Hirmer, Tobias; Ochs, Michaela; Henrich, Andreas

Vertical Search Scenarios within a Digital Study Planning Assistant

Date of secondary publication: 04.04.2024

Version of Record (Published Version), Conferenceobject

Persistent identifier: urn:nbn:de:bvb:473-irb-934071

Primary publication

Hirmer, Tobias; Ochs, Michaela; Henrich, Andreas (2023): „Vertical Search Scenarios within a Digital Study Planning Assistant". In: RWTH Aachen (Hrsg.), CEUR Workshop Proceedings, Aachen, Germany, S. 239–246, unter: https://ceur-ws.org/Vol-3630/LWDA2023-paper22.pdf.

Vertical Search Scenarios within a Digital Study Planning Assistant

Tobias Hirmer[1], Michaela Ochs[1] and Andreas Henrich[1]

[1]University of Bamberg, Chair of Media Informatics, An der Weberei 5, 96047 Bamberg, Germany

Abstract

The process of study planning can be complex and challenging for students. Due to decentralized information and complex study structures, a demand for individualized assistance arises. Search functionality plays an essential role in providing such support to the students in various ways. Therefore, we introduce the vertical search scenario of a digital study planning assistant (DSPA), which is currently developed, as a promising area of research for the information retrieval community. As a main contribution, we explore four exemplary search scenarios, including major challenges and potential solutions for each.

Keywords

Digital Study Planning Assistant, Vertical Search, Search Scenarios, Recommendation Systems

1. Introduction

Due to the various ways in which modules can be applied to current study programs, students face a multitude of potential paths they can choose before attaining their degree. In order to support their individual decision process, efficient and effective retrieval of study-planning related resources provided by the university is an essential aspect. However, the findability and quality of these resources are not always ideal [1]. Apart from this, the resources do not consider the individual students' backgrounds and prior knowledge [2]. Sophisticated search functionality, integrated into a digital study planning assistant (DSPA), may address these issues and allow for more effective study planning, tailored to the students' goals such as competency or career goals. We think that this application domain forms an interesting new scenario of a vertical search application that has been little explored so far.

Projects that seek to support individual study planning have developed throughout recent years, yet the aspect of what role search plays within these applications is often unclear. As an early digital study assistant project, *SIDDATA* connects various data sources in order to support individual and goal-based studying [3]. Another project which aims to support individual study planning is *AI Study Buddy* [4]. Further research is focusing on goal-oriented

LWDA'23: Lernen, Wissen, Daten, Analysen. October 09–11, 2023, Marburg, Germany

tobias.hirmer@uni-bamberg.de (T. Hirmer); michaela.ochs@uni-bamberg.de (M. Ochs); andreas.henrich@uni-bamberg.de (A. Henrich)

https://www.uni-bamberg.de/minf/team/hirmer/ (T. Hirmer); https://www.uni-bamberg.de/minf/team/michaela-ochs/ (M. Ochs); https://www.uni-bamberg.de/minf/team/henrich/ (A. Henrich)

0000-0002-5281-0342 (T. Hirmer); 0000-0002-3850-8585 (M. Ochs); 0000-0002-5074-3254 (A. Henrich)

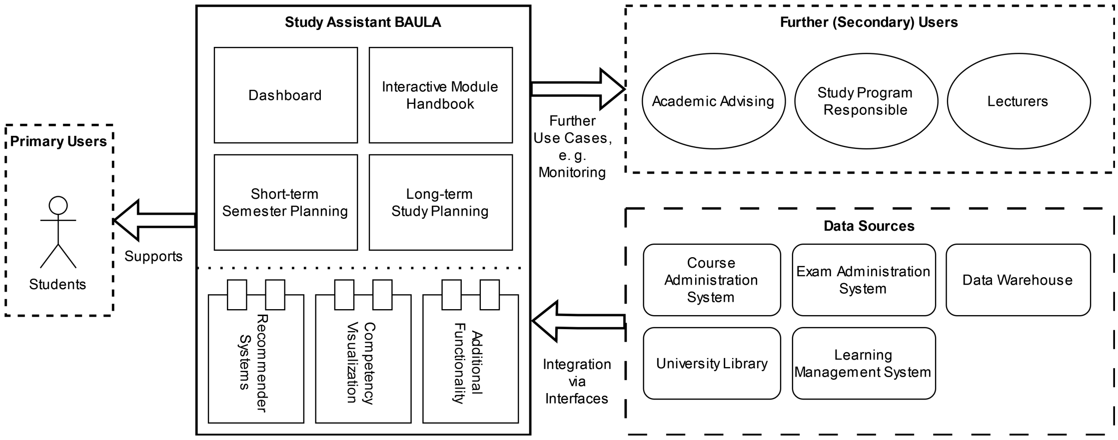

Figure 1: The concept of the DSPA *BAULA* at the University of Bamberg [7]

recommendations [5, 2]. While Jiang et al. [2] used different models of LSTM to predict next suitable courses, Basavaraj and Garibay [5] analyzed curriculum structure in combination with student performance data to generate personalized recommendations.

As a focus of this research, we would like to present relevant vertical search scenarios that became apparent during the work towards a DSPA at the University of Bamberg [6]. This DSPA is named *BAULA* as an acronym for "Bamberger Assistentin zur Unterstützung der Lehrveranstaltungskoordination und -auswahl" (engl. "Bamberg assistant for supporting course coordination and selection"). We think that these search scenarios represent challenging research opportunities for the information retrieval community. In the following, we will first describe the DSPA application and its data sources. Second, we will introduce four major search scenarios within this vertical, explore related obstacles, and suggest possible solutions. Last, we will address overall challenges of this research and conclude.

2. The Digital Study Planning Assistant BAULA

To assist the aforementioned findability of information and to support students in shaping their studies, a DSPA is developed within two projects at the University of Bamberg. The concept of this DSPA is depicted in Fig. 2. The general goal of it is to support students in the whole process of study planning, including tasks such as searching for suitable courses and modules as well as deciding when to attend these. Besides supporting students, the DSPA might also provide further support for *academic advising*, and monitoring features for *study program responsibles* and *lecturers*. To address the main goal of supporting students, the DSPA provides four core functions based on requirements that were elicited previously [6, 7]:

- The *Dashboard* offers students an overview over their current study path as well as relevant statistics like their average grade.
- The *Interactive Module Handbook* extends the current PDF-based handbook with interactivity like a faceted search and details on demand.

- The planning of the current semester is supported by the *Short-term Semester Planning*. Here, a course search as well as a timetable function is provided.
- In contrast to the existing projects mentioned earlier, our approach also includes a function to support the *Long-term Study-Planning*. Here, students are encouraged to plan beyond the current semester by providing hints on relevant prerequisites, for example.

Besides this core system, the DSPA is modular extendable to enable further functionality. Two possible extensions are *Competency Visualization* [8] and *Recommender Systems* [9].

The process of study planning requires data from various sources and university systems. For the Dashboard, personalized and aggregated data about individual and cohort student data from the *Data Warehouse* and the *Exam Administration System* is necessary to generate helpful information. The study structure comes from the Exam Administration System. It contains detailed information about modules and their hierarchical structure within a study program, which are used to visualize the interactive module handbook. Moreover, the study structure is required to enable long-term study planning on module level. For short-term semester planning, the *Course Administration System* is integrated via an interface to retrieve course-specific data such as room and timetable information.

3. Search Scenarios within a DSPA

Within the DSPA described above, four exemplary search scenarios can be identified as interesting research areas. These scenarios will be described in more detail in the following section. They aim to facilitate goal-oriented studying that encourages students to make self-determined and informed decisions. By providing effective search functionalities, the DSPA can assist students in navigating their study path effectively.

A central theoretical concept that is related to this goal is the concept of *competency*. Following the conceptual understanding of [10], we define competency as a combination of a specific, domain-bound *content* and a taxonomy *level*, which represents the extent to which a certain content is acquired. Here, we apply Bloom's taxonomy [11] that was revised by Anderson [12] and defines six levels of complexity that are described by the action verbs *remember, understand, apply, analyze, evaluate* and *create*. In contrast, by content, we define a specific thematic item that is taught within the context of a module. In the following paragraphs, we first introduce search scenarios which are related to competency-related search. Then, we introduce a use case that is based on module content and last, we illustrate a use case that unites both aspects.

3.1. Refining Search with Bloom Taxonomy Levels

Scenario: A central research area that is relevant to the context of the DSPA is not only searching for contents of a module (*content-based retrieval*), but searching for a module that addresses a specific level of competence (*competence-based retrieval*). In the former approach, the classic module retrieval scenario, results may be retrieved based on their content and thus concrete thematic units such as “breadth first search”. In contrast, the latter approach of retrieval relies not only on the content, but a combination of the content and a taxonomy level, for example “breadth first search” + “implement”. By creating the possibility of competence-based search within the DSPA, students are supported by a more precise search for modules, which may be realized as a filter in the form of faceted search or a ranking factor for the results.

Challenges: A challenge in this respect is the varied representation of Bloom's taxonomy information in the module descriptions. While standard sets of (verb) operators might be provided by the university to formulate learning outcomes in the descriptions, module descriptions may contain verbs that have not been provided by these guidelines or the relevant competency details are often contained not only in (nominalized) verbs, but also in adverbs and more complex phrases, partly split within the sentence. For certain contents, competence information might not even be present in the sentence at all. Another challenge is that operators are not exclusively assigned to only one taxonomy level, but can appear within several levels.

Solution approaches: Natural language processing techniques such as language models may capture the semantic meaning of the competencies rather independent of the concrete words that are used in the competencies and thus, word variation and ambiguity in the assignment of verbs to a certain level of taxonomy could be addressed based on similarity. Also, a more detailed analysis of the syntactic and semantic structure via text/dependency parsing and rule-based techniques could prove to be helpful.

3.2. Goal-Based Search

Scenario: Goal-based search may be implemented in various ways. A first approach would be having students define their own study goals and what they want to learn in textual form within the DSPA (e. g. during the onboarding process). A second approach would be having students choose goals from an existing competency standard they can select such as [13]. Based on the input given in these two approaches, modules can be retrieved and recommended.

Challenges: A central challenge in the success of this approach is addressing the vocabulary mismatch, which appears in both scenarios – student vocabulary vs. module vocabulary as well as competency standard vocabulary vs. module vocabulary. Students may use different words or search for more concrete items that may not be contained in the more general module descriptions (e. g. *React* vs. *Web-Frameworks*). Also, their input can have various forms such as lists of topics or more detailed explanations. Competency standards often use more abstract terms than module descriptions. Further difficulties are related to extracting the information from the competence standards automatically. Also, the assignment of verbs to taxonomy levels differs across study programs. A specific taxonomy for CS students is e. g. given by [14].

Solution approaches: In order to address the vocabulary mismatch, the module as well as the competence vocabulary may be extended (e. g. Linked Open Data) to give more context, that aligns both vocabularies. Moreover, transformer-based models like BERT [15] might also be adaptable to match module contents with competence descriptions. For the different forms of student inputs, query expansion may also be helpful.

3.3. Search with Job Descriptions

Scenario: An interesting research scenario within the context of the DSPA involves searching with job descriptions. Considering their long-term study plans, students may input job descriptions into the DSPA, that can use this information to recommend relevant modules to the students. The job descriptions may be provided either as raw textual input or through a convenient drag and drop feature.

Challenges: One of the primary challenges in this scenario is the vocabulary mismatch between job portals and academic documents. This issue has been previously recognized and studied [16]. On one hand, job descriptions, particularly those related to computer science, tend to employ more concrete language and often mention specific tools or technologies. On the other hand, module descriptions in an academic context often take a more abstract or comprehensive approach, providing a broader understanding of the module's content. They further may also not be aligned with the current market situation and requirements.

Solution approaches: In order to overcome the vocabulary mismatch, several potential solutions can be considered. One approach involves using embeddings, which represent textual information as vectors, thereby abstracting from the individual terms. This allows for a more thorough analysis of the semantic meaning behind the text. Additionally, the utilization of terminological reference sources such as ontologies or Linked Open Data (e. g. Wikipedia) can be beneficial in addressing this challenge. These approaches can help establish connections between terms from different domains, enabling the DSPA to bridge the gap between job descriptions and module descriptions more effectively.

3.4. Search for Advanced and Related Modules

Scenario: Another search scenario might be that students aim to search for modules that deepen the understanding of a specific, already successfully completed module or search for related modules that are similar to this module. Therefore, module contents might be represented as paths within a knowledge graph (e. g. WikiData) or as embedding in a vector space (e. g. word2vec/doc2vec). Modules can be (inter)connected on various levels, making it challenging to explore the hierarchical structure and to concretely define the relation between two modules.

Challenges: A main challenge is to define if a relation between two modules exists and how it can be categorized. Modules can contribute to each other, covering similar topics or give a more thorough understanding of a specific aspect. It is important to define this relation and further discuss possible solutions on how these relations can be identified. The relation between modules might also be affected by the study program of the requesting student.

Solution approaches: Similarity of modules can be identified using word embeddings on learning materials [17]. To identify advanced modules, that deepen the knowledge provided by a specific module, integrating the search for competence level as described before can contribute to a solution. Since these competence levels are organized sequentially as well, an advanced module can be identified searching for the topics of a specific module with a higher competence level. To further categorize different relations between modules, a graph-based approach might be applicable. As mentioned, the contents of a module can be represented as a knowledge graph. One assumption might be, that advanced modules expand the leaves (topics) of the queried module, whereas related modules contribute to the parent nodes of the leaves.

4. Data and Challenges

Beyond the already described scenario-specific challenges, this section will shortly address further challenges covering data issues and self-organizing skills. As previously mentioned, the data provided by the university is insufficient to meet the current need for personalization [2]. Additionally, quality is reduced due to a variety in the extent and wording of module and course information, which is highly dependent on the specific lecturer. Besides poor data quality, quantity and accessibility of data is challenging. To the best of our knowledge, large-scale corpora of (German) module catalogs currently do not exist in the form of one large data set, but are spread across university websites and repositories.

Besides challenges regarding the data, the DSPA and the respective search scenarios need to assist without directly guiding students too much. Assistance should not lead to help students the easiest way through study or making study planning obsolete limiting the self-organization skills of students. Also, the influence of more information on modules or courses (e. g. grade distribution) might lead to negative effects like a reduced grade [18]. Here the described search scenarios above should help students in making a more informed and conscious decision, without taking too much responsibility from them.

5. Conclusion

While the four search scenarios described above each search for modules on the basis of different information needs or different formulations of the information need, it is of course also conceivable to search for other things in the scenario of a DSPA. An example would be the search for Open Educational Resources (OER). This could be used, for example, to make alternative suggestions as to how any missing prior knowledge could be acquired. For this purpose, a pool would first have to be crawled and indexed on the basis of directories for OER. Then, the question would be how to identify appropriate OER based on the required prior knowledge listed in module descriptions. This further example of a search scenario in the DSPA context illustrates the need for appropriate search services as well as the challenges for information retrieval.

Acknowledgments

This research was conducted as part of the projects “Developing Digital Cultures for Teaching” (DiKuLe) and “Learning from Learners” (VoLL-KI) and was funded by “Stiftung Innovation in der Hochschullehre” (Foundation for Innovation in Higher Education) as well as “Künstliche Intelligenz in der Hochschulbildung” (Artificial Intelligence in Higher Education).

References

[1] M. Ochs, T. Hirmer, A. Henrich, Studierende und die Studienplanung: Untersuchung von Herausforderungen und Entwicklungsperspektiven eines digitalen Studienplanungsassistenten, 2023. URL: https://doi.org/10.5281/zenodo.8037697.

[2] W. Jiang, Z. A. Pardos, Q. Wei, Goal-based Course Recommendation, in: Proceedings of the 9th International Conference on Learning Analytics & Knowledge, LAK19, Association for Computing Machinery, New York, NY, USA, 2019, pp. 36–45. doi:10.1145/3303772.3303814.

[3] C. M. König, N. Guhr, M. H. Breitner, Ein Anreizsystem zur erfolgreichen Implementierung eines digitalen Studienassistenten, in: N. Gronau, M. Heine, K. Poustcchi, H. Krasnova (Eds.), WI2020 Community Tracks, GITO Verlag, 2020, pp. 33–48. doi:10.30844/wi_2020_s3-koenig.

[4] M. Wagner, H. Helal, R. Roepke, S. Judel, J. Doveren, S. Goerzen, P. Soudmand, G. Lakemeyer, U. Schroeder, W. M. P. van der Aalst, A Combined Approach of Process Mining and Rule-Based AI for Study Planning and Monitoring in Higher Education, in: M. Montali, A. Senderovich, M. Weidlich (Eds.), Process Mining Workshops, Lecture Notes in Business Information Processing, Springer Nature Switzerland, Cham, 2023, pp. 513–525. doi:10.1007/978-3-031-27815-0_37.

[5] P. Basavaraj, I. Garibay, A Personalized "Course Navigator" Based on Students' Goal Orientation, in: Proceedings of the 2018 ACM International Conference on Supporting Group Work, GROUP '18, Association for Computing Machinery, New York, NY, USA, 2018, pp. 98–101. doi:10.1145/3148330.3154508.

[6] T. Hirmer, J. Etschmann, A. Henrich, Requirements and Prototypical Implementation of a Study Planning Assistant in CS Programs, in: 2022 International Symposium on Educational Technology (ISET), 2022, pp. 281–285. doi:10.1109/ISET55194.2022.00066.

[7] M. Ochs, T. Hirmer, A. Henrich, Concept and Possible Impacts of a Study Planning Assistant in Higher Education, in: 2023 International Symposium on Educational Technology (ISET), Hong Kong, 2023. (in press).

[8] T. Hirmer, N. Heyne, A. Henrich, Die kompetenzorientierte Studienplanung - Entwicklung eines Tools zur Unterstützung von (Lehramts-)Studierenden, in: DELFI 2021, Gesellschaft für Informatik e.V., Bonn, 2021, pp. 121–126.

[9] M. Ochs, T. Hirmer, K. Past, A. Henrich, Design-Focused Development of a Course Recommender System for Digital Study Planning, in: A. Abelló, P. Vassiliadis, O. Romero, R. Wrembel, F. Bugiotti, J. Gamper, G. Vargas Solar, E. Zumpano (Eds.), New Trends in Database and Information Systems, Communications in Computer and Information Science, Springer Nature Switzerland, Cham, 2023, pp. 575–582. doi:10.1007/978-3-031-42941-5_50.

[10] E. Klieme, D. Leutner, Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG, Zeitschrift für Pädagogik 52 (2006) 876–903.

[11] B. S. Bloom, Taxonomy of Educational Objectives, McKay, New York, 1968.

[12] L. W. Anderson, D. R. Krathwohl, A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives, Longman, 2007.

[13] O. Zukunft, Empfehlungen für Bachelor- und Masterprogramme im Studienfach Informatik

an Hochschulen, 2016. URL: https://dl.gi.de/handle/20.500.12116/2351.
[14] U. Fuller, C. G. Johnson, T. Ahoniemi, D. Cukierman, I. Hernán-Losada, J. Jackova, E. Lahtinen, T. L. Lewis, D. M. Thompson, C. Riedesel, E. Thompson, Developing a computer science-specific learning taxonomy, in: Working Group Reports on ITiCSE on Innovation and Technology in Computer Science Education, ITiCSE-WGR '07, Association for Computing Machinery, New York, NY, USA, 2007, p. 152–170. URL: https://doi.org/10.1145/1345443.1345438. doi:10.1145/1345443.1345438.
[15] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805 (2018). arXiv:1810.04805.
[16] C. Gallardo Pérez, A. González Eras, P. Quezada S., P. Ludeña González, Comparing competences on academia and occupational contexts based on similarity measures, in: WEBIST 2015 : 11th International Conference on Web Information Systems and Technologies : Proceedings, Scitepress, Setúbal, Portugal, 2015, pp. 540–546. doi:10.5220/0005491405400546.
[17] N. Seidel, C. M. Rieger, T. Walle, Semantic Textual Similarity of Course Materials at a Distance-Learning University, in: CSEDM@EDM, 2020.
[18] C. Holman, S. J. Aguilar, A. Levick, J. Stern, B. Plummer, B. Fishman, Planning for success: How Students Use a Grade Prediction Tool to Win Their Classes, in: J. Baron, G. Lynch, N. Maziarz, P. Blikstein, A. Merceron, G. Siemens (Eds.), Proceedings of the Fifth International Conference on Learning Analytics And Knowledge, ACM, New York, NY, USA, 2015, pp. 260–264. doi:10.1145/2723576.2723632.