis and lexicon-syntactic patterns will also work well for relation extraction from e-learning content. Lexicons such as WordNet or other domain-specific web-based lexicons could also be used for concept extraction and relations, particularly for compound domain-specific terms. The extraction of domain-specific terms and concepts has improved by introducing seed words in the ontology learning pipeline.

The three main approaches used for relation extraction are based on Dependency analysis, patterns, and dictionaries. *Dependency analysis* uses parsing trees to extract relations between terms [29]. Dependency paths are used for finding relationship patterns [30]. They identify relations between two specific concepts by extracting the shortest path between those concepts in the parsing tree. Dependency analysis is an appropriate approach in e-learning because the extracted relations between concepts can be very useful in the learning process. The proposed educational content is good when it contains new, previously not shown, thought-provoking connections between concepts, and their explicit representation benefits learning. Research reveals that relationships between songs, extracted through Dependency analysis, can be successfully used for music recommendation [31]. Automatic extraction of dependencies between learning resources can be used to recommend external resources and support learning.

A *pattern-based approach* is rule-based, with rules often presented as lexical or syntactic patterns. A lexico-syntactic variant of this pattern-based approach is appropriate for

extracting taxonomic or non-taxonomic relations to support ontology learning. Regular expressions are used to extract well-known domain-independent or domain-specific relations. The pattern-based approach is easily applicable in the e-learning domain, as many domain-dependent patterns are known in advance, so there is no need to extract them using complex algorithms. When a precise evaluation of content similarity is required, well-known lexico-syntactic patterns for this content domain are used. The main limitations of pattern-based approaches are related to the need to create specific patterns. This task is time-consuming and challenging to maintain as the domain evolves. It requires domain experts and NLP experts to collaborate and can fail when texts deviate from the expected linguistic structure. It also has poor scalability across domains or subdomains, and this makes pattern reuse difficult.

Dictionary-based approaches have high precision, but their application in general lexicons such as WordNet is very limited, as usually domain-specific terminologies are not included in these dictionaries or have significant differences in meaning. Some domains, such as medicine [32] or mathematics, have good semantic lexicons. They offer a wide range of predefined concepts and relations and can be used to extract terms, concepts, and taxonomic and non-taxonomic relations. The semantic organization of terms in the lexicons, as a set of similar words (synsets) and predefined associations such as hypernymy, meronymy, etc., makes them very useful for (semi-) automatic ontology learning and for direct use by learners. The usefulness of dictionary-based approaches depends heavily on the available dictionaries. Finding a helpful dictionary or different linguistic resources in international languages such as English or Russian is easy. However, the availability of such resources in many other languages is limited. Therefore, educational content and curricula usually rely heavily on the following components: standardized terminology, controlled vocabularies, official curriculum frameworks, and competency frameworks, all represented in various natural languages. Dictionary-based approaches can be beneficial for building semantic ontologies, including all the terminology above, with high accuracy, reducing ambiguity (e.g., "function" in math vs. programming), aligning learning materials to official standards, supporting automatic curriculum mapping, and improving search, recommendation, and adaptivity in e-learning systems. The authors found many good papers related to ontology learning, using dictionary-based approaches, but none of them are especially directed to the development of ontologies for education. In the authors' opinion, most well-working dictionary-based ontology learning techniques are directly applicable for the development of tutoring domain ontologies.

Statistical Techniques for Ontology Learning and Their Application in the Educational Domain

Statistical techniques are like types of "black box" techniques. These techniques do not use concept semantics or relations between lexemes in the text, nor do they use semantic reasoning. Their main idea is that the co-occurrence of lexical units in a text often means that they are related or identical. Clustering, LSA, Association Rule Mining, Term Subsumption, Co-occurrence Analysis, and Contrastive Analysis are the most frequently used statistics-based ontology learning techniques. These techniques draw conclusions from statistical distributions of entity types across large amounts of previously selected texts (corpora) and do not consider underlying semantics. Such techniques require large, high-quality text corpora and are mainly used for term extraction, concept extraction, and taxonomic relation extraction. Frequently used statistical techniques for term/concept extraction are C/NC-value, Contrastive Analysis, Co-occurrence Analysis, LSA, and Clustering. In ontology learning, the algorithm LSA is used for concept extraction. The basic idea is that terms occurring together will be close in meaning.

The *C/NC-value technique* is particularly applicable to the extraction of multi-word terminology. It takes multi-word terms as input and returns a score for each of them [20]. This score combines two values—the C-value and the NC-value. The C-value tends to find a group of terms valid in the corpus. The NC-value considers the context of multi-word terms and tries to find longer strings that appear more frequently in the corpus [20]. The C/NC-value is a useful approach for term extraction in domains containing a significant amount of multi-word terms.

Contrastive analysis is a technique for filtering terms obtained through the term-extraction procedure. This technique uses two types of corpora: a relevant corpus (target domain) and an irrelevant corpus (contrastive domain). It also uses two measures of ontological learning—domain relevance and consensus. Such situations are occasional in the e-learning domain, and contrastive analysis is rarely used to develop learning ontologies in the e-learning domain.

Co-occurrence analysis is a concept-extraction technique that identifies lexical units frequently used together in texts. It applies to related-term extraction or to the identification implicit associations between various terms. Co-occurrences can appear on different levels: the phrase level, paragraph level, or document level. Various co-occurrence measures are used to evaluate the relationships between terms (e.g., Cosine Similarity, Dice Similarity, Mutual Information, etc.). Researchers report precision and recall of 60–70%. This approach can benefit learners in interactive ontology learning tasks. The ontology learning algorithm proposes related terms and asks learners about the type or correctness of the proposed relations.

LSA assumes that terms that occur together are close in meaning. LSA uses singular value decomposition on the term-document matrix as a mathematical technique. This approach is mainly used to find similar words and can be very useful for learners in interactive ontology learning to understand semantic relations in depth and find synonyms [33].

Statistical techniques for relation extraction include direct usage of statistical methods and machine-learning approaches. Term subsumption, FCA, Hierarchical Clustering, and ILP are statistics-based techniques used mainly for relation extraction during the ontology learning process. Term subsumption is used to extract hierarchical relations between terms. Hierarchical clustering is frequently applied to identify taxonomic relationships by grouping extracted terms into clusters based on similarity measures (e.g., Jaccard Similarity or Cosine Similarity). Two main strategies are used when building a cluster hierarchy—divisive clustering (top–down approach) and agglomerative clustering (bottom–up approach) [34].

Statistical methods rely mainly on term frequencies, co-occurrence, and distributional patterns. Usually, they do not capture comprehensive pedagogical meaning, e.g., prerequisite relationships between concepts, instructional intent, and cognitive difficulty levels. Educational materials use discipline-specific terminology and hierarchical concepts. Statistical techniques struggle to distinguish synonyms vs. related but different concepts. Ontologies frequently contain rich relations, and understanding complex dependencies is also very important in education, but statistical methods tend to detect only simple associations. Statistical methods often extract irrelevant terms or propose vague concept clusters or weak or spurious relations. Such an approach results in noisy ontologies that require extensive manual cleanup. Also, statistical methods cannot be easily adapted to evolving curricula. All this limits the usefulness of statistical methods for ontology development in adaptive educational systems.

Logic-Based Techniques for Ontology Learning in the Educational Domain

Logic-based techniques are mainly used to verify correctness and consistency and to extract relations and axioms. Finding contradictions in statements or relations between concepts can also be very useful for learners in the knowledge understanding process. The most frequently used logic-based techniques are inductive logic programming and logical inference [35].

Logical inferences are used mainly for extracting hidden relations from the existing ones in the ontology, using transitivity or inheritance rules. The main problem when using this method is extracting conflicting relations. Apart from ontology maintenance, logical inferences can be directly used in the tutoring process to support learners' logical thinking and active participation in the learning process.

ILP is a machine learning discipline that derives hypotheses from background knowledge and a set of examples using logic programming. In inductive logic, programming rules are extracted from a collection of concepts and their relations. A methodology for automatically building domain ontology using Deep Learning techniques to identify taxonomic and semantic relations between concepts is proposed in [36]. The relation classification model is trained using Wikipedia and WordNet through the distant supervision technique. Learners can be encouraged to compare all the automatically extracted hypotheses to their own conclusions on the related topics and discuss similarities or differences.

Logic-based techniques can ensure high semantic precision and full explainability. They are excellent for hierarchy extraction, determining prerequisites based on ordering in syllabi, modules, or textbooks, and applying pedagogical theories. Their use is restricted due to the need for specific expert-defined rules.

LLM-based knowledge extraction is a new approach that combines elements of classical knowledge extraction techniques but operates at a qualitatively different level. It is a new knowledge extraction paradigm that combines classical statistical NLP, logic-, and linguistically rule-based knowledge extraction [37]. An LLM-based approach mainly relies on statistical AI, but being a logic-based approach, it can infer implicit relationships. It can also be used to integrate background knowledge and to directly restructure text into knowledge graphs or ontology triples (in JSON, RDF, or OWL formats) [38,39]. So, this new approach combines lexical, syntactic, semantic, and pragmatic cues.

Hybrid classical techniques combine linguistic, statistical, and logical approaches to ontology learning, applying them in previously defined or flexible orders. Since various methods are most useful at different levels of the ontology learning process, a hybrid approach is the most common technique for the entire ontology learning process. Some automatic hybrid ontology-building strategies start from an initial (possibly small) ontology and extend it later through text processing. Most frequently, term extraction and statistical models are combined. Term extraction methods are first used to identify the essential terms from learning resources. Then, statistical models such as Term Frequency-Inverse Document Frequency (TF-IDF) or co-occurrence analysis are applied to determine the importance of each term within the domain. The importance helps to identify which terms should be included as classes or concepts in the ontology.

NLP-based and machine learning techniques can also be successfully combined. NLP-based techniques can be used for dependency parsing and named entity recognition (NER) to identify key entities and their relationships in a sentence. Then, machine learning models (e.g., supervised or unsupervised learning, such as clustering or classification) can be used to categorize these entities into higher-level concepts or to determine relationships between them.

Hybrid approaches are the most useful for automatic ontology extension during learning content evolution. Examples of statistical and rule-based technique combinations

include learning patterns using machine learning and then using the learned patterns as rules in the pattern-based approach. Ontology learning, supported by external semantic resources (seed words, thesauri, and other ontologies), is also a frequently used tool in e-learning, where most ontology learning tasks are of the ontology enrichment and evolution type.

Ontology learning from educational textual content aims to extract concepts, relations, learning objectives, prerequisites, skills, and pedagogical structures from learning resources (e.g., textbooks, syllabi, LMS content). Traditional techniques cannot extract missing (implicit) concepts, complex relations, and higher-level abstractions and struggles with synonyms/homonyms. LLM-based techniques can understand some context and conceptual meaning, extract fine-grained concepts, skills, and competencies, and identify synonyms, paraphrases, and pedagogical terminology. Some LLMs can also automatically convert text into structured knowledge. LLM-based techniques offer greater automation in relation extraction and ontology generation, but they cannot propose ontology validation, based on strict logical grounds.

3.2.3. Ontology Learning from Semi-Structured or Structured Sources

Ontology learning from semi-structured or structured sources refers to the automated or semi-automated extraction of concepts, relations, and hierarchical structures from structured data formats. Such structured sources mainly contain relational databases (tables, foreign keys), CSV files, spreadsheets, and existing knowledge bases (Wikidata, DBpedia). Semi-structured sources are XML, JSON, files, HTML pages (including tables, lists, infoboxes, and Web 2.0 sources), log files with formatting, Wikipedia infoboxes or template-based documents, and RDF datasets. Sources of most of listed types can contain valuable information for education.

Ontology learning from databases is a relatively new and prominent research area. Both relational and non-relational databases are appropriate. Some ontology learning approaches rely on the *automatic translation of UML diagrams into ontology*. There are well-working rules for such translation, but the logical complexity of the resulting ontology is highly restricted. Another ontology learning approach uses LOD. The process involves sending SPARQL queries to the LOD, analyzing the results, and, often, using them for further ontology development or updates. Due to the continuous enrichment of warehouses with open, linked data, this approach is gaining growing popularity. A significant problem in its use is the heterogeneity and dynamics of LOD, which make the possibilities for metadata extraction and corresponding ontology development challenging to predict.

Ontology Learning from Databases in the Educational Domain

Ontology learning from educational databases is an automated extraction of concepts, relationships, and hierarchies from (semi)structured data (e.g., tables, logs, learning analytics) to build or refine educational ontologies. Learning ontologies from databases involves establishing semantic correspondences between the data models and ontology-based knowledge models. In particular, when using relational databases in learning OWL ontologies, it is necessary to use the correspondences between the relational model and the ontology model, which are based on the relationships between the relational model and Descriptive Logic (DL).

Ontology learning from RDB includes two main tasks: constructing ontologies from RDB schema and extracting ontology instances from RDB data (ontology population) [40]. There are two main approaches to extracting the database schema for ontology building: transformation-based and mapping-based. They both work by applying rules. In [41], a set of rules is defined to analyze database components and convert them to corresponding on-

tology components. Researchers also define rules to explore and extract ontology elements from stored procedures, user-defined functions, views, multiple inheritance, and other database characteristics, treating these as constraints on tables and their columns. Some research explores relational and logical inference, as well as machine learning methods to extract instances from databases [42].

Many Learning Management Systems (LMSs) use relational databases to store data about learners. Ontology learning from databases can extract the semantics of the data stored in LMS databases and describe the learners' performance, navigation, or communication. The resulting ontology can then be used to support personalized learning. The main difficulties in ontology learning from databases stem from the relatively poor explicit semantics implemented in the relational database model. Standardized rules are used to map the following:

- Database tables onto OWL classes;
- Simple attribute to DatatypeProperty;
- Composition attribute to DatatypeProperty;
- Multi-valued attribute to DatatypeProperty;
- Primary key to DatatypeProperty;
- Bi-directional relationship to ObjectProperty;
- One-to-many relationships to OWL restrictions;
- Subtype relations (IS-A) to OWL:subClassOf.

However, the semantics of relationships and entity names in databases are usually unclear. A database schema is a relatively simple, incomplete model of related practical domains, and it is usually an insufficient resource for extracting domain knowledge to develop a domain ontology. Methods for extracting knowledge from databases are more useful for ontology enrichment, evolution, and population.

When applied to educational databases, ontology learning techniques can extract concepts, relationships, constraints, and semantics from LMS databases, Student information systems, Learning analytics dashboards, Assessment systems, Curriculum repositories, Course catalogues, and Library or resource databases, and they can generate, enrich, or actualize learning domain ontologies, learner profile ontologies, or resource description ontologies. This is of great importance for dynamically updating ontologies in Intelligent Educational Systems (IESs) in response to ongoing changes in educational systems. In this way, data (raw or processed using learning analytics) are represented as structured data and are added to knowledge structures that support semantic search, personalized recommendations, adaptive learning, and teaching/learning decision support.

Ontology Learning from UML Documents and Its Usage in the Educational Domain

Instructional systems rely on structured representations of knowledge, learning content, or learner models. UML class diagrams, activity diagrams, or use-case diagrams are used to describe domain concepts, relationships, workflows, or processes in some e-learning systems. UML diagrams can also contain inheritance hierarchies, process steps, SCORM metadata, MS Learning Design metadata, etc.

Ontology learning from the UML approach uses well-defined rules to translate UML diagrams into ontology, based on semantic correspondences between the two models. The UML-to-OWL translation algorithm maps UML symbols to OWL identifiers and UML elements to OWL axioms [43]. Some of the essential transformation rules are the following:

- A UML class is transformed into an OWL class;
- A UML association or association class between two or more classes is transformed into an OWL class;
- A UML attribute of a class is transformed to an OWL datatype property;

- A UML role associated with a UML association and a UML class is transformed to an OWL object property between the two OWL classes;
- A UML generalization set is transformed to a set of OWL class axioms (i.e., subClassOf type axioms);
- The disjointness constraints are transformed to OWL DisjointClasses.

UML descriptions help represent the structure of learning resources in a simple graphical way. Standard rules work well for transforming a UML description into an OWL ontology. For example, such transformations perform a semantic description of Educational Adaptive Hypermedia [44]. A semantic description of the learning content structure resulting from translation enables the automation of the organization of the personalized learning process. The transformation of UML content diagrams of tutoring content developed by students can also be used for automated assessment of students' knowledge.

Ontology learning from UML diagrams can be used to improve interoperability between systems by automating the creation of a formal semantic model (from a UML diagram to an ontology), enabling different e-learning systems to share data more effectively and exchange it. Ontologies, quickly generated from UML diagram can also allow for the structured storage of e-learning content, such as courses, assessments, student progress, etc. Such ontologies can also be used to recommend courses or assessments to students based on their prior knowledge or learning path. Students' evaluations can also benefit from methods for creating ontologies from UML. Students can make UML diagrams during the assessment, which can be converted to ontologies, and evaluation can be automated by a comparison of these ontologies to domain ontologies, modeling the course knowledge. Despite the various possibilities for using ontology learning from UML in education, the authors found no research papers related to the extraction of educational ontologies from UML documents.

Ontology Learning from Web Sources and Its Applications in the Educational Domain

Web 2.0 semi-structured resources from social media (folksonomies, XML documents, social networks, web services, etc.) can be used for ontology learning. User-generated content, collaboration, tagging, social interaction, social networks, e-commerce-related data, and web services messages give essential information for personalized learning. Tags created by learners show how they conceptualize information, which can guide adaptive systems. Tags also include information about the learner's interests and vocabulary. Learners' connections, groups, or interactions also reveal their educational interests and capabilities. Folksonomy mining and text mining from Web 2.0 content, social network analysis, learning path extraction, and methods for analyzing learner behavior are specific techniques for information extraction from Web 2.0. Using thematic searches, the authors found some publications related to learning educational ontologies from Web 2.0, but these do not contain significant results.

E-learning 2.0-based educational systems are most frequently used in higher education. These systems are flexible regarding resource adoption, knowledge acquisition, and the use of cloud computing. A rich dataset for learners' descriptions is stored and used in Web 2.0-based e-learning systems. In such systems, combining Web, Semantic Web, and social network-based technologies is easy. There are only a few studies on ontology learning from social web data [18,45], and none of them are related to e-learning. Some of these methods do not use HTML markup. These methods remove it and then process the extracted plain text. Thus, valuable information about the concepts and their relations discussed on webpages is discarded and not used. Other methods use HTML tags but do not exploit the entire document structure and paths. The HTML pages are handled by extracting semantics from titles, subtitles, bold, italic, underlined texts, tags, hyperlinks, lists, etc.; the related

full text is also used for extracting knowledge. Digital textbooks, educational websites, MOOCs, forums, and open knowledge repositories are also web sources. They provide a rich base for extracting domain concepts, relationships, and pedagogical structures. So, web sources include various types of resources that can enrich several educational ontology types, and a wide variety of ontology learning techniques can be useful for automating ontology development.

Research [45] proposes Xhtml TREE Mining methods (called XTREEM) for ontology learning from web documents, exploiting the specific domain and language-independent semi-structure of webpages. Combining these methods, manual selection of high-quality e-learning content from the web, and some NLP-based techniques, can yield good results in automating ontology development and evolution. Other semi-structured resources used for ontology learning are Wikipedia, Webopedia, dictionaries, and thesauruses such as WordNet. Wikipedia is a rich source of knowledge for automatic processing. Many research projects on ontology extraction [46–51] use the Wikipedia corpus and its category system as a large-scale taxonomy. Wikipedia's pages are also used to disambiguate terms. In [50], Tramontana and Verga propose a well-working algorithm for expanding domain ontologies using open semi-structured resources such as Wikipedia. Bhatt's research suggests a dictionary-based ontology learning method for developing multilingual ontologies [51].

A well-defined structure of textual HTML documents can help with terminology and relation extraction. It enables the simultaneous application of methods for ontology learning from text and methods that exploit the structure and presentation of web documents.

Despite a wide variety of sources and possible techniques for automated development of educational ontologies using web sources, the authors found only two significant research studies in this area. One study [52] proposes a mechanism for the automatic construction of concept maps from online discussion forums, which can be used in an e-learning environment. Another paper [53] presents a semi-automatic domain ontology development process based on knowledge extraction from existing SCORM educational content found in online educational systems. These studies are relatively old and, in our view, not very significant. So, research on educational ontology learning from the web is limited.

Challenges in ontology learning from web sources are related to data and content quality (web data can be noisy, inconsistent, and unreliable, making it difficult to extract meaningful concepts and relationships); domain ambiguity (terms in the educational domain can have multiple meanings); a need for the integration of heterogeneous sources (e.g., academic papers, blogs, educational platforms); dynamic and evolving content; and semantic heterogeneity.

To summarize, ontology learning from web sources can provide educational systems with dynamic, context-aware, and personalized knowledge. By applying automatic extraction of concepts and relationships from vast web-based resources, ontologies can evolve more efficiently, offering tailored learning experiences and more accurate knowledge representation. However, developers must address challenges such as data quality, integration, and scalability to ensure the effective use of ontologies in e-learning systems.

Using the Linked Open Data (LOD) Cloud for the Development of Ontologies for Intelligent Tutoring

The LOD cloud (including simple RDF resources, dictionaries, and thesauri) contains a large amount of freely available knowledge, represented in machine-processable formats (RDF, RDFs, OWL, OBO, etc.) as thesauri or linked ontologies. This knowledge is easily usable as background knowledge by many ontology learning methods in support of knowledge extraction from texts or other types of learning objects to create more enriched ontologies. The knowledge contained in the LOD cloud spans almost all scientific domains and can support ontology learning in the context of every e-learning course. There are

also suitable ontologies covering most scientific domains that can serve as initial variants for the development of course or learning content ontologies. By harnessing the power of linked data, ontologies can be created or dynamically managed to structure, organize, and enrich learning content. Thus, a more personalized, context-aware, and efficient learning experience is achieved, reducing information overload and helping students engage with the learning material in a meaningful way.

The main advantages of using LOD to automate the development of ontologies for e-learning are interoperability, access to rich knowledge, up-to-date data, and scalability (LOD datasets are designed to scale and handle large amounts of information). LOD can support ontology development by enhancing domain knowledge representation, by linking to open educational resources, linking with pedagogical data in LOD, dynamically pulling in related articles, videos, or exercises from LOD datasets to provide additional learning support, and linking to sources like Linked Data for Education, which could offer frameworks for assessment models, etc.

The most relevant LOD datasets for education and tutoring systems are DBpedia, Wikidata, YAGO, Open Educational Data Repositories, Competency, and Skills Ontologies Linked as LOD. DBpedia provides concept definitions, type hierarchies, semantic relations (e.g., part-of, subclass, relatedTo), and entity links to other datasets, which are very useful for building domain ontologies in almost all areas. Wikidata is a rich, human-curated knowledge graph comprising multilingual labels, well-structured statements, and fine-grained educational domain concepts. Open Educational Data Repositories contain education-focused datasets, OpenCourseWare data, Curriculum vocabularies, LODE (LOD for Education), and BabelNet (multilingual lexical and semantic resource). All these datasets and vocabularies can be very useful for the rapid development of educational ontologies, but the authors could not find serious research on educational ontology learning based on the LOD cloud.

Challenges in using LOD for the development of e-learning ontologies are data quality and consistency (not all LODsets are of the same quality). Some datasets may have incomplete or inconsistent data, integration complexity, and dynamic, evolving data. This is due to the LOD data changing over time, requiring constant updates and maintenance of the ontology to ensure consistency.

Ontology Reuse for the Development of Ontologies for Intelligent Tutoring

As the development of ontologies from scratch is difficult, labor-intensive, and time-consuming, and some previously developed ontologies can be found in almost every learning domain, finding and using appropriate ontologies is essential for easier ontology development. Search engines can help in finding suitable initial versions of learning domain ontologies, and ontology learning can support its primary adaptation. SPARQL queries are also useful solutions for extracting knowledge from previously developed ontologies, downloaded from the Internet, or created in other courses for intelligent e-learning. The facts obtained from these queries can support ontology extension, evolution, or evaluation.

Ontology mapping, merging, or partitioning can also support the development of educational domain ontologies using previously developed ontologies. For example, the ontology model for e-learning activities and actions can be mapped to the terminological domain ontology to define essential relationships between course terminology and learning activities.

Usually, ontology reuse in Intelligent Tutoring Systems (ITSs) involves ontology maintenance and is performed in the following steps: identify relevant existing ontologies; analyze the compatibility of existing ontologies with the ITS's requirements; adapt or extend existing ontologies; and evaluate and test the resulting ontology. Ontology learning

is used mainly in the extension phase. Some well-known ontologies that could be reused or adapted for ITS are the IEEE Learning Object Metadata (LOM) Ontology [54], useful for describing learning materials, resources, and metadata, and the Competency Ontology (COMP) [55], useful for competencies modeling and skills development in students.

Ontology reuse is a powerful method for automating the development of e-learning ontologies. By leveraging existing ontologies from domains such as pedagogy, student modeling, and domain knowledge, ITS developers can create more robust, scalable, and efficient systems. However, careful adaptation and integration are necessary to ensure the ontologies meet the specific requirements of the ITS and remain flexible enough to accommodate evolving educational practices.

Challenges in ontology reuse for ITSs include domain-specific variability, integration issues (combining multiple ontologies from different domains can lead to conflicts or redundancies), and scalability (some ontologies might not scale well when dealing with large, complex educational environments or massive amounts of student data).

Using LLMs for Automating Learning Domain Ontology Development

LLMs can assist in almost all steps of ontology learning from e-learning content, including Concept Extraction (automatically identifying and extracting key concepts and terms from e-learning content) and Relation Extraction (discovering relationships between concepts, identifying hierarchical, associative, and dependency relationships, etc.). LLMs can also understand and process natural languages. This capability can ensure the extraction of definitions and descriptions of concepts directly from the text, enabling the understanding of terms in the e-learning domain and supporting automated semantic modeling. LLMs can also synthesize information from multiple sources to identify overarching themes and concepts or to identify differences in the meaning of closely related terms. LLMs can also be used to classify and cluster educational content into relevant categories or topics, thereby proposing suitable texts for specific ontology learning tasks. For tasks such as generating semantic annotations for educational content, LLMs can also be beneficial. These annotations link terms, concepts and their definitions, interrelationships, and contextual uses, thus enriching the ontology with semantic layers. Additionally, LLMs can propose related terms, synonyms, or alternative terms, considering the context, which can supplement the ontology.

Using LLMs for ontology learning in the context of e-learning is a promising approach, thanks to leveraging advanced NLP capabilities, which can significantly improve the efficiency and accuracy of ontology building. Thus, LLMs can automate the extraction of concepts, relationships, and knowledge from large sets of unstructured educational data, making them highly valuable for the e-learning domain. The limitations of LLM-based techniques are lower transparency, harder-to-justify decisions, and specific needs in interaction (prompt design, structured extraction formats).

3.3. Evaluation of Learned Ontologies in the e-Learning Domain

Automatically developed ontologies usually do not accurately represent all the knowledge from the sources used for ontology learning. They may contain terminological, linguistic, or logical errors or inconsistencies, so evaluation is essential for the development of high-quality ontologies. There are four main approaches to evaluating learned ontologies: gold standard-based, application-based, data-driven, and human evaluation. Human evaluation and the gold standard are used frequently in e-learning. Gold standard-based evaluation is applicable, as complete ontologies have recently been developed across many fields, and parts of these ontologies are used to evaluate the quality of ontologies learned from learning resources.

Ontology evaluation tasks can have two main goals: examining an ontology based on its characteristics (white box evaluation) and measuring its overall performance on a specific task (black box or task-based evaluation). White box evaluation focuses on two directions. The first is empirically evaluating how many (of all) components are learned correctly (empirical evaluation). The second is to consider (and, if possible, correct) the precision or specifics of the learned model by checking each learned element (model-based evaluation). This empirical evaluation is essential when the initial ontology is learned from textual learning sources. However, during ontology enrichment in the e-learning domain, quality evaluation of models for selected entities is predominantly performed. In e-learning, modeling every concept or relation is essential. It can be defined in a specific way in line with learning goals, so model-based evaluation by teachers or experts is highly significant. Usually, humans perform this type of evaluation. Teachers, learning content developers, or even students can make human-based evaluations. Student evaluations during ontology learning or interactive ontology development can make evaluation cheaper while also benefiting the learning process. The drawback of this approach is that the quality of the evaluation needs to be guaranteed, and teachers or resource developers should conduct the final evaluation.

4. Overview of Existing Research on the Automation of Ontology Development for E-Learning

E-learning is a complex domain that embraces e-learning standards; pedagogical, psychological, technological, and tutoring domains; and learner-described knowledge. This specific complexity makes the application of ontology learning methods more difficult. To the authors' knowledge, the current study is the first attempt to survey the application of ontology learning to support ontology development in the e-learning domain. Table 2 summarizes research on the automation of ontology development in the e-learning domain. NLP algorithms are frequently used to support the extraction of the concept–relation–concept triple from textual resources.

Table 2. Summary of the research on the automation of ontology development in e-learning.

Paper	Source	Type of Contribution	e-Learning Subdomain	Ontology Type	Used Methods	Evaluation in Education	Year
[54] **	Slides	Approach	Tutoring Domain	Concept map	Concept extraction	Visualization of course contents	2017
[48] **	Books	Framework	Tutoring Domain	Concept hierarchy	Combined	High school physics	2021
[55] **	Textual learning content	Method	Tutoring Domain	Evolution of some ontology	Pattern-based	Not presented	2011
[56] **	Heterogeneous documents	Method	Tutoring Domain	Concept hierarchy	WordNet	Reference ontology	2011
[52] **	Online forums	System	Tutoring Domain	Concept hierarchy	Fuzzy domain ontology extraction	Reference ontology	2008

Table 2. *Cont.*

Paper	Source	Type of Contribution	e-Learning Subdomain	Ontology Type	Used Methods	Evaluation in Education	Year
[53] **	SCORM educational content	Method	Tutoring Domain	Ontology	WordNet	Not presented	2009
[57]	Heterogeneous text documents	Method	Tutoring Domain	Concept hierarchy	Fuzzy domain ontology extraction	Reference ontology	2013
[58]	Textual learning content	Method	Tutoring Domain	Concept hierarchy	Pattern-based	Not presented	2010
[59]	RDB of LMS Moodle	Method	User profile	OWL ontology	Transformation rules	In LMS Moodle	2013
[60]	Various	Review	Resource recommendation	Learner profile, learning domain	Combined	Proposes and discusses evaluation metrics	2023
[61]	Neural network-based	Method	Learner profile	Learner profile	Deep learning	Evaluation in agent-based system	2020
[62] *	Lecture slide text	Approach	Tutoring Domain	Educational ontology	LLM-based	Not presented	2024
[63]	Various	Review	Tutoring domain	Domain ontology	LLMs-based	Not presented	2025
[64] *	Various	Review	Tutoring domain	Knowledge graphs	Combined classical	Focused on knowledge graphs	2024
[65] **	Learner data	Method	Learner profile	Learner ontology	Data mining	For personalized learning	2020
[66]	Various	Review	Tutoring domain	Knowledge graphs	Combined classical	Focused on knowledge graphs	2025

The symbol * indicates indexation in the scientific database Scopus, and ** indicates indexation in Web of Science. Papers without these symbols are not indexed in these databases and are found in Google Scholar.

Atapattu et al. investigated the effectiveness of automated approaches in extracting concepts and auto-generating concept maps from lecture slides [54]. Experts evaluated auto-generated concept maps. The research presents the development of a set of NLP algorithms to support the extraction of the concept–relation–concept triple from tutoring presentations. The natural layout of the lecture slides was also used to help organize extracted concepts in a hierarchy. Thus, various applications can use auto-generated concept maps, including knowledge organization and reflective visualization of course content.

A framework for automatically constructing and extending Educational Domain Ontology, called 'ADOL', is proposed in [48]. This ontology learning framework can automatically convert domain textbooks into a corresponding ontology. This was tested in a high school physics course, and researchers found it feasible and efficient. A semi-automatic interactive ontology learning process to facilitate domain ontology enrichment is presented in [55]. Web information sources, such as the Glossary of Programming Terms Used in C++ and lexicosyntax patterns, can support the extraction and recognition of needed relationships from textual resources.

In [56], Gaeta et al. introduce a hierarchical clustering algorithm to derive a concept hierarchy from the textual learning content. Then, background knowledge from WordNet was used to label the extracted hierarchy and detect synonyms. The effectiveness of this method was tested in seven domains: tourism, art, user modeling, problem-solving, databases, and information systems. The F-measure across the experiments ranged from 0.69 to 0.85.

Research by Lau et al. proposed a mechanism for generating a concept map using a fuzzy domain ontology extraction algorithm [52]. Such a mechanism constructs concept maps by extracting information from messages posted to online discussion forums. The context-sensitive text mining method and the fuzzy domain ontology extraction algorithm were applied to automatically generate concept maps that represented the knowledge structures of the learning content to support students' learning.

SCORM-based hierarchical organization of the learning content in learning objects was used during ontology generation [53]. Navigation rules were applied during ontology creation in correspondence with SCORM rules. For example, a SuggestedOrder relation between two concepts was generated if the item corresponding to the source concept had a Sequencing Control Choice=True. Statistical and data mining algorithms were implemented to identify concepts and their relationships. Transformation rules were used in [59] to build an OWL ontology from the Relational Database used by an LMS Moodle to store information about learners, learning content, and the entire tutoring process. Database-to-ontology transformation rules were enriched by analyzing stored data to detect constraints on disjointedness and totality in hierarchies. Data analysis was also used to recover some missing aspects during the mapping of the conceptual data model to the relational model. The generated ontology also had non-taxonomic relations.

A study [61] presented a robotic multi-agent system that constructed an ontology to analyze student learning behavior in the context of English speaking and listening. A deep neural network (DNN) method was used. The agent integrated three types of intelligence: perception, computational, and cognition.

Recent advancements in LLMs have offered a novel opportunity to automate and refine ontology learning, yet the authors found only one related study in the educational domain [62]. This research applied LLMs to process lecture slide texts for domain ontology extraction and used them to enhance student performance predictions.

Another study [63] discusses two publications related to ontology learning, LLMs, and e-learning. However, a key drawback of both publications is that they do not focus explicitly on educational applications. Instead, they primarily investigate ontology learning for other purposes and merely propose, at a conceptual level, that the learned ontologies could be applied in educational contexts, without providing empirical validation or concrete educational use cases. For example, other research [67,68] investigated how LLMs can support and automate collaborative ontology engineering by gradually shifting from human-centered to LLM-centered workflows, and the authors discussed the usage of the generated ontologies in education.

Recent research on semantic knowledge modeling in education presents a systematic literature review that examines ways of constructing knowledge graphs (KGs) and their applications in the field of education [64]. It highlights the methodologies used to build educational KGs, knowledge extraction techniques, and how knowledge graphs are used across five key educational domains—including adaptive and personalized learning, curriculum design, concept mapping, and semantic search. The research covers automated and semi-automated knowledge graph construction techniques used in education, but the discussion is high-level and descriptive, rather than deeply analyzing ontology/knowledge graph learning algorithms or LLM-based automation.

Other research [66] used data mining techniques to analyze learner data and support the enrichment of learner ontology. The ontology was used as a conceptual backbone to structure learner interactions.

All the research publications on automating ontology development in the e-learning domain presented in Table 2 concern learning ontologies labeled in English. This is because most scientific research is conducted and published in English, as an international language. According to the authors' experience, bilingual ontologies are often more helpful than monolingual ontologies in e-learning. The following section presents a use case illustrating the advantages of bilingual ontologies.

5. Use Case and Discussion

5.1. Use Case

One of the most valuable specifics of ontology learning in the educational field is the possibility of using semi-automatic ontology development or evaluation as a learning task. Such an approach can both reduce the cost of developed ontologies and deepen the understanding of learning content. To demonstrate how ontology learning can enhance the educational process, a small use case on ontology learning from textual content in a programming course at the Technical University of Sofia is presented.

Domain ontologies developed to represent scientific knowledge can be used to support e-learning tasks but are usually composed of many concepts and propose general knowledge models. However, it is essential to represent a correct model of learning content in the same way as is presented in learning resources. In a simple use case, the authors tested the usage of linguistic techniques for ontology learning from a textual content in a programming course. This course was in Bulgarian and also included English-language terminology. Initially, a small Java programming ontology was developed and used in testing two approaches: monolingual (learning ontology in Bulgarian) and bilingual (Bulgarian and English-language terminology, as well as external English-language textual sources and a dictionary).

Initial versions of the small course domain ontologies describing the programming course content were developed using terminology extraction from text and then manually evaluated first by two advanced students (as a final learning task) and then by the teacher. Linguistic analysis of the text was performed using the GATE framework, and Protégé was used for ontology maintenance. GATE is open source and oriented to English texts, so the authors added a restricted set of Bulgarian-specific linguistic rules to support the linguistic analysis of Bulgarian texts. In addition, during semi-automatic ontology development, we used external sources such as WordNet, Wikipedia, and the web-based Glossary of Computer Programming Terms. Table 3 shows the results.

Table 3. Comparison of monolingual and bilingual ontology learning.

Content	Precision	Recall	Correct Learned Hierarchies, %	Fully Described Concepts, %
In Bulgarian	0.84	0.92	72%	54%
Bilingual	0.92	0.95	83%	47%

The developed course domain ontologies were lightweight linguistic ontologies that contained 82 concepts and 34 relations each. The extracted terminology was categorized (by the teacher) into three categories to make ontologies helpful for personalizing e-learning: prerequisites (defined carefully by the teachers, using manually added properties), terminology defined and explained in the course content, and terminology related to future or

additional learning (also manually marked by the teachers). Only the terminology of the second category was explained comprehensively in the ontologies. Students were asked to evaluate the correctness of this terminology and propose corrections, if needed.

As the results show, most of the terminology was extracted correctly in general, but full descriptions of concepts (using properties or relations) were extracted approximately half as well. Students were interested in discussing and correcting such incompletely extracted knowledge. This activity helped students understand and remember the course terminology.

Based on the use case, the conclusions about the usage of ontology learning in e-learning are as follows:

- Ontology learning can help with initial educational domain ontology development, but manual evaluation and maintenance are critical for developing high-quality ontologies.
- The use of bilingual resources, including widely used international language content and dictionaries, can ensure higher quality of the learned ontology due to cross-linguistic connections and specific strategies for using more high-quality language resources.
- Involving some advanced students in interactive ontology learning and evaluation tasks is engaging and useful for them and can reduce the time and effort it takes for professionals to complete the evaluation.

After development, the course domain ontology can be mapped to other domain ontologies that model related courses to allow for an explanation of the terminology used in this course (but not explained) with relevant definitions or descriptions.

During tutoring, every course evolves in some way (by manual addition of new content, by removal, or by recommending new external sources to some students). The most effective way to update the ontology describing continuously evolving content is dynamic ontology learning. For example, when a new textual learning object is added to the course, ontology learning from the text should be performed to extract new concepts from the text and insert them in the appropriate place in the course domain ontology. NLP techniques work well for knowledge extraction from unstructured or semi-structured textual e-learning content, as these resources are grammatically correct and have a clear linguistic structure. Pattern-finding techniques work well with e-learning content, as this type of textual content has a well-known structure (e.g., contains definitions, explanations, and examples). As e-learning content is grammatically correct and exploits implicit grammatical knowledge presented in texts (i.e., morphological, syntactic, and semantic information), it usually helps extract knowledge from the textual learning content. Recently, ontology learning by usage of LLMs has become a very active and emerging research area, based on earlier ontology learning approaches and machine learning [61]. LLMs also can be used for the text analysis and semantic processing of learning content in Bulgarian. Some advanced learners can be actively involved in the ontology maintenance process, contributing to ontology enrichment and evaluation by using interactive ontology development tools, visualizing ontology elements, and suggesting corrections. And teachers should perform the final evaluation of added corrections in the evolved ontologies.

LLMs are especially effective in education, as educational texts are well-structured, pedagogically structured, and context-dependent. On the other hand, LLM-generated ontologies can be plausible but can contain errors or logical contradictions, which is unacceptable in education. So, interactive development and evaluation of automatically generated ontologies by LLMs with experts, teachers, or learners is a must. LLMs can help in ontology learning from educational data by supporting almost every step of the ontology learning pipeline, especially where traditional NLP struggles with pedagogical meaning [63]. LLMs can support interactive semi-automatic ontology development, in

which some students can take part. Educational knowledge is curriculum-dependent, pedagogically constrained, and continuously evolving. Interactive ontology development, evolution, and evaluation will ensure the quality of developed ontologies, understanding, and trust among the learners and teachers involved. LLMs can significantly assist in all the previously discussed steps of knowledge extraction from textual e-learning content, including concept identification, relation extraction, and instance classification. Learners can participate, and this can be highly beneficial. Learners can confirm whether concepts are understandable, can flag confusing or ambiguous definitions, and can suggest missing concepts from a learner perspective. Learners can help identify incorrect prerequisite relations and misjudged difficulty levels. Learners can also participate in collaborative ontology learning (learning-by-building), in which they actively construct parts of the ontology, interacting with each other or with LLMs. Such activities improve the cognitive understanding of tutoring content.

The current use case is intended solely to illustrate the bilingual term extraction aspect and collaborative ontology development by involving students. It also shows that some advanced students like to participate in interactive ontology development, thereby reducing costs and helping achieve specific tutoring goals. This use case was conducted as a "proof-of-concept" study, and we acknowledge that there is still much work to be done on ontology learning. The practical application of all other ontology development techniques in education requires experimentation in an advanced intelligent e-learning environment or advanced intelligent tutoring system. Application of the reviewed techniques to complex educational scenarios across different contexts will be a future research goal.

5.2. Discussion

Recent developments in ontology technologies and their use to support the learning process and provide high-level services to teachers and students have contributed to the development of intelligent educational environments. In general, education is a complex field that requires the use of specific and frequently changing ontologies. Automation of ontology development is vital for supporting personalized learning content development. Learning domain ontologies are an essential part of adaptive resources, and their automated acquisition makes resource development easier, cheaper, and faster. Ontology learning also plays a significant role in resource recommendation, enabling automatic ontology evolution according to changes in the learning content or learner characteristics during the learning process.

There is a substantial amount of research on ontologies in e-learning, and it is a well-established area in educational technology and semantic computing, but, according to the authors' experience, the usage of ontologies in practical e-learning systems is rare. There is plenty of research literature on ontologies. They are considered useful, but they have limited application in practical, real-world e-learning systems (e.g., Moodle, Canvas, Blackboard, Google Classroom). There are several reasons for this gap, with high development costs being the main ones. Thus, automation of ontology development is crucial for its practical usage. The primary motivation of this study was to analyze approaches and methods for automating ontology development and to discuss their practical usage in the e-learning domain. The analysis reveals a wide variety of methods and techniques for ontology learning, but its application in education has important specifics.

There are only a few studies on ontology learning in the educational domain, and most of them are relatively old (see Table 2). Table 2 also shows an increasing interest in ontology learning in the educational domain over the last 3–4 years, due to the impact of LLMs on the automation of ontology development. The presented use case is intended only to illustrate the bilingual term extraction aspect and collaborative ontology development by involving

students. It also shows that some advanced students like to participate in interactive ontology development, which reduces the cost of ontology development and helps achieve particular tutoring goals. There is an increasing need for the practical application of all the other techniques discussed in this survey. Such an application can be performed in a well-developed intelligent tutoring system.

5.2.1. Finding Summary

This survey aimed to identify the development of the types of ontologies used in e-learning, as well as which ontology learning techniques are valuable and useful (RQ1). Since a significant part of educational content is presented as well-structured, grammatically correct textual sources, the linguistic and pattern-based approaches for ontology learning from text, combined with LLM-based techniques, can benefit initial domain ontology development, evolution, or enrichment. In this context, the recently developed framework for ontology learning, ADOL [48], which automatically and efficiently converts domain textbooks into a corresponding ontology, looks promising. So, some LLMs can do, but reasoner-based and expert/teacher manual evaluation is strongly recommended before using automatically generated ontologies. LLMs are better for the fast generation of simple short classifications (OWL hierarchies or concept maps) of learning domain terminology but have limited capabilities for extracting complex relationships. For extracting complex dependency patterns, pattern-based approaches can give better results. Text mining and clustering approaches are rarely used for ontology learning in the educational domain due to the lack of large static and well-annotated textual corpuses.

Ontology learning from UML models is typically used to generate structural, conceptual ontologies in specific domains, where the static structure of systems, or process descriptions, are used. UML diagrams contain classes, essential attributes, and relationships, so the ontologies automatically derived from them propose valuable domain representations.

Ontology learning from web sources is typically used to automatically build or enrich mainly learning-domain ontologies, but some social media sources are very useful for the enhancing or populating of learner profile ontologies. Human-in-the-loop ontology engineering should be used for pedagogical ontology development rather than ontology learning, as pedagogical logic is challenging to infer automatically. Only some LLM-based techniques are useful in the automation of pedagogical ontology development.

Pedagogical knowledge can also be learned from LMS databases, student information systems, and learning analytics databases. Ontology learning from database techniques can be combined with LLMs, which can interpret table semantics, propose pedagogical abstractions, or generate ontology hierarchy or axioms from data patterns. In this process, human validation via experts or educators is a must.

ARM and pattern-based techniques are most useful for automating learner profile ontologies. Techniques, based on information extraction from databases, are also very useful for learner profile ontology maintenance (mainly for enrichment and population) as LMSs store large amounts of structured learner data, including course enrolments, assessment scores, activity logs, learning paths, skill acquisition records, etc., in their databases. Automating ontology development is also useful for populating domain ontologies and learner ontologies by extracting data from databases.

The LOD cloud is primarily used to develop and enrich domain-level, knowledge-rich ontologies by mainly modifying free domain ontologies, available across almost all scientific domains. Some popular learner profiles or pedagogical ontologies are also available, but an adaptation of these ontologies for every system or course is needed.

Linguistic, dictionary-based, and pattern-based classical ontology learning approaches are most frequently used for learning domain ontologies, as textual learning content is

the most common. Ontology learning and the population of learner ontologies from the social web or learner-described data stored in educational systems are essential for personalization in response to dynamic changes in learners' states. Methods for ontology learning that utilize data from databases, UML diagrams, or the social web, tailored to the specifics of data stored in educational systems, should be enhanced and used more frequently for intelligent tutoring or assessment in the future. Cross-language algorithms for ontology learning also work well for learning domain ontologies. Using ontologies is the most important for personalized tutoring and recommendation tasks. However, reasonable classifications and logically rich ontologies can also be used directly to support learning, thinking, and comprehension. So, strengthening axiom-learning algorithms is also needed to produce suitable ontologies for e-learning.

Next, this paper explores the significant specifics of ontology learning in education (RQ2). Learner profiles can vary across different personalized learning systems and require professional development by experts or teachers. However, in many cases, ontology learning from databases can be very useful for learner ontology maintenance or population, as many LMSs use relational databases to store learner-related data. Ontologies modeling pedagogical information are also specific to e-learning systems and need professional development. Suitable resources for the automatic acquisition of these types of ontologies are rarely available. Therefore, ontology learning does not apply to pedagogical ontologies. The reuse of previously developed ontologies is the only way to make the development of pedagogical ontologies easier. However, professionals should make changes and evaluations. A significant aspect of the semi-automatic development of educational domain ontologies is the possible participation of some learners. This can make the development of some domain ontologies cheaper, as their maintenance can be part of the educational process. Also, involvement in semi-automatic ontology development and evaluation can be very useful for a deep understanding of the learning content.

RQ3: How can a combination of LLMs and ontology learning techniques make ontology development and ontology evolution easier and more effective in e-learning?

With LLM-assisted term extraction, relation suggestion, and concept clustering, much of the manual labor related to language processing can be automated. It is particularly valuable for educational domains where pedagogy, topics, concepts, or learners' properties evolve often. LLMs can build a first version of small learning domain ontology, which is very useful. This ontology can then be evaluated and enhanced manually by professionals or used by learners in learning or assessment tasks. As disciplines evolve (new topics, methods, terminology), some LLMs can help keep ontologies current without enormous human effort. LLMs contribute to ontology development. Modern LLMs offer high-quality term and concept extraction, synonym detection, and the identification of related concepts, subtopics, and pedagogical concepts (e.g., misconceptions, skills, and difficulty levels). This reduces the need for manual domain analysis. LLMs can also propose the extraction of some prerequisite relationships, part-of hierarchies, similarity relations, and conceptual dependencies. It accelerates ontology schema development and reduces the workload for experts. LLMs can generate initial ontological models, including concepts and relations. Ontology learning methods complement LLMs capabilities by providing formal grounding. Ontology learning can be used to confirm, adjust, or reject some of these components, achieving both creativity and precision.

Learners can directly use LLMs during activities related to ontology development and evaluation. Teachers can also use LLMs during manual ontology maintenance or evaluation. LLMs can monitor new content or changes in a knowledge domain or learning content and propose updates.

Integrating LLMs into ontology learning yields a significant advancement in ontology engineering in education. LLMs offer high-level semantic understanding, linguistic reasoning, and rapid structure generation, while ontology learning, followed by logical reasoning and validation, can add complex relationships and specific elements, helping correct mistakes. Together, they enable the efficient, scalable, and adaptive development of educational ontologies, which are essential for intelligent e-learning environments.

Furthermore, this research seeks to identify the types of e-learning tasks for which the automation of ontology development is applicable and important (RQ4). The overview in previous sections revealed that one of the most natural applications of domain ontologies, learned or enriched from the learning content, is in recommender systems [19]. Some course ontologies can automatically be extended using related external resources that particular students like or use during their learning. These enriched ontologies can then be used to recommend resources to other students with similar needs or preferences. Learning needs and available resources are dynamic and require continuous changes in content describing ontologies. This work is complex and very expensive when performed manually. Therefore, in this case, ontology learning is essential. In addition, subject ontologies can assist the automated dynamic generation of e-learning exercise problems or tests based on the learner's knowledge and performance. Automatically learned domain ontologies can also support the automatic semantic annotation of learning content during resource development. Semantically annotated content is the most significant resource for organizing personalized learning, including tasks such as content recommendation and mapping learning content to dynamically generated learning paths.

Ontology learning is very useful for adapting the learning process to the vast ever-changing array of educational resources. It also helps ensure interoperability between e-learning platforms, tools, and content formats.

According to the authors, ontology learning and the population of learner ontologies from the social web or learner-described data stored in educational systems are essential for personalization based on dynamic changes in learners' states. Therefore, such ontology learning methods should be developed in the future. LOD, Wikipedia, WordNet, and other dictionaries are also very useful for automating ontology extraction and maintenance for learning. These resources are public, freely available, and cover more and more natural languages and scientific domains. They can support automated terminology extraction from textual learning content, synonym or homonym recognition, and the development of multilingual ontologies. Therefore, ontology learning from web-based linguistic and scientific sources is very promising.

Ontology learning can benefit e-learning by enabling easier, faster ontology development and maintenance, supporting the organization of personalized learning, enabling rich, contextualized knowledge representations, improving content discovery, supporting interoperability and reusability of educational content, and also enhancing collaborative learning.

The comprehensive analysis reveals that the main challenges in the ontology learning field are related to the following issues (RQ5): the diversity and complexity of educational content, ambiguity of educational terminology, scarcity of high-quality labeled data, interdisciplinary knowledge representation, difficulties in extracting pedagogically meaningful relations, dynamic nature of educational knowledge, learner-centered aspects, and evaluation specifics. Educational resources come in many formats: textual content, videos, lecture slides, quizzes, student discussions, LMS logs, etc. They are also related to many different subjects (math, literature, and physics) that use vastly different approaches but are restricted by tutoring strategies and goals, vocabularies, and conceptual structures. This requires the use of complex information extraction methods, adapted to the specifics

of every course, and the results need to be evaluated by professionals. The scarcity of high-quality, labeled data related to the educational content makes the use of statistical ontology learning approaches difficult; modern LLM-based approaches fit well in this area. The heterogeneity of learning content (due to the variety of subjects and resource formats) makes the automatic extraction of concepts, relations, and hierarchies difficult. Many terms in education are domain-dependent and context-dependent. For example, different definitions of the same concept can be used across different levels of courses on the same subject. Disambiguating terms automatically require deep contextual understanding. Curricula, standards, learning objectives, and teaching methods change regularly. Ontologies must be updated frequently; automating this process is not easy. Version control and consistency maintenance are non-trivial. Educational content often integrates multiple subjects (e.g., STEM projects). Building cross-domain ontologies that avoid overlap and inconsistencies is complex. Learner-centered aspects are hard to formalize. Education involves more than domain knowledge—learning styles, learning difficulties, preferences, cognitive load, and skill progression over time. Modeling of all such knowledge about learners requires professional development and standardization.

Other challenges are related to the quality of automatically developed ontologies. Ontology learning is still a developing field, and every learning approach has some precision and recall below 100%, as shown in the presented use case. Since the complete ontology learning process is a sequence of stages, each of which depends on the results of the previous stage, minimizing errors at each step is crucial. As errors at each ontology learning stage are normal parts of the development process, the results of ontology learning need evaluation. Learners can perform evaluation tasks during learning, but professionals should make a final evaluation after every automatic change to prevent errors.

Ontology learning is the most valuable for developing and evolving tutoring domain ontologies, as they are the most commonly used in adaptive intelligent e-learning systems and are closely related to the learning content of every course. Logically simple and terminology classification ontologies are easy to learn and most often used for supporting resource recommendation tasks and the organization of personalized learning. Changes in learning domain ontologies during course evolution are the most frequent, and additional resources for supporting learning domain ontologies (including textual and web-based content) are usually available. Learner, pedagogical, and resource description ontologies can be specific to a given course and, at the same time, represent data according to established standards. Therefore, previously developed standardized ontologies can serve as initial versions, and ontology learning can help evolve and populate these ontologies.

5.2.2. Limitations

This survey explored research on ontology learning methods in the educational context. It is limited to the analysis of approaches for automating the development of ontologies applicable in e-learning. Next, the survey restricted the research to full-text access (mainly to open access proceedings papers and journal articles) written in English. However, since most scientific research is published in English, the findings, analyses, and conclusions can serve as the basis of further in-depth exploration. Another limitation is the discussion of only the methods for ontology development from text-based sources, such as plain text, textual databases, and textual web documents. The survey did not discuss ontology learning from videos, pictures, or other graphical sources.

5.2.3. Future Directions of Research

Based on this survey, the authors outline future research related to automating ontology development for educational purposes. Future work will include the exploration

of collaborative ontology learning (including in a multilingual environment) in which students actively participate in ontology construction and refinement. The use of LLMs can help generate initial versions of concepts and relationships (including draft OWL ontologies) from learning materials, and students can validate, correct, and enrich these results as a part of their learning tasks. Traditional extraction techniques (e.g., pattern-based or statistical methods) can serve as baseline or verification mechanisms to ensure semantic consistency. In this context, the authors plan to make a detailed comparison of traditional knowledge extraction techniques and LLM-based approaches and identify effective strategies for their combined use in ontology learning within intelligent e-learning environments. The authors will prove the hypothesis that LLMs have difficulties in extracting complex relationships and the hypothesis about the possibilities of LLMs to extract pedagogical knowledge. Questions about selecting the most appropriate LLM model for conducting ontology learning experiments (including model training experiments, aiming to enhance LLMs' ontology learning capabilities) are also very interesting.

Given the encouraging results related to multilingual ontology learning, the authors plan to explore bilingual or multilingual collaborative ontology learning scenarios in which students work with content in different languages. LLMs can assist with translation and cross-lingual concept alignment, while traditional methods ensure terminological accuracy using mathematical, logic-based formal reasoning.

The next future work concerns the detailed evaluation of how collaborative ontology learning supported by LLMs affects student learning outcomes. A comparative study could assess concept understanding, knowledge retention, and critical thinking when students engage with LLM-assisted ontology building versus traditional learning activities. Another promising direction is studying how student interactions, such as edits, annotations, and disagreements, can drive ontology evolution. LLMs can analyze interaction logs to suggest improvements, enabling continuous and data-driven ontology development, whereas traditional techniques have structural constraints. The authors will work on integrating these techniques into practical e-learning systems.

6. Conclusions

In contemporary intelligent e-learning environments, ontologies can support structured domain knowledge and competency model presentation, classification of learning resources, learner modeling (skills, mastery, preferences), and adaptive and personalized learning. This complex domain needs better ontology development, evolution, and evaluation techniques. Traditional ontology development is mainly manual and time-consuming and requires knowledge modeling skills. Automatically developed ontologies can contain errors and inconsistencies and are difficult to update when curricula or knowledge domains change. The authors' analysis reveals that most existing ontology learning techniques are not evaluated and tested in the educational domain.

Most frequently, ontology learning is used for learning domain ontologies. Textual educational content and many other external textual or semantic resources are suitable sources for learning domain ontologies. During the tutoring process, these ontologies often have updates, making automation in ontology development essential. Automated acquisition of ontologies from learning content can make adaptive resource development and maintenance considerably easier and cheaper.

To manage large amounts of educational content in a dynamic, evolving environment, scalable methods are necessary. Still, most ontology learning methods are developed and tested using specific types of learning content. However, extending the ontology to accommodate content with a different structure or using different terminology is difficult. Automated identification of the exact meanings of terms in an educational context is of great

importance for learning, but terms can sometimes be ambiguous, especially in multilingual environments. It is challenging to identify the exact meaning of most of these vague terms during ontology learning, so humans should participate in the process of validation or fine-tuning. Easy-to-use, intuitive, and personalizable interfaces for ontology visualization are needed to support human involvement in ontology development. The integration of ontologies into existing e-learning platforms (e.g., LMS, courseware) is also a challenge. It can require significant adaptation and technical work.

The specificity of the e-learning domain can help to overcome some of the difficulties in ontology learning. For example, validation by learners, including social web participants, collaborative tagging, and folksonomy is helpful for both ontology learning and simplifying ontology evaluation. In intelligent e-learning environments, the semi-automatic and interactive ontology acquisition and evaluation process is applicable, and it tends to produce suitable ontologies. It can benefit learning and help reduce costs while increasing the quality of ontology development.

The complexity of domain knowledge in some subjects can be challenging to represent in an ontology. For instance, advanced topics in engineering, medicine, mathematics, or law require a deep understanding of both content and context, or very high accuracy in knowledge representation. Ensuring that ontologies are compatible with various e-learning platforms, tools, and content formats is also a challenge. Standardization of ontologies for e-learning is crucial for seamless integration across systems. Another challenge is related to learner privacy; collecting data on learner profiles and interactions to build personalized experiences must be performed carefully to protect privacy and comply with legal regulations.

There is interest in multilingual and multicultural ontologies for education to support cross-language, cross-region sharing of learning resources. For example, ontologies for multilingual educational resources have been flagged as a promising direction.

Ontology learning is very important for ontology development and maintenance. There are significant specifics of ontology learning for educational ontologies, yet only a few studies address it in the educational domain. Thus, there is a strong need for targeted research on ontology learning to support the automation, evolution, and evaluation of ontology development, particularly for educational ontologies. Fully automatic ontology generation cannot ensure ontology quality in education, but the involvement of students can, in many cases, be useful both as an educational activity and as part of ontology development and maintenance. With the development of LLMs, ontology learning applying this innovation is a highly promising approach.

Author Contributions: Conceptualization, T.I. and V.T.; methodology, T.I. and V.T.; investigation, T.I.; data curation, T.I.; writing—original draft preparation, T.I. and V.T.; writing—review and editing, T.I. and V.T.; visualization, T.I. and V.T.; funding acquisition, T.I. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been accomplished with financial support by the European Regional Development Fund within the Operational Programme "Bulgarian National recovery and resilience plan", procedure for direct provision of grants "Establishing of a network of research education institutions in Bulgaria", and under project No. BG-RRP-2.004-0005 " Improving the research capacity and quality to achieve international recognition and resilience of TU-Sofia(IDEAS)".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Acknowledgments: The authors would like to thank the Research and Development Sector at the Technical University of Sofia for the financial support. This research is also supported by the project UNITe BG16RFPR002-1.014-0004 funded by PRIDST.

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

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ARM	Association Rule Mining
FCA	Formal Concept Analysis
ICT	Information and Communication Technologies
IES	Intelligent Educational Systems
ILP	Inductive Logical Programming
ITS	Intelligent Tutoring Systems
LLMs	Large Language Models
LMS	Learning Management Systems
LOD	Linked Open Data
LOM	Learning Object Metadata
LSA	Latent Semantic Analysis
NLP	Natural Language Processing
TF-IDF	Term Frequency-Inverse Document Frequency

References

1. Zhu, Z.T.; Yu, M.H.; Riezebos, P. A Research Framework of Smart Education. *Smart Learn. Environ.* **2016**, *3*, 4. [CrossRef]
2. Rico-Bautista, D.; Medina-Cardenas, Y.; Coronel-Rojas, L.A.; Cuesta-Quintero, F.; Maestre-Gongora, G.; Guerrero, C.D. Smart University: Key Factors for an Artificial Intelligence Adoption Model. In *Advances and Applications in Computer Science, Electronics and Industrial Engineering*; García, M.V., Fernández-Peña, F., Gordón-Gallegos, C., Eds.; AISC; Springer: Singapore, 2021; Volume 1307, pp. 153–166. [CrossRef]
3. Huang, L.-S.; Su, J.-Y.; Pao, T.-L. A Context Aware Smart Classroom Architecture for Smart Campuses. *Appl. Sci.* **2019**, *9*, 1837. [CrossRef]
4. Iqbal, H.M.N.; Parra-Saldivar, R.; Zavala-Yoe, R.; Ramirez-Mendoza, R.A. Smart Educational Tools and Learning Management Systems: Supportive Framework. *Int. J. Interact. Des. Manuf.* **2020**, *14*, 1179–1193. [CrossRef]
5. Ilić, M.; Mikić, V.; Kopanja, L.; Vesin, B. Intelligent Techniques in E-Learning: A Literature Review. *Artif. Intell. Rev.* **2023**, *56*, 14907–14953. [CrossRef]
6. Popchev, I.P.; Orozova, D.A. Towards Big Data Analytics in the E-Learning Space. *Cybern. Inf. Technol.* **2019**, *19*, 16–24. [CrossRef]
7. Terzieva, V.; Ilchev, S.; Todorova, K.; Andreev, R. Towards a Design of an Intelligent Educational System. *IFAC-PapersOnLine* **2021**, *54*, 363–368. [CrossRef]
8. Ilchev, S.; Alexandrov, A.; Ilcheva, Z. Design of a Laser Projection System for Intelligent Learning Environments. In *Proceedings of International Conference on Data Science and Applications*; Saraswat, M., Roy, S., Chowdhury, C., Gandomi, A.H., Eds.; LNNS; Springer: Singapore, 2022; Volume 288, pp. 89–103. [CrossRef]
9. Peng, H.; Ma, S.; Spector, J.M. Personalized Adaptive Learning: An Emerging Pedagogical Approach Enabled by a Smart Learning Environment. *Smart Learn. Environ.* **2019**, *6*, 9. [CrossRef]
10. Bontchev, B.; Antonova, A.; Dankov, Y. Educational Video Game Design Using Personalized Learning Scenarios. In *Computational Science and Its Applications—ICCSA 2020*; Gervasi, O., Murgante, B., Misra, S., Garau, C., Blečić, I., Taniar, D., Apduhan, B.O., Rocha, A.M.A.C., Tarantino, E., Torre, C.M., et al., Eds.; LNTCS; Springer: Cham, Switzerland, 2020; Volume 12254, pp. 829–845. [CrossRef]
11. Ivanova, T.; Terzieva, V.; Ivanova, M. Intelligent Technologies in E-Learning: Personalization and Interoperability. In Proceedings of the International Conference on Computer Systems and Technologies '21, Ruse, Bulgaria, 18 June 2021; ACM: New York, NY, USA, 2021; pp. 176–181.
12. Blagoev, I.; Vassileva, G.; Monov, V. A Model for E-Learning Based on the Knowledge of Learners. *Cybern. Inf. Technol.* **2021**, *21*, 121–135. [CrossRef]

13. Trichkova-Kashamova, E.; Paunova-Hubenova, E.; Boneva, Y.; Dimitrov, S. Criteria and Approaches for Optimization of Innovative Methods for STEM Education. In Proceedings of the 22th IFAC Conference on Technology, Culture and International Stability (TECIS 2024), Sofia, Bulgaria, 11–13 September 2024; IFAC Papers Online. Elsevier: Waterford, Ireland, 2024; Volume 58, pp. 123–128. [CrossRef]
14. Villegas-Ch, W.; García-Ortiz, J. Enhancing Learning Personalization in Educational Environments through Ontology-Based Knowledge Representation. *Computers* **2023**, *12*, 199. [CrossRef]
15. Kaur, P.; Sharma, P.; Vohra, N. An Ontology Based E-Learning System. *Int. J. Grid Distrib. Comput.* **2015**, *8*, 273–278. [CrossRef]
16. MaduraiMeenachi, N.; Sai Baba, M. A Survey on Usage of Ontology in Different Domain. *Int. J. Appl. Inf. Syst. (IJAIS)* **2012**, *4*, 46–55. [CrossRef]
17. Khadir, A.C.; Aliane, H.; Guessoum, A. Ontology Learning: Grand Tour and Challenges. *Comput. Sci. Rev.* **2021**, *39*, 100339. [CrossRef]
18. Konys, A. Knowledge Systematization for Ontology Learning Methods. *Procedia Comput. Sci.* **2018**, *126*, 2194–2207. [CrossRef]
19. Rahayu, N.W.; Ferdiana, R.; Kusumawardani, S.S. A Systematic Review of Ontology Use in E-Learning Recommender System. *Comput. Educ. Artif. Intell.* **2022**, *3*, 100047. [CrossRef]
20. Asim, M.N.; Wasim, M.; Khan, M.U.G.; Mahmood, W.; Abbasi, H.M. A Survey of Ontology Learning Techniques and Applications. *Database* **2018**, *2018*, bay101. [CrossRef]
21. Iyer, V.; Mohan, L.; Bhatia, M.; Reddy, Y.R. A Survey on Ontology Enrichment from Text. In Proceedings of the 16th International Conference on Natural Language Processing, Hyderabad, India, 18–21 December 2019; NLP Association of India. International Institute of Information Technology: Hyderabad, India, 2019; pp. 95–104.
22. Lande, D.V.; Dmytrenko, O.O. Using Part-of-Speech Tagging for Building Networks of Terms in Legal Sphere. In Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems (COLINS 2021), Kharkiv, Ukraine, 22–23 April 2021; pp. 87–97.
23. Aparna, K.; Bhakta, P.; Vijaykumar, S. A Review on Different Approaches of Pos Tagging in NLP. In Proceedings of the Information Technology & Bioinformatics: International Conference on Advance IT, Engineering and Management—SACAIM-2022, Ulaanbaatar, Mongolia, 28–29 October 2022; Volume 1, pp. 47–51. [CrossRef]
24. Blandón Andrade, J.C.; Zapata Jaramillo, C.M. Gate-Based Rules for Extracting Attribute Values. *Comput. Sist.* **2021**, *25*, 851–862. [CrossRef]
25. OpenNLP—Apache OpenNLP Library. Available online: https://opennlp.apache.org (accessed on 22 December 2025).
26. Stanford CoreNLP API. Available online: https://stanfordnlp.github.io/CoreNLP/api.html (accessed on 22 December 2025).
27. WordNet-Based Java Library. Available online: https://github.com/extjwnl/extjwnl (accessed on 22 December 2025).
28. Braga, M.; Milanese, G.C.; Pasi, G. Investigating Large Language Models' Linguistic Abilities for Text Preprocessing. *arXiv* **2025**. [CrossRef]
29. Sen, S.; Tao, J.; Deokar, A.V. On the Role of Ontologies in Information Extraction. In *Reshaping Society Through Analytics, Collaboration, and Decision Support*; Iyer, L.S., Power, D.J., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Volume 18, pp. 115–133. [CrossRef]
30. Kang, S.; Patil, L.; Rangarajan, A.; Moitra, A.; Jia, T.; Robinson, D.; Ameri, F.; Dutta, D. Extraction of Formal Manufacturing Rules from Unstructured English Text. *Comput.-Aided Des.* **2021**, *134*, 102990. [CrossRef]
31. Byeon, H.; Chunduri, V.; Narang, G.; Alghayadh, F.Y.; Soni, M.; Ramesh, J.V.N. Deep Learning Model for Recommendation System Using Web of Things Based Knowledge Graph Mining. *Serv. Oriented Comput. Appl.* **2025**, *19*, 57–76. [CrossRef]
32. Unified Medical Language System. Available online: https://www.nlm.nih.gov/research/umls/index.html (accessed on 22 December 2025).
33. Jiang, X.; Tan, A. CRCTOL: A Semantic-based Domain Ontology Learning System. *J. Am. Soc. Inf. Sci.* **2010**, *61*, 150–168. [CrossRef]
34. Zepeda-Mendoza, M.L.; Resendis-Antonio, O. Hierarchical Agglomerative Clustering. In *Encyclopedia of Systems Biology*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 43, pp. 886–887.
35. Ismail, R.; Abu Bakar, Z.; Abd Rahman, N. Extracting knowledge from English Translated Quran using NLP Pattern. *J. Teknol.* **2015**, *77*, 67–73. [CrossRef]
36. Panchenko, A.; Faralli, S.; Ruppert, E.; Remus, S.; Naets, H.; Fairon, C.; Ponzetto, S.P.; Biemann, C. TAXI at SemEval-2016 Task 13: A Taxonomy Induction Method Based on Lexico-Syntactic Patterns, Substrings and Focused Crawling. In Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), San Diego, CA, USA, 16–17 June 2016; Association for Computational Linguistics: Stroudsburg, PA, USA, 2016; pp. 1320–1327.
37. Mukanova, A.; Milosz, M.; Dauletkaliyeva, A.; Nazyrova, A.; Yelibayeva, G.; Kuzin, D.; Kussepova, L. LLM-Powered Natural Language Text Processing for Ontology Enrichment. *Appl. Sci.* **2024**, *14*, 5860. [CrossRef]
38. Zhu, Y.; Wang, X.; Chen, J.; Qiao, S.; Ou, Y.; Yao, Y.; Deng, S.; Chen, H.; Zhang, N. LLMs for Knowledge Graph Construction and Reasoning: Recent Capabilities and Future Opportunities. *World Wide Web* **2024**, *27*, 58. [CrossRef]

39. Babaei Giglou, H.; D'Souza, J.; Auer, S. LLMs4OL: Large Language Models for Ontology Learning. In *The Semantic Web—ISWC 2023*; Payne, T.R., Presutti, V., Qi, G., Poveda-Villalón, M., Stoilos, G., Hollink, L., Kaoudi, Z., Cheng, G., Li, J., Eds.; Springer Nature: Cham, Switzerland, 2023; Volume 14265, pp. 408–427. [CrossRef]
40. Ma, C.; Molnár, B. Ontology Learning from Relational Database: Opportunities for Semantic Information Integration. *Vietnam J. Comp. Sci.* **2022**, *9*, 31–57. [CrossRef]
41. Lakzaei, B.; Shamsfard, M. Ontology Learning from Relational Databases. *Inf. Sci.* **2021**, *577*, 280–297. [CrossRef]
42. Lin, L.; Xu, Z.; Ding, Y. OWL Ontology Extraction from Relational Databases via Database Reverse Engineering. *J. Softw.* **2013**, *8*, 2749–2760. [CrossRef]
43. Xu, Z.; Ni, Y.; He, W.; Lin, L.; Yan, Q. Automatic Extraction of OWL Ontologies from UML Class Diagrams: A Semantics-Preserving Approach. *World Wide Web* **2012**, *15*, 517–545. [CrossRef]
44. Papasalouros, A.; Retalis, S.; Papaspyrou, N. Semantic Description of Educational Adaptive Hypermedia Based on a Conceptual Model. *J. Educ. Technol. Soc.* **2004**, *7*, 129–142.
45. Brunzel, M. The XTREEM Methods for Ontology Learning from Web Documents. In Proceedings of the 2008 Conference on Ontology Learning and Population: Bridging the Gap Between Text and Knowledge, Patras, Greece, 21–22 September 2008; IOS Press: Amsterdam, The Netherlands, 2008; pp. 3–26.
46. Kawakami, T.; Morita, T.; Yamaguchi, T. Building up Ontologies from the EnglishWikipedia and Comparing with YAGO. *Trans. Jpn. Soc. Artif. Intell.* **2020**, *35*, C-J32_1-14. [CrossRef]
47. Gorodetsky, V.; Tushkanova, O. Learning an Ontology of Text Data. In *Advances in Fuzzy Systems and Soft Computing: Selected Contributions to the 10th International Conference "Integrated Models and Soft Computing in Artificial Intelligence" (IMSC-2021), Kolomna, Russia, 17–20 May 2021*; CEUR Workshop Proceedings; CEUR-WS: Aachen, Germany, 2021; pp. 1–8.
48. Chen, J.; Gu, J. ADOL: A Novel Framework for Automatic Domain Ontology Learning. *J. Supercomput.* **2021**, *77*, 152–169. [CrossRef]
49. Navarro-Almanza, R.; Juárez-Ramírez, R.; Licea, G.; Castro, J.R. Automated Ontology Extraction from Unstructured Texts Using Deep Learning. In *Intuitionistic and Type-2 Fuzzy Logic Enhancements in Neural and Optimization Algorithms: Theory and Applications*; Castillo, O., Melin, P., Kacprzyk, J., Eds.; Springer International Publishing: Cham, Switzerland, 2020; Volume 862, pp. 727–755. [CrossRef]
50. Tramontana, E.; Verga, G. Ontology Enrichment with Text Extracted from Wikipedia. In Proceedings of the 2022 5th International Conference on Software Engineering and Information Management (ICSIM), Yokohama, Japan, 14–16 January 2022; ACM: New York, NY, USA, 2022; pp. 113–117. [CrossRef]
51. Bhatt, B. Unsupervised Multilingual Ontology Learning. In Proceedings of the 2019 14th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP), Macau, China, 20–22 November 2019; IEEE: Chiang Mai, Thailand, 2019; pp. 1–7. [CrossRef]
52. Lau, R.Y.K.; Song, D.; Li, Y.; Cheung, T.C.H.; Hao, J.-X. Toward a Fuzzy Domain Ontology Extraction Method for Adaptive E-Learning. *IEEE Trans. Knowl. Data Eng.* **2009**, *21*, 800–813. [CrossRef]
53. Capuano, N.; Dell'Angelo, L.; Orciuoli, F.; Miranda, S.; Zurolo, F. Ontology Extraction from Existing Educational Content to Improve Personalized E-Learning Experiences. In *Proceedings of International Conference on Semantic Computing*; IEEE: Piscataway, NJ, USA, 2009; pp. 577–582. [CrossRef]
54. Atapattu, T.; Falkner, K.; Falkner, N. A Comprehensive Text Analysis of Lecture Slides to Generate Concept Maps. *Comput. Educ.* **2017**, *115*, 96–113. [CrossRef]
55. Ivanova, T. Adaptive Open Corpus E-Learning and Authoring, Using Collaborative Ontology Learning. In Proceedings of the 9th International Conference on Emerging eLearning Technologies and Applications (ICETA), Starý Smokovec, Slovakia, 27–28 October 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 83–87. [CrossRef]
56. Gaeta, M.; Orciuoli, F.; Paolozzi, S.; Salerno, S. Ontology Extraction for Knowledge Reuse: The e-Learning Perspective. *IEEE Trans. Syst. Man Cybern.-Part A Syst. Hum.* **2011**, *41*, 798–809. [CrossRef]
57. Casali, A.; Deco, C.; Romano, A.; Tomé, G. An Assistant for Loading Learning Object Metadata: An Ontology Based Approach. *Interdiscip. J. e-Ski. Lifelong Learn.* **2013**, *9*, 077–087. [CrossRef]
58. Ivanova, T. A Semi-Automatic Ontology Learning Method for E-Learning Resources Terminology Extraction. In Proceedings of the International Conference on Interactive Collaborative Learning (ICL2010), Hasselt, Belgium, 15–17 September 2010; pp. 1030–1034.
59. Louhdi, M.R.C.; Behja, H.; El Alaoui, S.O. A Novel Method for Generating an E-Learning Ontology. *Int. J. Data Min. Knowl. Manag. Process* **2013**, *3*, 151. [CrossRef]
60. Khoiruddin, M.; Kusumawardani, S.S.; Hidayah, I.; Fauziati, S. A Review of Ontology Development in the E-Learning Domain: Methods, Roles, Evaluation. In Proceedings of the 2023 International Conference on Computer, Control, Informatics and Its Applications (IC3INA), Bandung, Indonesia, 4 October 2023; pp. 262–267. [CrossRef]

61. Lee, C.-S.; Wang, M.-H.; Kuan, W.-K.; Ciou, Z.-H.; Tsai, Y.-L.; Chang, W.-S.; Li, L.-C.; Kubota, N.; Huang, T.-X.; Sato-Shimokawara, E.; et al. A Study on AI-FML Robotic Agent for Student Learning Behavior Ontology Construction. In Proceedings of the 2020 International Symposium on Community-Centric Systems, Tokyo, Japan, 23–26 September 2020; pp. 1–6. [CrossRef]
62. Li, G.; Tang, C.; Chen, L.; Deguchi, D.; Yamashita, T.; Shimada, A. LLM-Driven Ontology Learning to Augment Student Performance Analysis in Higher Education. In *Knowledge Science, Engineering and Management*; Springer Nature: Singapore, 2024; Volume 14886, pp. 57–68. [CrossRef]
63. Li, J.; Garijo, D.; Poveda-Villalón, M. Large Language Models for Ontology Engineering: A Systematic Literature Review. 2025. Available online: https://www.semantic-web-journal.net/system/files/swj3864.pdf (accessed on 22 December 2025).
64. Abu-Salih, B.; Alotaibi, S. A systematic literature review of knowledge graph construction and application in education. *Heliyon* **2024**, *10*, e25383. [CrossRef]
65. Chang, M.; D'Aniello, G.; Gaeta, M.; Orciuoli, F.; Sampson, D.; Simonelli, C. Building ontology-driven tutoring models for intelligent tutoring systems using data mining. *IEEE Access* **2020**, *8*, 48151–48162. [CrossRef]
66. Giglou, H.B.; D'Souza, J.; Mihindukulasooriya, N.; Auer, S. Llms4ol 2025 overview: The 2nd large language models for ontology learning challenge. *Open Conf. Proc.* **2025**, *6*, 1–17. [CrossRef]
67. Yang, T.; Ren, B.; Gu, C.; He, T.; Ma, B.; Konomi, S.I. Leveraging LLMs for Automated Extraction and Structuring of Educational Concepts and Relationships. *Mach. Learn. Knowl. Extr.* **2025**, *7*, 103. [CrossRef]
68. Doumanas, D.; Bouchouras, G.; Soularidis, A.; Kotis, K.; Vouros, G. From human-to LLM-centered collaborative ontology engineering. *Appl. Ontol.* **2024**, *19*, 334–367. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.