

Enhancing Information Literacy Skills: A Game Design for Seeking Information and Making Queries

Paavo Arvola and Tuulikki Alamettälä

Tampere University, Tampere, Finland
forename dot surname at tuni dot fi

Abstract. Information literacy is an essential skill in today's society and information search is an essential part of it. These skills need constant training and maintenance and thus, innovative and engaging approaches to teach and learn information literacy and searching are needed. In this paper, we present a serious game to enhance information literacy skills, focusing on skills in seeking information. The experimental research setting widely used in search engine development provides means to gamify information seeking.

Keywords: Information Retrieval, Search Engines, Information Literacy.

1 Introduction and Related Work

Information literacy refers to the skills to recognize the information need, locate the information, evaluate the information found, and finally to use the chosen information ethically and effectively [15]. These abilities are widely seen as civic skills that everyone needs in today's society. Information search is an important part of information literacy. It is the first step in the interaction with available information [13]. Without relevant search results, it is difficult to proceed further: into evaluating and using information. Earlier research indicates that users often have problems with query formulation and reformulation [19]. An ability to use the search engine is essential, and thus the terms *search (engine) literacy* and consequently, *algorithmic literacy* have been adopted in the academic discussions. These kinds of literacies mean the ability to make effective use of a wide array of search engines, including a familiarity with their full functionality as well as their limitations [10].

The pedagogical approach of utilizing games in education is called *game-based learning* (GBL) or *digital game-based learning* (DGBL). Definitions of GBL mostly emphasize that it is a type of game play with defined learning outcomes. [18] Games that are used in (D)GBL are mainly *serious games*. A common definition of serious games is "games that do not have entertainment, enjoyment, or fun as their primary purpose" [16]. Serious games have the goal to transfer knowledge, to teach skills or to stimulate behavioural change.

DGBL has been found to enhance students' motivation, participation, confidence, perception, attitude and learning performance [7, 21]. Accordingly, several researchers have implemented educational games to help students develop information literacy.

The studies have shown positive effects [14, 22]. Games are widely accepted by new generations of learners and encourage creativity in information literacy training [9].

Competition is classified as a characteristic of a serious game [4]. Competition may have potential to optimize GBL. Earlier research has reported several benefits that competition in games may bring. Competition is considered as a useful technique to increase motivation and enhance students' learning by improving students' engagement and persistence in learning activities [6, 7, 22]. Competition has been found to be beneficial also in information literacy education [4]. Nonetheless, also some negative consequences of competition have been reported. Competition may cause feelings of anxiety, discouragement, frustration, and pressure [5]. However, virtual competition has been proposed to address these disadvantages as there is no direct competition between each player in virtual competition, and players compete with computers instead of real persons [8].

There are gamified approaches to information retrieval or search [17] of which our approach has similarities to *Query Performance Analyzer* [20], which was built upon a standard test collection of Finnish newspaper articles, topics, and related relevance assessments. The system did not consider judging documents' relevance, but it did measure query performance using the laboratory model and metrics available at that time. In a more recent approach called *Query Aspect Game* [8], relevance judgements for a topic are constructed with different queries and used for classification.

2 Game Design

We have designed a game that trains information search skills using multilingual Wikipedias. For technical details, see the demonstration publication in Arvola and Alamenttälä [1]. It has also been used in workshops (hackathon) e.g., in Arvola and Alamenttälä [2] and in undergraduate level university education.

The aim is for the teachers (hosts) to facilitate competitions for students (players) to search information from Wikipedias on a certain topic and measure the outcome, enabling scores for individuals or groups and provide leaderboards. Our game mechanics has its theoretical foundation in so called *Information Retrieval Laboratory evaluation model* [11, p. 4-5] and evaluation of information interaction [13].

In general, the laboratory model aims to measure and quantify the quality of the search results delivered by a search engine based on query and matching method. The main components of the framework consist of (1) a (test) collection of documents to search, (2) test topics and corresponding queries, (3) relevance judgements, i.e., data on which documents in the collection are relevant for the given topic, and (4) evaluation metrics. Our game design is an application of the laboratory model where students are players, and a teacher works as a host who selects, modifies, or creates topics for the players to play. Instead of competing algorithms or search engines, the players' queries about a topic are measured with a given search engine. So, the query is the search method to be measured.

First, the host – such as a teacher – develops a topic according to his or her preferences. The host can use topics that have been made earlier by someone else or create an own.

For example, the host creates a topic named “Vaccination in preventing diseases to spread”¹. In addition to the topic name, the host adds a topic description for the players (the students) to grasp what kind of documents they should search for. The language of the game is set here as well. As Wikipediae are considered here as document collections, virtually any language is viable, and the relevance judgements translated to other languages [1]. After setting the topic, the teacher judges the result documents relevant or not relevant according to his/her expertise using the assessment tool (Similar to Fig 5). It is worth noting that the relevance judgements are usable ever after.

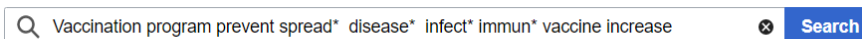


Fig 1. Standard search box. Best performing query for the Vaccine topic in the English Wikipedia edition

The user input is by default a standard search box (Fig 1). This allows the player to freely insert any query they see fit following the *query language* provided by the search engine specifications². Another alternative is to restrict the possibilities by offering a fixed set of query words and let the player choose the best possible combination of them (Fig 2).

Fig 2. An alternative form of game. Instead a query box, the player is given a restricted set of possible query expressions. Here for the German edition of the Vaccine topic the user is required to select three words for the query. (Note a misspelled word).

The game delivers feedback based on the player’s query performance in relation to the gold standard (relevance judgements). The feedback is immediate and comes in the form of precision-recall measures as well as corresponding diagrams, such as Euler diagrams (Fig 3) and precision-recall curves.

¹ Topic translated from Finnish matriculation exam

² <https://www.elastic.co/elasticsearch/>

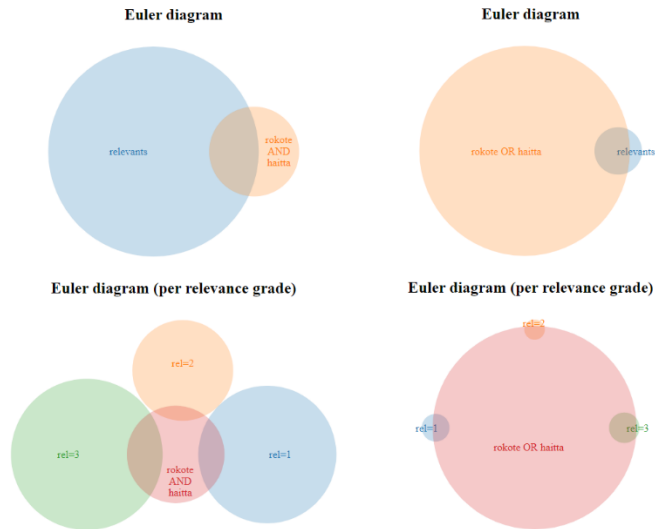


Fig 3. Euler diagrams³ as feedback for the player on how two different result sets overlap with the relevant document sets (binary and graded) using the queries “rokote AND haitta” and “rokote OR haitta”.

Obviously, there are many other metrics available [3], of which the widely used Cumulated Gain (CG) [12] metrics for ranked lists is used in this study. This metrics supports graded relevance used in our game, while some relevant documents are more relevant than some others. A normalized CG curve is shown to the player (Fig 4).

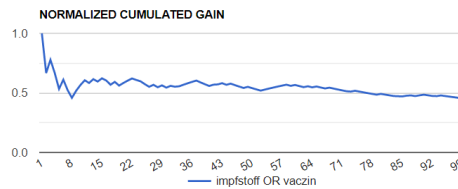


Fig. 4. Normalized Cumulated Gain [12] curve of a query evaluation.

Players themselves may judge the documents as well (Fig 5). When players contribute to the relevance judgements in a gamified and *collaborative* effort, it results in more complete recall bases.

³ <https://github.com/benfred/venn.js>

37 hits

No	Document	Relevance score (Graded) (Subst score)	Cumulated Gain
1	Koronavirustaudit Koronavirustaudit todettiin ensiksi SARS-CoV-2-viruse, kalliolla oireisiksi, joka aiheaa vuonna 2019 maantieteellisesti laajalle levinneen SARS-CoV-2 tartunnan	3	1
2	FINCA:n Alueella 17 miljoonaa suomalaisella koronavirusleikkiä - verkkosivut.6. 25.6.2020. Viitattu 25.6.2020. Kotonaan koronavirus voi isoa robotillaan ihmiskokoon	4	4
3	SARS-CoV-2-palautus leikkely Tutkimus: Madrasan koronavirus on tavallinen Veikkosivut: 15.12.2020. Viitattu 15.12.2020. Myös Madrasan koronavirus on tavallinen Veikkosivut	2	6
4	Oxford-Amerikan COVID-19 leikki Viitattu 24.12.2021. Arvola, P., Alamettälä, T.: Koronavirus & Johonnan koronavirus - kaksi oireisiksi etua The Timesin 11.3.2021. Viitattu 14.12.2021	3	9
5	Johonnan & Johonnan Koronavirus, Terveystieteiden tutkimuskeskus 2021. Janssenin koronavirus koronavirus (16.45) -vastausta koronavirus adenoiruksen tutkimus	3	8

Rank

Link to document and snippet

Graded relevance score (0,1,2,3 or ?)

Relevance judgement for collaboration

Fig. 5. Result list presented to the player after executing a query. The host can use the same view for judging the relevance.

3 Discussion and Conclusions

The laboratory research setting is widely used in search engine development, spam filtering, (text) classification, and recommender systems, and many others. To provide a realistic, familiar and inspiring environment, we have adopted the online Wikipedias as the test collections and a freedom for the host to create or use virtually any topic - or language - for the players to play with. In our game design query or seeking success is measured explicitly, which enables competitive elements with leaderboards, see [4].

Consequently, the game presented in this paper is a challenge, where the players control the game with queries and exploit the explicit and measurable feedback. This results preferably in growing competence after seeing and comparing the different search strategies.

Acknowledgements. This work was partially funded by the European Union Erasmus+ programme with grant number 2021-1-FI01-KA220-SCH000029713

References

1. Arvola, P., Alamettälä, T.: IRVILAB: Gamified Searching on Multilingual Wikipedia. In: Proceedings of ACM SIGIR '22 Conference, pp. 3329–3333. ACM, New York (2022). <https://doi.org/10.1145/3477495.3531662>
2. Arvola, P., Alamettälä, T.: Information retrieval workshop. In: 30th BOBCATSSS Symposium - Book of Abstracts, p. 43. (2022). <https://dx.doi.org/10.5281/zenodo.6484810>
3. Butcher, S., Clarke, C.L., Cormack, G.V.: Information retrieval: Implementing and evaluating search engines. Mit Press (2016).
4. Cagiltay, N. E., Ozelik, E., Ozelik, N. S.: The effect of competition on learning in games. Computers & Education 87, 35–41 (2015). <https://doi.org/10.1016/j.compedu.2015.04.001>
5. Chen, C. H., Liu, J. H., Shou, W. C.: How competition in a game-based science learning environment influences students' learning achievement, flow experience, and learning behavioral patterns. Journal of Educational Technology & Society 21(2), 164–176 (2018).
6. Chen, C. H., Shih, C. C., Law, V.: The effects of competition in digital game-based learning (DGBL): A meta-analysis. Educational Technology Research and Development 68(4), 1855–1873 (2020). <https://doi.org/10.1007/s11423-020-09794-1>

7. Chen, S. Y., Chang, Y. M.: The impacts of real competition and virtual competition in digital game-based learning. *Computers in Human Behavior* 104, 106171 (2020). <https://doi.org/10.1016/j.chb.2019.106171>
8. Di Nunzio, G. M., Maistro, M., Zilio, D.: Gamification for IR: The Query Aspects Game. In *CLiC-it/EVALITA*. (2016).
9. Encheva, M., Tamarro, A. M., Kumanova, A.: Games to Improve Students Information Literacy Skills. *International Information & Library Review* 52(2), 130–138 (2020). [10.1080/10572317.2020.1746024](https://doi.org/10.1080/10572317.2020.1746024)
10. Hockly, N., Dudeney, G., Pegrum, M.: *Digital literacies*. Routledge, London (2014).
11. Ingwersen, P., Järvelin, K.: *The Turn - Integration of Information Seeking and Retrieval in Context*. The Kluwer International Series on Information Retrieval (2005).
12. Järvelin, K., Kekäläinen, J.: Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems* 20 (4), 422–446 (2002). <https://doi.org/10.1145/582415.582418>
13. Järvelin, K., Vakkari, P., Arvola, P., Baskaya, F., Järvelin, A., Kekäläinen, J., Keskustalo, H., Kumpulainen, S., Saastamoinen, M., Savolainen, R., Sormunen, E.: Task-Based Information Interaction Evaluation: The Viewpoint of Program Theory. *ACM Trans. Inf. Syst.* 33(1) 1–30 (2015). <https://doi.org/10.1145/2699660>
14. Li, Z., Zou, D., Xie, H., Wang, F. L., Chang, M.: Enhancing information literacy in Hong Kong higher education through game-based learning. In Jong, M., Shih, J.L., Looi, C.K., Huang, M., Chen, X., Xie, Y., . . . Tu, S. (eds.) *The 22nd Global Chinese Conference on Computers in Education: Conference proceedings*, pp. 595–598. South China Normal University, Guangzhou (2018).
15. Library and Information Association. What is information literacy? (2018). Retrieved from <https://www.cilip.org.uk/page/informationliteracy>
16. Michael, D. R., Chen, S. L.: *Serious games: Games that educate, train, and inform*. Thomson Course Technology, Boston (2005).
17. Muntean, C. I., Nardini, F. M.: Gamification in Information Retrieval: State of the Art, Challenges and Opportunities. In: *Proceedings of the 6th Italian Information Retrieval Workshop (IIR'2015)*. (2015).
18. Plass, J. L., Homer, B. D., Kinzer, C. K.: Foundations of game-based learning. *Educational psychologist* 50(4), 258–283 (2015).
19. Rieh, S.Y. Xie, H.: Analysis of multiple query reformulations on the web: The interactive information retrieval context. *Information Processing & Management* 42(3), 751–768 (2006). <https://doi.org/10.1016/j.ipm.2005.05.005>
20. Sormunen, E., Hokkanen, S., Kangaslampi, P., Pyy, P., Sepponen, B.: Query performance analyser- a web-based tool for IR research and instruction. In: *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 450-450. (2002).
21. Subhash, S., Cudney, E. A.: Gamified learning in higher education: A systematic review of the literature. *Computers in human behavior* 87, 192–206 (2018). <https://doi.org/10.1016/j.chb.2018.05.028>
22. Zou, D., Zhang, R., Xie, H., Wang, F. L.: Digital game-based learning of information literacy: Effects of gameplay modes on university students' learning performance, motivation, self-efficacy and flow experiences. *Australasian Journal of Educational Technology* 37(2), 152–170 (2021).