Investigating the learning perspective of Searching as Learning, a review of the state of the art

Paulo J. A. Gimenez, Marcelo Machado, Cleber Pinelli, Sean W. M. Siqueira

1Programa de Pós-Graduação em Informática (PPGI)
Universidade Federal do Estado do Rio de Janeiro (UNIRIO)
Av. Pasteur, 456, Urca – Rio de Janeiro – RJ – Brasil
{paulo.gimenez, marcelo.machado}@edu.unirio.br
{cleber.pinelli, sean}@uniriotec.br

Abstract. Current search engines are not designed to facilitate learning as they do not lead the user to develop more complex skills. Searching as Learning (SAL) emerged as a research area from the intersection of information search and learning technologies in order to advance the study of searching as a learning process. However, we wonder how have the learning theories and approaches been explored in SAL. Through a systematic review of the literature, we identified 65 papers that report SAL solutions. We analyzed them, seeking to answer (i) which learning theories, approaches and methods support the search as a learning process, and (ii) what metrics, procedures, or treatments were used to measure learning during the searching process. We uncover the learning perspective in the SAL literature, discussing the learning paradigms, the mechanisms influencing the learning process, the search session design for learning, and the knowledge gain measurement strategies.

1. Introduction

Performing search tasks to acquire new skills and knowledge is a commonplace. In distance education (Cho and Shen, 2013), the use of search systems become indispensable due to the shift from a physical space of social interaction to a digital space. Coronavirus emergence has raised the need for a transformation in education, which is intensively supported by educational technology (Williamson et al., 2020). Searching as Learning (SAL) is a recent research area that considers issues involving learning during the search process (Moraes et al., 2018). It focuses on the impact, influence, and consequences of using search engines as learning technologies (Machado et al., 2019).

There are two main tasks performed by learners while using search systems, the decision-making process of formulating and reformulating queries (Liu et al., 2010); and the navigation where the learner interacts with information (Herder and Juvina, 2004). In the first task, the learner has to properly use the key terms to retrieve relevant information and develop successful search strategies. In the second task, the learner is responsible for selecting the sources. He should be able to identify reliable sources and to establish relationships between concepts to process information towards knowledge construction. The dynamics of these two tasks contribute to learning as the learner is engaged in his process. Knowledge gain in this context comes from web search behavior and information use analysis (Wilson, 1999).
Information retrieval and interaction communities have supported the understanding of the learning process during search tasks as well as the development of search engines that target learning outcomes (Rieh et al., 2016). Vakkari (2016) surveyed empirical studies on the relationship between information searching and learning. On the same perspective, Rieh et al. (2016) critically reviewed the literature, highlighting the learning process during the search process, focusing on critical and creative thinking. Then, Hoppe et al. (2018) reinforced the current challenges of a SAL agenda, over information retrieval and psychopedagogy perspectives. Marchionini (2018) reviewed historical aspects of the learning literature, giving attention to learning theories and highlighting changes that occur on neural, cognitive, affective, and behavioral dimensions. Machado et al. (2019) revealed the presence of few studies rooted in learning theories in SAL, which may represent a research gap. SAL is an approach that suits better for informal learning perspective. However, in order to incorporate this approach to distance education, the supervision of learner’s progress is necessary, since learning assessment is an important step in formal education. In this context, detecting, measuring, and predicting knowledge gain (Gadiraju, 2018; Gadiraju et al., 2018) in SAL situations might become the focus due to the current pandemic crisis. For that reason, the understanding of learning theories and the mechanisms to estimate learning progress are important research issues for SAL.

Through a systematic review protocol, we identified 65 research papers that presented studies with solutions that support search as a learning process. Then, we listed some findings from the analysis of these studies considering: (i) learning theories are used to understand and define the learning process, with emphasis on the use of the main known approaches such as behaviorism, cognitivism and constructivism; and (ii) the techniques used to measure knowledge gain in searching as a learning process, including learner’s metacognition, search success performance, and predictive model analysis.

The remaining of this paper is organized as follows: Section 2 presents the systematic review methodology; Section 3 presents and analyzes the learning approaches and techniques by the literature review in SAL; Section 4 deals with issues related to the threat of validity of the present work. Finally, Section 5 concludes the paper and develops some perspectives of future work.

2. Methodology
The purpose of this systematic literature review is to identify, evaluate and report the available studies considering the research questions (Kitchenham and Charters, 2007). This review was organized based on the main activities proposed by Kitchenham and Charters (2007): planning, conducting and reporting the study.

The systematic review aims to answer the following research questions:

- **RQ1**: Which learning theories, approaches and methods support the search as a learning process?
- **RQ2**: What metrics, procedures and treatments are used to measure learning during a searching process?

The Selection Criteria process used three exclusion criteria (EC):

- **EC1**: The paper does not deal with the SAL context or learning associated with the search process OR
EC2: The paper was not published in a peer-review conference, event or journal OR
EC3: The paper was not written in the English language.

Then the search query was formulated and executed in eight digital libraries, as presented in Table 1.

Initially, to create the search string, we considered the research questions, and the scope was defined using the PICOC method (Petticrew and Roberts, 2008): Population="Web Search", Intervention = “Search as Learning”, Comparison="", Outcome="Solutions" and Context="Education"). Therefore, we intended to evaluate Searching as Learning proposals in Web-based settings that were proposed with an educational purpose. However, as SAL is a rather new research agenda, we only considered Intervention terms and their alternate spelling and synonyms, not applying other filters that could leave some sort of solutions out of the analysis. Then, an initial string was formed considering synonyms and alternate spellings from the intervention term, concatenated using Boolean OR operator. A set of potential primary studies were also defined to validate the search string accuracy in the selected databases and check whether the search retrieved the verification studies. The final search string was defined as follows:

\begin{equation}
("Search as Learning" OR "Searching as Learning" OR "Search as a Learning" OR "Searching as a Learning" OR "Learning on Search" OR "Learning on Searching")
\end{equation}

The platform Parsif.al\(^1\) was used to catalog the papers and manage the selection activity (Kitchenham and Charters, 2007). The systematized process covered objectives, PICOC items, research questions, search string, keywords and synonyms, selecting the sources, inclusion and exclusion criteria collected during Nov-Dec/2019. The process comprised the following steps:

\begin{itemize}
  \item **Step 1** The execution of the search string, considering the selected sources.
  \item **Step 2** The merge of the results from all databases and removal of duplicates.
  \item **Step 3** The evaluation of the papers based on their titles and abstracts, considering the selection criteria.
  \item **Step 4** The evaluation of the papers based on their introduction and conclusion, considering the selection criteria.
  \item **Step 5** The full reading of the papers and their analysis, considering the research and mapping questions.
\end{itemize}

First, each paper was evaluated by two researchers. In the case of divergence, a third researcher was assigned to the evaluation. Only papers accepted by two researchers were considered.

With the final search string applied in all sources, a set of 577 papers was obtained in Step 1. In Step 2, the duplicated papers were removed, and a set of 461 (79.89\%) papers remained. After reviewing the titles and abstracts, in Step 3, 104 (18.02\%) papers were included. After evaluating the introduction and conclusion of each paper in step 4, only 65 (11.26\%) remained and were accepted to the next stage (Step 5).

\(^1\)https://parsif.al/
3. The learning side of Searching as Learning

Since there is not a unique definition for learning or a unique way to assess it, SAL may be seen according to the set of theories and pedagogical principles guiding the learning process. There are different approaches applied to comprehend the human mind as paradigms. Learning theories were created under these paradigms. Even having antagonistic comprehension of learning, one theory does not invalidate another, and they complement and refine educational practices (Schunk, 1991). Learning approaches result from the attempts of psychologists to organize observations, hypotheses, cues, laws, principles and suppositions that have been presented from human behavior (Lefrancois, 2012).

Picciano (2017) examined theoretical frameworks and models focused on online education. In that work, Behaviorism is described with a focus on how people behave, while Cognitivism is based on the concept where the mental process has an important role in learning, establishing the connection between environmental stimuli and the individual’s responses. Constructivism, represented mainly by Vygotsky and Piaget, has a focus on the individual’s development, where knowledge is constructed as a result of individual experience or social interaction (Lefrancois, 2012).

In order to answer RQ1, some analyses were done to identify and correlate the dimensions of macro learning approaches, the mechanisms that influence learning, and session design to structuring the environment and the interaction that supports learning. These elements of analysis were structured and collected based on the learning approaches perspective (Picciano, 2017). The panel of learning factors influencing searching as learning was also considered (Hansen and Rieh, 2016).

1. **Learning Paradigms (LP)**: Represents sets of principles, statements or ideas that define, explain or predict learning. It supports the understanding of how people learn. Some studies explicitly declare the learning paradigm they are using, such as Liu and Song (2018), who constructed two types of learning-related tasks according to the first two levels of Cognitive Learning Mode classification (Lee et al., 2015). On the other hand, some other studies had the learning paradigm inferred based on their characteristics, as they do not mention it, such as Han et al. (2019), who explored mental models and dimensions of cognition and emotion. Table 2 summarizes the studies according to these paradigms that are predominantly represented based on our analysis.

2. **Mechanisms Influencing Learning Process (MILP)**: Represents the features, elements or actions used to influence or stimulate learning. From the behavioral perspective, the use of techniques to reinforce search behavior, which is learning-related (Harasim, 2011). From the cognitivist perspective, it is usually concerned

### Table 1. Results of the systematic conduction process.

<table>
<thead>
<tr>
<th>Source Name</th>
<th>URL</th>
<th>Step 1 (duplicates)</th>
<th>Step 2 (removed by criteria)</th>
<th>Step 3+4</th>
<th>Step 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACM Digital Library</td>
<td><a href="https://dl.acm.org/dl.cfm">https://dl.acm.org/dl.cfm</a></td>
<td>13</td>
<td>12</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>El Compendex</td>
<td><a href="http://www.engineeringvillage.com">http://www.engineeringvillage.com</a></td>
<td>36</td>
<td>21</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>Google Scholar</td>
<td><a href="https://scholar.google.com">https://scholar.google.com</a></td>
<td>297</td>
<td>39</td>
<td>228</td>
<td>38</td>
</tr>
<tr>
<td>IEEE Digital Library</td>
<td><a href="http://ieeexplore.ieee.org">http://ieeexplore.ieee.org</a></td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ISI Web of Science</td>
<td><a href="http://www.isiknowledge.com">http://www.isiknowledge.com</a></td>
<td>14</td>
<td>10</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Science@Direct</td>
<td><a href="https://www.sciencedirect.com">https://www.sciencedirect.com</a></td>
<td>144</td>
<td>3</td>
<td>137</td>
<td>4</td>
</tr>
<tr>
<td>Scopus</td>
<td><a href="http://www.scopus.com">http://www.scopus.com</a></td>
<td>31</td>
<td>21</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Springer Link</td>
<td><a href="http://link.springer.com">http://link.springer.com</a></td>
<td>33</td>
<td>8</td>
<td>21</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>577</td>
<td>116</td>
<td>396</td>
<td>65</td>
</tr>
</tbody>
</table>
Table 2. Summary of the studies according to learning paradigms.

<table>
<thead>
<tr>
<th>Paradigm</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behaviorist</td>
<td>(Lu and Hsiao, 2017), (Zhuang et al., 2016), (Mao et al., 2016), (Moraes et al., 2018), (Wilson and Wilson, 2013)</td>
</tr>
<tr>
<td>Cognitivist</td>
<td>(Kodama et al., 2017), (Moraes et al., 2018), (Taiibi et al., 2017), (Wilson et al., 2016), (Syed and Collins-Thompson, 2016), (Bhattacharya and Gwizdka, 2019), (Al-Tawil et al., 2019), (Azpiazu et al., 2017), (Crescenzi, 2016), (Han et al., 2019), (Liu and Song, 2018), (Johnson, 2018), (Jansen et al., 2007), (Smith and Rich, 2019)</td>
</tr>
<tr>
<td>Constructivist</td>
<td>(Ghosh et al., 2018), (Tibau et al., 2018b), (Freund et al., 2016), (Komlodi and Caidi, 2016), (Weingart and Eickhoff, 2016), (Tibau et al., 2018a), (Yu et al., 2018b), (Al-Tawil et al., 2019), (Ibieta et al., 2019), (Zapata et al., 2015), (Zhang, 2017), (Meyers, 2018), (Cho et al., 2017), (Vakkari et al., 2019), (Ibieta et al., 2019)</td>
</tr>
</tbody>
</table>

Table 3. Summary of the studies according to mechanisms that influence the learning process.

<table>
<thead>
<tr>
<th>MILP</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reinforcements</td>
<td>(Zapata et al., 2015)</td>
</tr>
<tr>
<td>Rewards</td>
<td>(Taiibi et al., 2017), (Gadiraju, 2018), (Zhuang et al., 2016), (Yu et al., 2018a), (Gadiraju et al., 2018)</td>
</tr>
<tr>
<td>Evaluation</td>
<td>(Rich et al., 2012), (Tibau et al., 2018b), (Liu and Song, 2018), (Johnson, 2018), (Smith and Rich, 2019), (Wilson and Wilson, 2013), (Vakkari et al., 2019)</td>
</tr>
<tr>
<td>Assistance or guidance</td>
<td>(Han et al., 2019), (Hinostroza et al., 2018), (Cho et al., 2017), (Moraes et al., 2018), (Ibieta et al., 2019)</td>
</tr>
</tbody>
</table>

3. Session Design for Learning (SDL): Represents the structural, environmental and interactive features of the session that are used to support and ensure learning. Under the lens of what learning is and how it occurs during an online search session, we are concerned about how session design features provide important information about what makes students more effective. These features are related to session control (Han et al., 2019), sharing (Meyers, 2018) and the kind of assistance. The search process environment converges to the learning process in some ways, such as an effective learning environment with a community-centered, knowledge-centric, student-centric, or overlapping (Anderson, 2011). Table 4 summarizes the papers according to the SDL perspective.

4. Recognizing Knowledge Gain

A major challenge of recognizing knowledge gain is through metrics, which is the topic of RQ2. Along with the evolution of the search mechanisms, ways to assess its processes
and interactions have emerged. We analyzed recent studies on how to estimate learning, identifying what we call as Measurement Records of Learning (MRL). It represents mechanisms and methods employed to indicate learning gains during search sessions by extracting data from interviews, search logs, self-reports, video recording, and pre and post-tasks (Gadiraju et al., 2018). It may involve predictive analysis, cognitive models, ontology and knowledge representation.

Predictive models, cognitive representation of search tasks and other methods may support the assessment of knowledge gain. Han et al. (2019) proposed a classification system for novice users’ mental models. Tibau et al. (2018b) investigated the decision-making process, which supports the understanding of the learner’s progress. Classical evaluations still predominate to measure the learning gain during the search process, evidenced by the use of pre- and post-session tests (Meyers, 2018). Liu and Song (2018) evaluated the knowledge points captured before and after the search tasks. Jansen et al. (2007) proposed the use of Anderson and Krathwohl’s Taxonomy of the cognitive domain and (Anderson et al., 2001) the revised Bloom’s Taxonomy of Educational Objectives. Alternatively, Moraes et al. (2018) used a vocabulary of learning tasks, and Smith and Rieh (2019) used metacognitive analysis to understand the knowledge-context enriching from retrieval results. Table 5 presents a summary according to MRL we identified.

### Table 5. Classification of Measurement records of learning (MRL).

<table>
<thead>
<tr>
<th>MRL</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre and post-tests</td>
<td>(Kiehl et al., 2012), (Meyers, 2018), (Gadiraju et al., 2018), (Cho et al., 2017)</td>
</tr>
<tr>
<td>Assisted Process</td>
<td>(Johnson, 2018), (Hinostroza et al., 2018), (Vakkari et al., 2019), (Ibieta et al., 2019)</td>
</tr>
<tr>
<td>Knowledge base</td>
<td>(Yu et al., 2018a), (Tibau et al., 2018b), (Liu and Song, 2018), (Gadiraju et al., 2018)</td>
</tr>
<tr>
<td>Ontologies or taxonomies</td>
<td>(Jansen et al., 2007), (Moraes et al., 2018), (Wilson and Wilson, 2013)</td>
</tr>
<tr>
<td>Cognition or mind models</td>
<td>(Han et al., 2019), (Smith and Rieh, 2019)</td>
</tr>
</tbody>
</table>

### 5. Concluding Remarks

Searching as Learning nowadays is pervasive in formal, informal and lifelong learning. Particularly, in distance (and remote) education, where the learner needs to be more engaged in the tasks, it raises the need for understanding and supervising this process. Then, mechanisms to evaluate learning progress must perform well in order to search as learning be considered as a reasonable strategy for pedagogical architecture. According to our view, learning theories and approaches should guide the development of the area in order to align with educational practices.
This paper presented a systematic literature review on SAL, focused on analyzing the different solutions for search as learning. The selected papers were analyzed based on two research questions: (i) the learning theories and paradigms that have been used, either directly cited or inferred from our analysis, and (ii) how knowledge gain has been measured in these studies. We believe that these items are of essential importance for connecting SAL with other areas of information systems for education, such as Learning Analytics and Educational Data Mining.

As future work, we consider a further analysis of learning theories and approaches according to education-related research areas to reinforce the importance of developing techniques to support learning. As a result, it would be possible to identify the intersection of the learning aspects applied by such approaches with the elements of the search process, highlighting which ones combine successfully or not to promote measurable learning gains. This correlation would guide the combination of new or revised learning approaches with more advanced search engines.

Acknowledgment
This study was financed in part by the ‘National Council for Scientific and Technological Development (CNPq) - Brazil’ - Process 315374/2018-7, Project ’Searching as Learning: the information search as a tool for learning’ and by the ‘Coordination for the Improvement of Higher Education Personnel’ (CAPES) – Brazil – Finance Code 001

References


Kodama, C., Jean, B. S., Subramaniam, M., and Taylor, N. G. (2017). There’s a creepy guy on the other end at google!: engaging middle school students in a drawing activity to elicit their mental models of google. *Information Retrieval Journal*, 20:403–432.


Meeting on Navigating Streams in an Information Ecosystem - Volume 47, ASIS&T ’10, USA. American Society for Information Science.


