The other side of the same coin: From learning-centric search systems to search-centric learning systems

Catherine L. Smith[a] and Soo Young Rieh[b]

[a] *Kent State University, Kent, Ohio, USA*
[b] *University of Texas at Austin, Austin, Texas, USA*

Abstract

This short paper proposes a framework for designing search-centric learning systems that support search as learning. Our argument draws on Jackson's purpose-centric design concepts for software, and from research on self-regulated learning, an established paradigm that intersects psychology, education, and learning sciences. In introducing these ideas we also examine searching for information as self-regulating activity and the design of experimental learning systems that support self-regulation. We argue that embedding search functionality within learning systems holds promise for better supporting students engaged in self-regulated learning.

Keywords 1

searching as learning; self-regulated learning; software design; metacognition

1. Introduction

Smith and Rieh [18] presented design goals focused on *information literate action* and the need for *learning-centric search systems* designed for supporting metacognitive engagement. One of the key ideas of learning-centric search systems was to better facilitate active engagement with information that would result in long-term learning and creative endeavor. In this paper, we flip that design goal over and focus on self-regulated learning to argue for the design of a *search-centric learning system*. Such a system would embed search functions within a learning system.

Our argument draws on two constructs. First is Jackson's conceptual design paradigm for software [11], which focuses on alignment between a users' *purpose* and functional *concepts* within a software application. More specifically, our goal is to focus on the users' purposes for information search during *self-regulated learning* (SRL) [16]. SRL is a psychological construct focused on cognitive, metacognitive and emotional processes students use when engaged in learning, and on how those processes affect learning in an academic setting [8, 21, 25]. This is a decidedly design-centric research orientation and we acknowledge that goals such as basic research are also essential.

This paper is organized as follows. The first three sections present ideas and selected work from purpose-centric design, self-regulated learning, and learning system design. Next we briefly examine results showing that searching for information is a process integral to SRL. We then present an example of a search-centric learning system, define the construct more broadly, and discuss a short scenario explicating the need for search concepts that better meet purposes for searching during SRL. The paper concludes with a brief summary. The paper contributes a framework for considering design goals for learning systems that support search as learning.

2. Purpose-centric design.

Jackson [11] proposed that good software design aligns user-centric purposes with software concepts. Within this paradigm, "a

Proceedings of the CIKM 2020 Workshops, October 19-20, Galway, Ireland
EMAIL: csmit141@kent.edu (A); rieh@ischool.utexas.edu (B)
ORCID: 0000-0002-8433-0829 (A); 0000-0002-2978-692X (B)

concept is a self-contained, reusable, increment of functionality that is motivated by a purpose defined in terms of the needs of an end user" [17]. Within a design, concepts exist at all levels of granularity and are independent of their instantiations in code. For example, Twitter's purpose is viral public expression. Twitter serves its purpose with three concepts: *tweet*, *hashtag* and *following*, each concept with a single purpose. The purpose of a tweet is short public posting (a variant of the concept *posting*). The purpose of a hashtag is to establish associations between tweets (a variant of the concept *label*). The purpose of following is to receive messages from a specific account. All three concepts may be used for similar purposes in other applications or they may be instantiated in concept variants with similar functionality (as tweet is for posting and hashtag is for label). Further, each of these concepts comprises sub-concepts, ideally, each with its own single purpose.

The purpose of a search application is to find information. Search applications use two concepts: *query* and *results*. The purpose of a query is to express an information need. The purpose of results is to expose the information sources most likely to meet the need. These purposes apply in many contexts thus these concepts have many applications. Examples of sub-concepts for query include *suggestion*, *completion*, and *structure*. The purpose of suggestion is to clarify the need by helping users reformulate queries. The purpose of completion is to minimize typing and typing errors. The purpose of structure is to improve the precision of results. The concept of structure includes sub-concepts such as *filter* and *logic*.

Good software uses concepts that each serve a single purpose, where the purpose is defined well enough to motivate one and only one concept. *Unmotivated* concepts serve no purpose and are of no intrinsic value to users; typically these involve patching over a design flaw or simply superfluous functionality. When software contains *redundant* concepts that fulfill the same purpose the application is confusing, hard to learn, and inefficient for users. Problems also arise when a concept serves more than one purpose; *overloaded* concepts are likely to require design tradeoffs that render the concept suboptimal for at least one purpose. Of course, unfulfilled purposes with no concept are often opportunities for new applications and enhanced designs. Jackson's full perspective on design includes a rich set of ideas that we do not cover here, however, the purpose/concept heuristics serves as a useful framework for considering design goals for a search-centric learning system. We return to purpose-centric concepts later in the paper.

3. Self-regulated learning

Hypothesizing a search-centric learning system provides an opportunity to focus on the purpose for information search within the context of a system designed for learning. The construct of SRL is particularly compelling as a framework because it is domain-independent and centers on the general behaviors and mental processes students use when engaged in effective learning. Also, its theories are embedded in much recent work on learning system design and related analytics [21].

SRL has been defined as "self-directive processes and self-beliefs that enable learners to transform their mental abilities, such as verbal aptitude, into an academic performance skill, such as writing." [25]. At its most basic, SRL posits the recursive use of cognitive and metacognitive skills in three phases during task-focused learning: *forethought*, *performance*, and *assessment*. Each of these may be variously named or decomposed, but there is consensus on a minimal three [16]. Experimental research often focuses on subsets of specific skills within each phase. Forethought generally encompasses interpreting, understanding, strategizing, and planning a learning task. Performance focuses on monitoring and control of plans and strategies while learning. Assessment includes using performance feedback, reacting, adapting, and reflecting on cognition. Theories differ on the roles, types, and importance of motivation, skill, context, individual differences, and prior knowledge that affect transitions between phases. Increased use of SRL skill reliably enhances learning outcomes [26], thus much work has been done on the design of instructional methods that enhance self-regulation [16].

SRL has a large, rich, and growing literature of empirical study and convergent theory covering task, affect, and motivational factors in individual, shared, and collaborative learning scenarios [16]. It is studied sufficiently to have spawned multiple handbooks, literature reviews, and meta-analyses [1, 10, 16]. Protocols and self-report instruments exist for measures of learning and self-regulation [1]. Current research uses behavioral logs collected in online learning environments [21]. SRL contrasts with the concept of self-directed learning (*SDL*), which unlike SRL, focuses on individual initiative and adult learners' formulation of their own learning objectives [13]. We acknowledge that SDL and other learning theories may be equally valid and useful for consideration in search-as-learning. It is not our objective to claim SRL is the only useful paradigm.

As implied above, SRL is a large and complex research domain that bridges several areas of psychology and practical aspects of education. Generally, results from experimental studies have informed models of factors affecting the use of cognitive and metacognitive skill as related to learning outcomes. As learning has moved to computers and then online, these methods and attendant research have moved to online learning systems.

4. Learning system design

Learning systems (*computer-based learning* systems; *CBLE*s) are designed for many purposes. Within the SRL community, designs derive from pre-computer classroom and tutoring approaches that enhance SRL and ultimately, learning outcomes. Studies on experimental SRL systems focus typically on methods for facilitating SRL, usage of SRL skill, and differential learning outcomes. Early experimental systems focused on SRL within the context of learning tasks such as homework assignments on a topic. The first published systems were domain-independent, general-purpose, and operated over the Encarta encyclopedia [2, 23]. Research with the early MetaTutor system focused on scaffolding learning goals for domain knowledge and students' use of SRL skills [2, 3]. A later version of MetaTutor used animated pedagogical agents to scaffold skills in SRL, with prompts and feedback delivered as student learning progressed [9]. Other examples include a dashboard that prompts forethought and provides feedback on learning behavior [14], a system that uses curricula structured in pedagogical concept maps to guide a course of study and facilitate SRL using prompts [12], and a system of prompts selected by learners [19]. These examples use some form of navigable content structure for the target learning domain.

Also an early design, the general purpose, domain-independent gStudy system was different [23]. The design sought to facilitate SRL through behaviors such as note-taking, labeling, glossary building, concept mapping, coaching, chatting, and collaborating. The system also included a learner's display of analytics derived from logged interaction behavior. Much of the functionality involved information search and interaction such as "indexing, annotating, analyzing, classifying, organizing, evaluating, cross referencing and searching " ([23] page 107). Later versions of the system (nStudy) incorporated a Web browser, webpage linking, tagging, hypertext authoring, and a library of information resources filterable on various bibliographic and user-generated metadata [22, 24].

Experimental systems from the SRL community have not used explicit models of the individual learner (but see [15] for a notable exception), however a large, parallel body of research in learner modeling has done so. Early learner models tracked and facilitated content navigation and summative assessment within a closed system, with data generated during observable behavior [5]. Modern systems use various forms of statistical modeling, where the product of the model is generally a visual display. *Open learner models* (*OLM*s) make their underlying data accessible to the learner, who may initiate, append, or update the data directly. OLMs may model states associated with SRL, including data and reports on reflection, planning, monitoring, and formative evaluation [5].

[10] reviewed 64 published OLMs designed for higher education. The vast majority (89%) of models supported learning in STEM specifically. Most of the OLMs (63%) operated within a closed learning system such as an automated tutor. The most common modeling objectives focused on predicting and tracking learners' attainment of domain knowledge. Within the three-phase view of SRL (forethought, performance, assessment), fewer than one-third of OLMs reviewed addressed any part of a learner's forethought, with support of performance and summative assessment more common.

The above brief review suggests that currently published learning systems often address domains where knowledge content can be structured to scaffold and support the attainment of domain knowledge. Importantly, learning also occurs in less structured domains where problems, goals, and standards for success are relatively underspecified. For example, success in information-intensive learning tasks such as writing a research paper require considerable SRL. This less structured learning scenario provides context for considering the purposes for search during learning.

5. Search in SRL learning systems

The development of experimental computer systems for SRL enabled researchers to trace students' use of strategy and skills during study. With those advances, the capture of data indicative of the internal SRL processes has been a key need, thus think-aloud methods are common. One early study used think-aloud during assignment completion in a hypertext encyclopedia [4]. The environment included a search function, which students were free to use. Utterances indicative of SRL were coded within the authors' four-part model of SRL. Monitoring (awareness of self, task, and context) included *identifying the adequacy of information* and *information content evaluation*. Strategy use (control and regulation of self, task, and context) included *coordinating information sources*; *selecting a new information source*; *goal-directed information search*; *free search* (searching with no articulated goal); and *evaluating content relative to a learning sub-goal*. Later work on how students sequenced SRL activities also used think-aloud in a closed hypertext environment [19]. Although the system did not offer query-based search, searching for information and judging information relevance were found among key metacognitive activities. The authors examined patterns of SRL processes, finding prominent effects of the SRL system on the position of search within the patterns of SRL activity.

The above results suggest that information search, interaction, and judgment are frequent and central aspects of SRL, even in relatively simple environments like a closed hypertext system. Where the options for searching are more complex, covering not only the Web but also tags, bookmarks, folders, saved work, online textual material, media, library resources, a learning-management system, and so forth, we expect the role of searching to also be more complex. As a central psychological process for learning, a learner's purpose for information search may involve accessing domain knowledge and self-regulation of learning. In the next section we consider how these purposes may fit concepts for a search-centric learning system.

6. Purpose and search-centric learning systems

For a search-centric learning system, good design provides search concepts that fit the user's purposes within the context of the learning application. The design of an intelligent textbook provides a clear example.

The purpose of the Inquire Biology textbook [6] is to assist students in learning complex concepts and their associations within the biology domain. One of the central concepts used in the textbook's design is question-and-answer search, with attendant concepts and sub-concepts such as question generation, vocabulary lookup, and term association search. These concepts fit the types of SRL strategies that work well in highly structured domains such as those found in STEM: memorization, knowledge elaboration, self-test, and self-questioning. Within the design, search is not an overloaded monolithic concept. Rather, each purpose for searching is met with a concept fit for purpose. One may consider the textbook a search-centric learning system, albeit one that does not search beyond its internal resources.

In the limited view presented in this paper, learning systems may have two distinct purposes: (1) to facilitate the learner's acquisition of special knowledge in a single domain (e.g., the Inquire Biology textbook) or (2) to facilitate the development of transferable knowledge and skill in any domain; for example: critical thinking, reading for comprehension, synthesis, and expository writing. How well a single system can fulfil both purposes is a matter for empirical study, but information search is essential in both cases. As the Inquire Biology textbook demonstrates Jackson's notion of fitting functional concepts to search purposes, we argue that new system concepts can be designed to fulfill the purposes for searching in the second case. Indeed, we have argued that the need for this view is compelling [18] due to psychological effects on metacognition associated with current system designs. Like others studying undergraduate learners [20] our recent observations of 100+ college students working on transfer-focused assignments revealed heavy reliance on Web search. Those observations led us to consider the ways in which search functionality may fit Jackson's definition of an overloaded concept. We believe there is need for design concepts that better facilitate information search purposes in the context of SRL.

For example, we consider Chris, a freshman nursing student taking two courses requiring a research paper. For a first-year writing course the paper can be on any topic. The paper needs to demonstrate research and writing skill; pedagogically this is learning meant to transfer to any general learning situation. For Chris's nursing class, the paper must go beyond the course content to demonstrate understanding of a chronic disease condition. Here the goal is to show deep knowledge and synthesis, so Chris wants to choose a condition that has already been introduced in class.

Considering Chris's goals through the lens of search system design, observation of Chris's current and past search behavior enables inference on the structure of the two tasks and topics. Within this task-centric view, we may infer Chris's more specific information goals and internal state as interaction proceeds over possibly multiple sessions. Having inferred tasks, topics, goals, and internal state, inferences may be updated with the goal of exposing only the information sources most likely to optimize learning and task completion. From Chris's perspective, the search concepts used for this purpose are queries and results in a Web search engine.

Flipping the design goal over embeds search functionality within a learning system. The purpose of SRL-focused systems is to facilitate the development and use of effective strategies for study and academic achievement. Here one or more OLMs may be in use, providing information that obviates the need for inference on individual differences and preferences for Chris's learning processes and skills. Further, specific sources of contextual information may be accessible in textbooks, readings, prior work on assignments, and other attendant sources. Chris's progress relative to instructional scaffolding may also be available. Before working on a paper, Chris is likely to engage in explicit forethought captured for later self-reflection. When Chris works on one of the papers, features of the assignment are accessible to the search system, along with concurrent evidence of engagement with search functions and tools and supports for SRL. This context provides rich data for the search system and for research examining the purposes for searching external information sources during SRL. For example, one such purpose is the notion of sourcing, a metacognitive skill used in reading for comprehension where the reader attends to "who says what" [7]. We argue that search functionality can be designed using concepts that fulfil varied and complex purposes for searching during SRL.

7. Conclusion

This paper makes three contributions to search as learning. First, we reviewed self-regulated learning as a useful paradigm for research on search as learning, focusing on how search activities may be conceptualized as self-regulated learning. Second, we introduced Jackson's [11] software design paradigm, focusing on alignment of the purposes for searching with functional concepts that fulfill those purposes. Third, we presented a design-centric framework for considering the purpose of searching in academic tasks, proposing that a search-centric learning system may fulfil those purposes with the design of new functional concepts. We look forward to discussing these ideas with IWILDS workshop attendees. [32, 18, 33, 34].

8. REFERENCES

[1] Araka, E. et al. 2020. Research trends in measurement and intervention tools for self-regulated learning for e-learning environments—systematic review (2008–2018). Research and Practice in Technology Enhanced Learning. 15, 1 (2020), 1–21.

[2] Azevedo, R. et al. 2004. Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia?

Contemporary Educational Psychology. 29, 3 (Jul. 2004), 344–370. DOI:https://doi.org/10.1016/j.cedpsych.2003.09.002.
[3] Azevedo, R. et al. 2004. The role of self-regulated learning in fostering students' conceptual understanding of complex systems with hypermedia. J. Educational Computing Research. 30, 1–2 (2004), 87–111.
[4] Azevedo, R. and Cromley, J.G. 2004. Does training on self-regulated learning facilitate students' learning with hypermedia? J Educational Psychology. 96, 3 (2004), 523.
[5] Bull, S. and Kay, J. 2007. Student models that invite the learner in: The SMILI:() Open learner modelling framework. International J Artificial Intelligence in Education. 17, 2 (2007), 89–120.
[6] Chaudhri, V.K. et al. 2013. Inquire Biology: A textbook that answers questions. AI Magazine. 34, 3 (2013), 55–72.
[7] Delgado, P. et al. 2020. To click or not to click: investigating conflict detection and sourcing in a multiple document hypertext environment. Reading and Writing. (2020), 1–24.
[8] Efklides, A. 2011. Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. Educational Psychologist. 46, 1 (2011), 6–25.
[9] Harley, J.M. et al. 2017. Let's Set Up Some Subgoals: Understanding Human-Pedagogical Agent Collaborations and Their Implications for Learning and Prompt and Feedback Compliance. IEEE Transactions on Learning Technologies. 11, 1 (2017), 54–66.
[10] Hooshyar, D. et al. 2020. Open learner models in supporting self-regulated learning in higher education: A systematic literature review. Computers & Education. (2020), 103878.
[11] Jackson, D. 2015. Towards a theory of conceptual design for software. 2015 ACM International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software (Onward!) (2015), 282–296.
[12] Kivimäki, V. et al. 2019. Curricular Concept Maps as Structured Learning Diaries: Collecting Data on Self-Regulated Learning and Conceptual Thinking for Learning Analytics Applications. J Learning Analytics. 6, 3 (2019), 106–121.
[13] Loyens, S.M. et al. 2008. Self-directed learning in problem-based learning and its relationships with self-regulated learning. Educational Psychology Review. 20, 4 (2008), 411–427.
[14] Molenaar, I. et al. 2020. Personalized visualizations to promote young learners' SRL: the learning path app. Proceedings of the Tenth International Conference on Learning Analytics & Knowledge (2020), 330–339.
[15] Nussbaumer, A. et al. 2015. A competence-based service for supporting self-regulated learning in virtual environments. J Learning Analytics. 2, 1 (2015), 101–133.
[16] Panadero, E. 2017. A review of self-regulated learning: Six models and four directions for research. Frontiers in Psychology. 8, (2017), 422.
[17] Perez De Rosso, S. et al. 2019. Declarative assembly of web applications from predefined concepts. Proceedings of the 2019 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software (2019), 79–93.
[18] Smith, C.L. and Rieh, S.Y. 2019. Knowledge-Context in Search Systems: Toward Information-Literate Actions. Proceedings of the 2019 Conference on Human Information Interaction and Retrieval (New York, NY, USA, 2019), 55–62.
[19] Sonnenberg, C. and Bannert, M. 2015. Discovering the effects of metacognitive prompts on the sequential structure of SRL-processes using process mining techniques. J of Learning Analytics. 2, 1 (2015), 72–100.
[20] Warwick, C. et al. 2009. Cognitive economy and satisficing in information seeking: A longitudinal study of undergraduate information behavior. JASIST. 60, 12 (2009), 2402–2415. DOI:https://doi.org/10.1002/asi.21179.
[21] Winne, P.H. 2018. Cognition and Metacognition Within Self-regulated Learning. (2018). In D. H. Schunk & J. A. Greene (Eds.), Educational psychology

handbook series. Handbook of Self-Regulation of Learning and performance (p. 36–48). Routledge/Taylor & Francis Group.

[22] Winne, P.H. et al. 2019. nStudy: Software for learning analytics about processes for self-regulated learning. J Learning Analytics. 6, 2 (2019), 95–106.

[23] Winne, P.H. et al. 2006. Supporting self-regulated learning with gStudy software: The Learning Kit Project. Technology Instruction Cognition and Learning. 3, 1/2 (2006), 105.

[24] Winne, P.H. and Hadwin, A.F. 2013. nStudy: Tracing and supporting self-regulated learning in the Internet. International Handbook of Metacognition and Learning Technologies. Springer. 293–308.

[25] Zimmerman, B.J. 2008. Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. American Educational Research Journal. 45, 1 (2008), 166–183.

[26] Zimmerman, B.J. and Schunk, D.H. 2011. Self-regulated learning and performance: An introduction and overview. Handbook of Self-regulation of Learning and Performance. 1–12.