Open Educational Resources Platform Based on Collective Intelligence

Marcela Ribeiro de Oliveira[1], Israel Barreto Sant'Anna[2], Guilherme Scariot Ramos[3],
Luis C. E. De Bona[4], Marcos Alexandre Castilho[5], Marcos Didonet Del Fabro[6] and Eduardo Todt[7]
Department of Informatics, Federal University of Paraná
Curitiba - Paraná, Brazil
{mro15[1], ibsantanna[2], gsr16[3], bona[4], marcos[5], marcos.ddf[6], todt[7]}@inf.ufpr.br

***Abstract*—Open Educational Resources (OER) are educational resources openly available to be used by educators and students and are an important tool to support education. A considerable effort has been made to build repositories that allow the sharing and reuse of these OERs. However, many of these repositories offer unsatisfactory search engines, resulting in a frustrating experience for users. The problem of content search is partly explained by the lack of appropriate metadata on resources and the lack of ranking mechanisms, for instance based on user interaction over the educational objects. In this paper, we present the *Plataforma Integrada do MEC* (MEC's Integrated Platform), a novel Digital Educational Resources platform, which employs concepts of social networks to create a collective intelligence system to improve and refine search results. Platform users are able to evaluate the available resources and also publish new ones. A ranking is associated with these users and it is used to determine the relevance of their actions on the platform. High ranking users can publish new content that will be considered relevant in the platform, eliminating the need for resources to be evaluated by an expert. Unlike other repositories, where authorization must be given for content publication, on this platform the user has the opportunity to actively contribute with new content. A prototype of this platform was developed, and made available as free software. Preliminary results indicate viability of the proposal and that the system opens a path to effective knowledge diffusion.**

***Index Terms*—open educational resources, educational platform, collective intelligence, social networks**

I. Introduction

Open Educational Resources (OER) describes any educational resources that are openly available for use by educators and students, without any need to pay royalties or licence fees. Their transformative power lies in the ease that such resources can be shared via the Internet [1].

An OER is characterized by its metadata. Some representative examples of metadata are name, creation date, tags and description. The OERs are stored in repositories. The purpose of OER repositories is to support educators in searching for content in a structured way, sharing their own resources, reusing existing materials and creating new resources through adapting or translating, and in collaborating with other members of the user community by commenting upon, reviewing, promoting and developing resources [2].

There are currently several digital repositories in Brazil where educational resources are stored and available for teachers and students. Many of them are maintained by public institutions, such as Federal or State Governments, e.g. *Portal do Professor*[1] (Teacher's Portal), *Banco Internacional de Objetos Educacionais*[2] (International Bank of Educational Objects), *Domínio Público*[3] (Public Domain), *Dia a Dia Educação*[4] (Day-to-day Education). Unlike other user-collaborated sources such as Wikipedia, where content is often explained in-depth, regardless of the reader's age and expertise, educational resources in these repositories are categorized by educational stages, such as elementary, middle and high school, giving both students and teachers access to content that is more properly suited to their current needs.

Despite this wide availability of educational resources, the task of selecting relevant educational resources is arduous and exhaustive to teachers and students [3]. According to Atenas, Havemann and Priego [4], searching and retrieving OER from repositories can be a challenging task as materials are difficult to find, retrieve and sometimes impossible to download in order to be adapted, translated or updated.

The aforementioned repositories are an example of how cumbersome this task can be. Their content can be of mixed quality, with several resources lacking important descriptive metadata which, allied to the often sub-optimal indexing from their search engines, results in the user having to sort through many irrelevant content when querying for specific educational resources. Since it is often easier for users to search for content on their preferred search engine, these repositories are rarely accessed.

Furthermore, the frequency with which new content is added to these repositories is very low, mainly because there is a great amount of bureaucracy to grant authorization to publish a new educational resource in the existing repositories. This bureaucracy intends to prevent the publication of inappropriate content, such as offensive or inaccurate materials, but ends up discouraging teachers from creating and sharing their own educational resources, which leads to a increased lack of interest in these repositories.

The platform's goal is to provide an interactive web platform with social network features, to ease the publication, search and ranking of open educational resources through a collective

[1] http://portaldoprofessor.mec.gov.br/index.html
[2] http://objetoseducacionais2.mec.gov.br/
[3] http://www.dominiopublico.gov.br/
[4] http://www.diaadia.pr.gov.br/

intelligence. Although collective intelligence may bring to mind the idea of group consciousness, in computer science it usually means the combining of behavior, preferences, or ideas of a group of people to create novel insights [5].

This collective intelligence is mainly present on the platform through a ranking system, where the community interaction with the available content directly affects the ranking of both resources and creators, resulting in a ranking system where the higher the score, the more reliable the user or the content is. This ranking is useful to sort the items retrieved by the search engine, showing the users the more relevant and reliable items first.

Beyond that, the ranking can also be used to help avoid the bureaucracy when publishing an educational resource. The platform trusts in the community to decide which user is allowed to publish resources, giving high ranking users the ability to publish and invite new users as publishers. The new resources go through an automatic filtering, where machine learning algorithms try to further mitigate the presence of inappropriate content. The result of this filtering is then validated by high ranking users, to avoid any wrongly classified resource.

Since it is a public and educational focused platform, brazilian teachers who sign up receive an initial ranking high enough to allow them to publish resources directly after registration.

The platform also intends to reduce the lack of metadata in resources by letting the users of a certain rank to enrich the information of the resources, with social tagging being one of the main tools in this regard, since because of their lack of predefined taxonomic structure, social tagging systems rely on shared and emergent social structures and behaviors, as well as related conceptual and linguistic structures of the user community [6], improving the chances of people finding relevant resources through different terms.

All of this features aim to harness the power of the collective intelligence to create a more attractive OER platform for teachers, rewarding them for their participation while improving the overall quality of the contents.

The initial platform prototype was developed and made available as free software, through a partnership with the *Ministério da Educação* (Ministry of Education - MEC).

We describe the platform main components, their interactions and our technological choices. Then, we ran stress tests to verify if it would scale when having a larger user base and number of interactions. The deployed prototype integrates the content from some of their more traditional repositories, together with new resources published by users. The platform currently has more than 30,000 published educational resources and 5,000 registered users, with an average of 1,200 resources downloaded per month. These numbers indicate the viability of a collaborative approach to OER repositories. It is available at http://plataformaintegrada.mec.gov.br/.

II. Architecture

The architecture consists of three main components: Web Service, Storage and User Interface, each of these components performs specific functions. Figure 1 presents this architecture.

The Web Service component consists of the platform implementation and it is responsible for communicating it with different applications. The User Interface is responsible for connecting users with the web service and allows them to interact with the platform in an intuitive way. The Storage component takes care of all data storage, such as educational resources metadata and information about users registered on the platform.

The following subsections explain in detail each of these components.

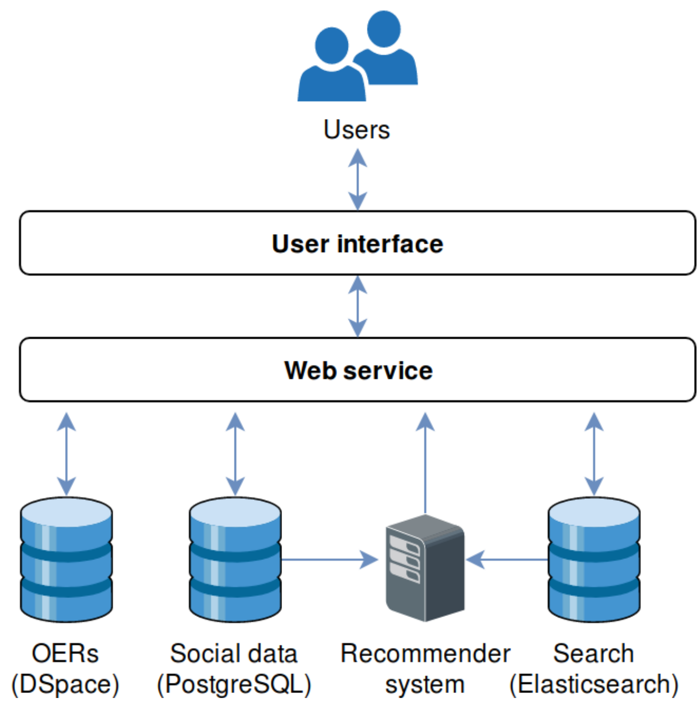

Fig. 1. Platform architecture overview

A. Web Service

Representational State Transfer (REST) is an architectural style for distributed hypermedia systems. REST provides a set of architectural constraints that, when applied as a whole, emphasizes scalability of component interactions, generality of interfaces, independent deployment of components, and intermediary components to reduce interaction latency, enforce security, and encapsulate legacy systems [7].

Web services are purpose-built web servers that support the needs of a site or any other application. Client programs use application programming interfaces (APIs) to communicate with web services. Generally speaking, an API exposes a set of data and functions to facilitate interactions between computer programs and allow them to exchange information. A Web API conforming to the REST architectural style is a REST API [8].

The platform was implemented as a REST API, facilitating the communication with different interfaces, such as websites, mobile apps and software packages.

Its source code was developed with Ruby on Rails and is available under the GNU/AGPL license at https://gitlab.c3sl.ufpr.br/portalmec/portalmec/.

The documentation of the API endpoints can be found at https://api.portalmec.c3sl.ufpr.br/docs-new/.

B. Storage

To better accommodate the different aspects of the project, the database structure was divided into three different parts.

The storage of educational resources and its files was assigned to DSpace, an open source repository application that allows the capture, storage, indexation, preservation and distribution of digital material, making the digital archiving an easy task and guaranteeing the long-term preservation of the resources [9]. It uses a relational database to store information about the organization of content and the resources metadata [10]. Beyond that, DSpace also implements the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), a low-barrier mechanism for repository interoperability [11], allowing the platform to easily integrate several OER repositories.

The search was delegated to Elasticsearch, a distributed, RESTful search and analytics engine, allowing for fast and reliable search results [12]. Besides being a search engine, Elasticsearch is also a document oriented database [13], storing the relevant searchable information of all available resources.

The main database, where all the remaining relevant data is saved, was stored in PostgreSQL, an open source object-relational database system, because of its highly relational model, consequence of the social network characteristics of the project. PostgreSQL also showed itself as a reliable Database Management System, capable of holding the high amount of data throughput required for the project and maintaining availability [14]. User interaction data, such as OER views and downloads, is shared with the recommender system and also includes actions from anonymous users. Elasticsearch also provides the recommender system with information about metadata similarities between educational resources.

Figure 2 shows the simplified relational database model from the main database, containing the more social relevant models. The main element is the *LearningObject* table, which stores the educational resources information, such as description, author, metadata, and publication date. The *Collection* table stores user's collections of educational resources, which are discussed in more details in Section III-A. The *User* table stores all the user information and relates to the *LearningObject* and *Collection* tables through many actions that can be taken in the platform, such as *Follow*, *Review*, *View*, *Like*, *Share* and *Download*.

Data Synchronization: Since each of these initially three independent parts have their own database, they need to be synchronized for the whole architecture to work.

An educational resource can be inserted in the platform in two ways: through the REST API, being stored in the main PostgreSQL database, or through the DSpace OAI-PMH Harvester. In both cases, the Ruby on Rails server propagates

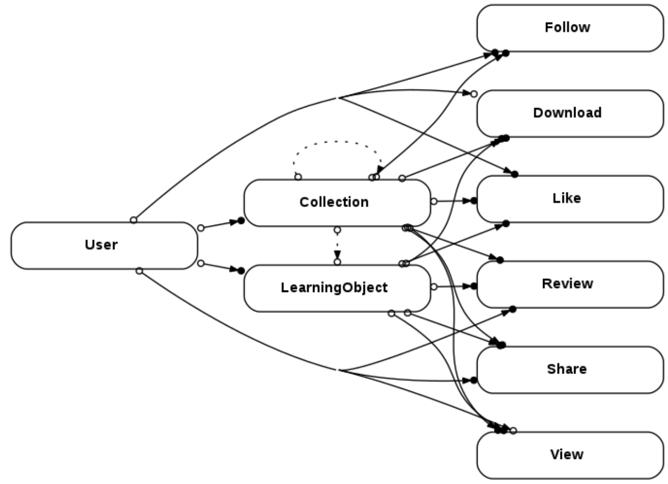

Fig. 2. Social Network Simplified Database Model

the new resource to the other databases, with the difference that in the latter, the server needs to check the DSpace repository for new resources and retrieve them.

This means there is a two-way synchronization between DSpace and the PostgreSQL database and a one-way synchronization between the PostgreSQL database and the Elasticsearch database.

The Rails server uses the Searchkick library to communicate with Elasticsearch, sending all the relevant information about the resource. For the DSpace communication there was no pre-existing Ruby library, so the DSpace Rest Client [5] library was created for this project, using the DSpace REST API to manage all the resources and synchronize the databases.

C. User Interface

The main user interface to the platform can be accessed at https://plataformaintegrada.mec.gov.br/. It was developed in JavaScript with the AngularJS framework. Figure 3 shows the homepage, where the search field is emphasized. On the homepage it is also possible to view the newest educational resources and collections.

Figure 4 shows an educational resource page, where a user can see its title, description, author, language, size and tags. On this page it is possible to download, like, review and report the resource, following the specification of resources' actions.

III. OER Collective Intelligence

The platform tries to use collective intelligence methods to find, publish and rank OERs. It also encourages community participation with social network features such as following and reviewing. A score number is attributed to each user and OER. Each score is calculated by our custom ranking system as the results are shown in descending order. The recommender

[5] *https://github.com/C3SL/dspace-rest-client*

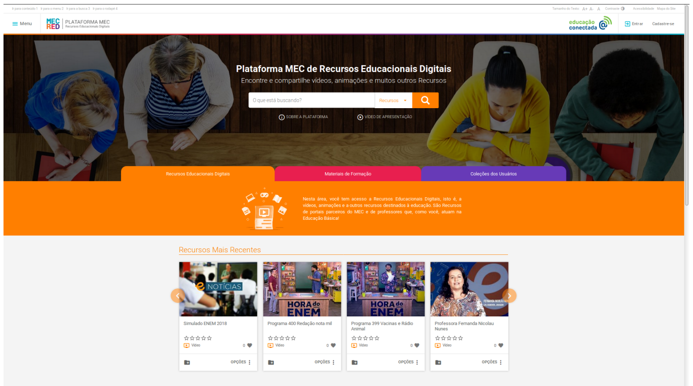

Fig. 3. Platform homepage

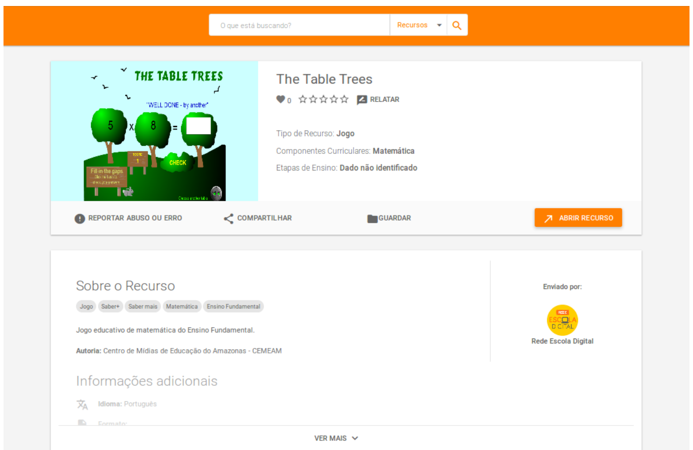

Fig. 4. Viewing an educational resource

system applies *hybrid filtering* methods to retrieve recommendations that are related to the user interaction data. These functionalities are described in the following sections.

A. Social Network Features

Following Users: If a user likes the content published by another user, it is possible to follow him to receive notifications about his activities, such as new resources and collections created by this user.

Liking Educational Resources: If a user enjoys a certain content he might express his appreciation by liking this content. This action is interpreted by the platform as an indicator of content that the user likes or not, this indicator is used to recommend similar resources to the user. Also, content liked by many users has its ranking increased.

Tagging Educational Resources: The platform allows users to freely assign keywords, so-called tags, to the resources or collections.

Allowing individuals to apply free text keywords to digital objects, potentially offers advantages in terms of personal knowledge management, serendipitous access to objects through tags, and enhanced possibilities to share content with emerging social networks among other users [15].

Reviewing Educational Resources: A user can tell their experience with an educational resource through a review. This review will contribute to the score of the educational resource and its publisher.

Creating Collections: Educational resources can be grouped in collections, allowing users to save and organize resources in any way that seems relevant to them. For instance, a teacher could create a collection for one of his classes, grouping together educational resources that could improve the students interest and comprehension of the class subject.

Besides educational resources, collections can also contain other collections, creating a hierarchy of collections similar to a directory tree, which helps users to keep their collections and resources organized and easy to find and share. In the previous example, the teacher could save his collection inside a more general collection about a subject, and create more collections for different classes inside it.

These collections can be private or public, so the teacher can choose to keep his collection to himself or share it with other users, endorsing the quality of those resources and explaining the purpose of the collection in the description. Users can follow public collections to be notified of resources added or removed from the collection.

B. Collaborative Ranking

Each educational resource, collection or user in the platform is associated with a score. The score of collections and resources are based on the reputation of their creators and the evaluations and interactions made by the community. The user score, on the other hand, is based on his interactions in the platform and the score of the resources and collections created by him. This creates a self-balancing ranking system where the quality of the content, as implied by the community interaction, influences the score of the creator and his future content.

The result is a ranking system where the higher the score, the more reliable the user, collection or resource is, making the ranking useful to sort the items retrieved by the search engine, and to choose users for tasks of greater responsibility, such as filtering inappropriate resources.

The goal is to constantly improve the resources, collections and users rankings; using collective intelligence to classify the educational contents.

Score: Educational resources and collections have the same score parameters, which are: thumbnail presence, description presence, number of views, number of likes, number of downloads, number of shares, reviews rating average, owner score and inclusions in collections by other users.

Meanwhile, users have the following parameters: owned educational resources included in collections by other users, average rate received in the reviews created by the user, average rating of the reviews from the user's educational resources and collections, best score between all the user's educational resources and collections, number of educational resources published by the user, number of followers, average score from the user's educational resources, average score from the user's collections and number of recently created resources.

The final score is obtained by the sum of the normalized scores of each parameter multiplied by a pre-defined weight, as shown in Equation 1. For instance, if an educational resource has 15 views, while the one with most views has 300 and the weight of the view parameter is 3, the view score will be 0.15, and the educational resource score will be the sum of all its parameters scores. The value for the parameters weights were defined empirically and can be changed if needed.

$$Score = \sum_{FisrtParameter}^{LastParameter} Weight * \frac{ResourceParameterScore}{BestParameterScore} \quad (1)$$

C. Recommendation System

The recommendation system plays a central role on the collaborative aspect of the platform. Since each user interaction with the website is stored on the database, it is possible to use this data to produce personalized OER recommendations, based on the user's previous activity. This is specially useful to bring relevant new resources to the user's attention, allowing them to see content that may be harder to find due to their initial low ranking. A *hybrid recommendation system* was implemented, which consists of a *collaborative filter* and a *content-based filter*.

Collaborative filtering: Collaborative filtering is a technique that tries to produce recommendations for a specific user, based on how others users have interacted with the available educational resources [16], that is, such an algorithm tries to find correlations between actions and users to produce recommendations. Every user collaborates to the system – with every new interaction with the website, the platform is able to ingest more data and improve the recommendation results.

In order to implement a collaborative filtering algorithm, Apache Foundation's PredictionIO technology was chosen, which provides an architecture for the development and deployment of machine learning algorithms [17]. The built-in implementation of the *Alternating Least Squares* (ALS) algorithm produced relevant results for users with a considerable amount of recorded data, but it showed some limitations, such as not being able to update recommendations without recalculating the entire ratings matrix. Although there are some proposed solutions to this problem [18], Apache Mahout's *Correlated Cross-Occurrence* algorithm was chosen, since it better fits data with multimodal user-item interactions [19] and avoids this intrinsic limitation of the base ALS algorithm.

Content-based filtering: Since *collaborative filtering* algorithms cannot produce good recommendations when there is not enough data about a user [20] (e. g., a user creates a new account, commonly known as the cold-start problem), a *content-based filter* was implemented to complement the collaborative results. The idea is to utilize both the OER's *metadata* as well as the user's recent history to produce recommendations. To find similar resources based on *metadata*, such as description and title, an effective yet simple solution is the *term frequency–inverse document frequency* (tf-idf) statistic, which is implemented as a function in Apache Lucene, a library that is available through Elasticsearch.

D. Searching Open Educational Resources

A user can search for OERs through a simple interface, and can be filtered by the following fields:

- Subject: different themes that an OER addresses, such as art, philosophy or mathematics
- Resource type: OER content type, such as images, maps, text or audio
- Educational stages: specifies the educational stage for which the OER is recommended, such as elementary, middle or high school
- Tags: OERs can be filtered through tags

Through these filters it is possible, for example, to find all the educational resources that contain images about the hydrologic cycle that are related to biology.

The results are ranked in descending order by the OER's score multiplied by the search score, which takes into account how relevant an OER is for the searched words and applied filters. This method produces better search results than Elasticsearch's default scoring, since it has additional, community-contributed parameters.

It is also possible to search for OERs with specific terms (tags) through a search method that integrates the results from the default search with the results from a search based on term clustering, which uses a set of metrics adapted for OERs [21]. This previous work showed, through empirical tests, that adding the term clustering ranking in the search results may yield more significant resources for a given term.

Once found, an OER can be downloaded, allowing its use in places that do not have a good internet connection, which is the case for many schools in the brazilian public school system. Collections can also be searched and downloaded.

IV. Experiments

In order to exemplify the impact of the collaborative ranking system in the search results, the same queries were searched and sorted using two different score functions: the default Elasticsearch score, which only takes into account the resource's metadata, and the same score multiplied by the educational resource ranking score. The ids for each of the first 10 results from these searches are presented in this Section, together with their respective Elasticsearch score and Collaborative ranking score.

The results for searching "DNA" and ranking with only the Elasticsearch score are shown in Table I, while the results for the same query, but using as ranking the Elasticsearch score multiplied by the collaborative ranking score are shown in Table II. As highlighted in the tables, some OERs are retrieved by both searches, but are presented in different positions due to the ranking score. By looking at some parameters from these resources is possible to see why an OER absent from Table I is shown in first place on Table II. The OER 6056 had a review score average of 4, while the other resources had no review at all, showing that the community finds its content more noteworthy. The other changes in positions are due to the OERs owners scores, improving the rank from the OER published by more reliable users.

TABLE I
SEARCHING "DNA" WITH RESULTS RANKED BY THE ELASTICSEARCH SCORE

Query = DNA			
Rank	OER id	Elasticsearch Score	Ranking Score
1	22636	111.90484	0.11074
2	23177	110.46753	0.106383
3	23330	110.46753	0.110243
4	22098	107.954475	0.106825
5	15717	105.92194	0.042553
6	23218	105.32358	0.106752
7	23086	104.27601	0.106567
8	21563	103.58426	0.106567
9	16408	101.81194	0.042553
10	13044	100.841064	0.071125

TABLE II
SEARCHING "DNA" WITH RESULTS RANKED BY THE ELASTICSEARCH AND RANKING SCORE

Query = DNA				
Rank	OER id	Elasticsearch Score	Ranking Score	Final Score
1	6056	87.750875	0.240522	21.106016
2	22636	111.90484	0.11074	12.392342
3	23330	110.46753	0.110243	12.173381
4	23177	110.46753	0.106383	11.747148
5	22098	107.954475	0.106825	11.551255
6	23218	105.32358	0.106752	11.262893
7	23086	104.27601	0.106567	11.126675
8	21563	103.58426	0.106567	11.053612
9	21159	100.313049	0.106881	10.721559
10	20567	98.897258	0.106383	10.520987

The results for searching "Água" (water) and ranking with only the Elasticsearch score are shown in Table III, while the results for the same query, but using the combined scores from Elasticsearch and the collaborative ranking score, are shown in Table IV. It is possible to see that the collaborative ranking scores in Table IV are higher than the ones in Table III, resulting in no intersection between the OERs. This happens because almost all the resources returned by the combined scores ranking were reviewed or liked, while the resources returned by Elasticsearch's default ranking were not. This could potentially result in resources that have no relation with the searched term showing first, but since the Elasticsearch score is part of the final score, the metadata from the resources must still have some relation to the query to show in the results.

V. CONCLUSION AND FUTURE WORK

We have presented an Integrated Search Platform for Open Educational Resources. The goal of the Integrated Platform is to shift the paradigm of educational content portals previously developed from a centralized view into a distributed approach. Mechanisms for resource ranking, user reputation and personalized recommendations based on user interaction data were added. The platform has several features that enable storing, organizing and than searching for educational resources, intending to provide a better experience and collaboration between users, not only a content-based engine, as in most part of existing portals. The first experimental results are promising – the use of collective intelligence formed by the ratings and actions of the thousands of users offers educators and students more adequate content than a simple generic web search.

The future of the platform lies in improving the artificial intelligence algorithms that automatically filter the educational resources, using different modalities, such as audio, text and image, in order to make improved decisions and reduce the number of resources needed to be validated by the users. The information learned by such algorithms can also be used to automatically set some of the resource's metadata, which will also be improved by the community through tools that facilitate metadata collaboration, further improving resource quality and the relevance of search results.

TABLE III
SEARCHING "ÁGUA" WITH RESULTS RANKED BY THE ELASTICSEARCH SCORE

Query = Água			
Rank	OER id	Elasticsearch Score	Ranking Score
1	4665	112.41242	0.042553
2	8861	107.54766	0.042553
3	11557	106.20375	0.042553
4	1962	106.20375	0.042996
5	12923	105.669266	0.042867
6	15514	105.669266	0.042553
7	5592	104.33827	0.042553
8	11584	104.33827	0.042553
9	18353	104.33827	0.042553
10	18634	104.33827	0.042682

TABLE IV
SEARCHING "ÁGUA" WITH RESULTS RANKED BY THE ELASTICSEARCH AND RANKING SCORE

Query = Água				
Rank	OER id	Elasticsearch Score	Ranking Score	Final Score
1	19133	36.710327582	0.597865	21.94782
2	13919	66.093416176	0.315941	20.88162
3	1279	100.837453097	0.199081	20.074821
4	10713	55.812713715	0.26554	14.820508
5	15185	67.426115707	0.218673	14.744271
6	166	64.022185141	0.215144	13.773989
7	6790	71.321551128	0.170766	12.179296
8	32473	76.540031712	0.114784	8.785571
9	29630	62.553600595	0.137144	8.578851
10	28875	80.336023347	0.106567	8.561169

APPENDIX

In order to better visualize the difference between the search results from the experiments reported in Tables III and IV, this appendix shows the results returned by both searches in the user interface.

Figure 5 shows the educational resources returned by the search using only Elasticsearch's score. Figure 6 shows the educational resources returned by the search using the Elasticsearch score multiplied by the collaborative ranking. In these figures, each tile is an educational resource, the image is the resource thumbnail, the text under the image is the resource name, the stars represents the resource reviews score average and the number on the left of the heart icon is the amount of likes received by the resource.

It is possible to notice that the search results are improved when using collaborative ranking. None of the results have a missing thumbnail and the educational resources with higher review score averages and number of likes are returned first.

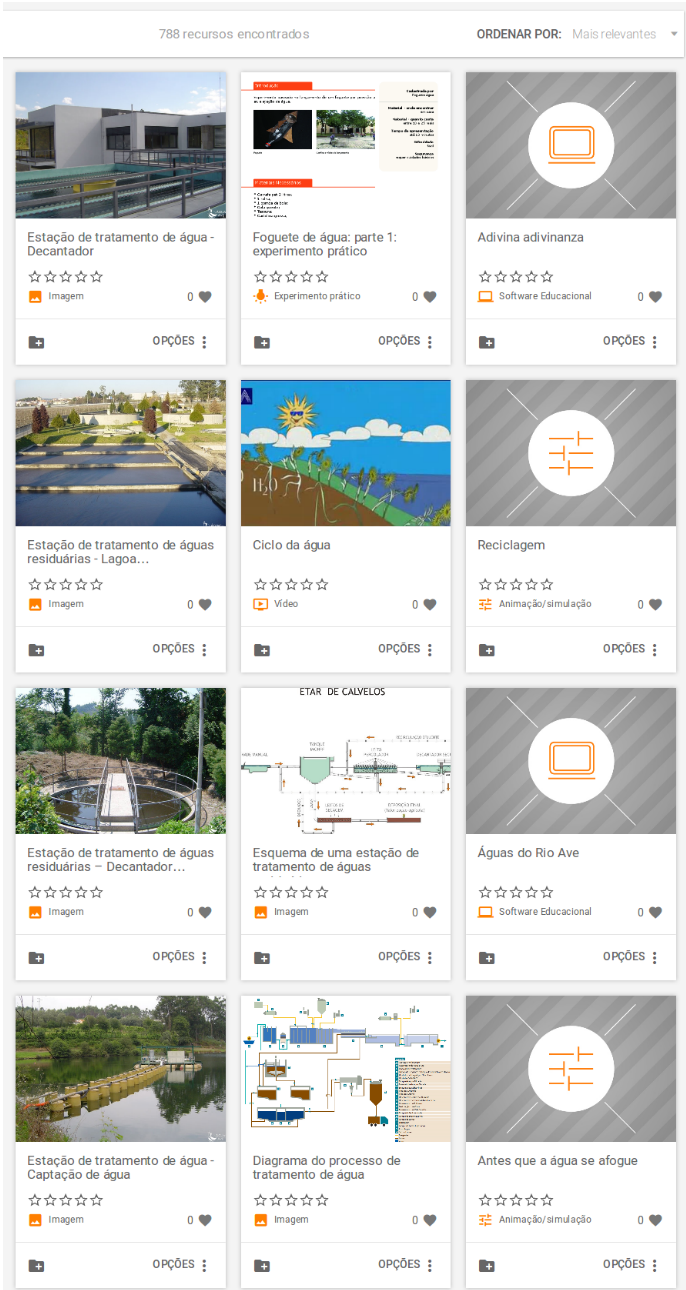

Fig. 5. Results from search using only Elasticsearch score.

ACKNOWLEDGMENT

The *Plataforma Integrada do MEC* is the outcome of a research project conducted by the *Centro de Computação Científica e Software Livre* (Scientific Computing and Free Software Center - C3SL) from the *Universidade Federal do Paraná* (Federal University of Paraná - UFPR) in association with the *Fundo Nacional de Desenvolvimento para a Educação* (National Development Fund for Education - FNDE), responsible for the implementation of public educational policies created by *MEC*.

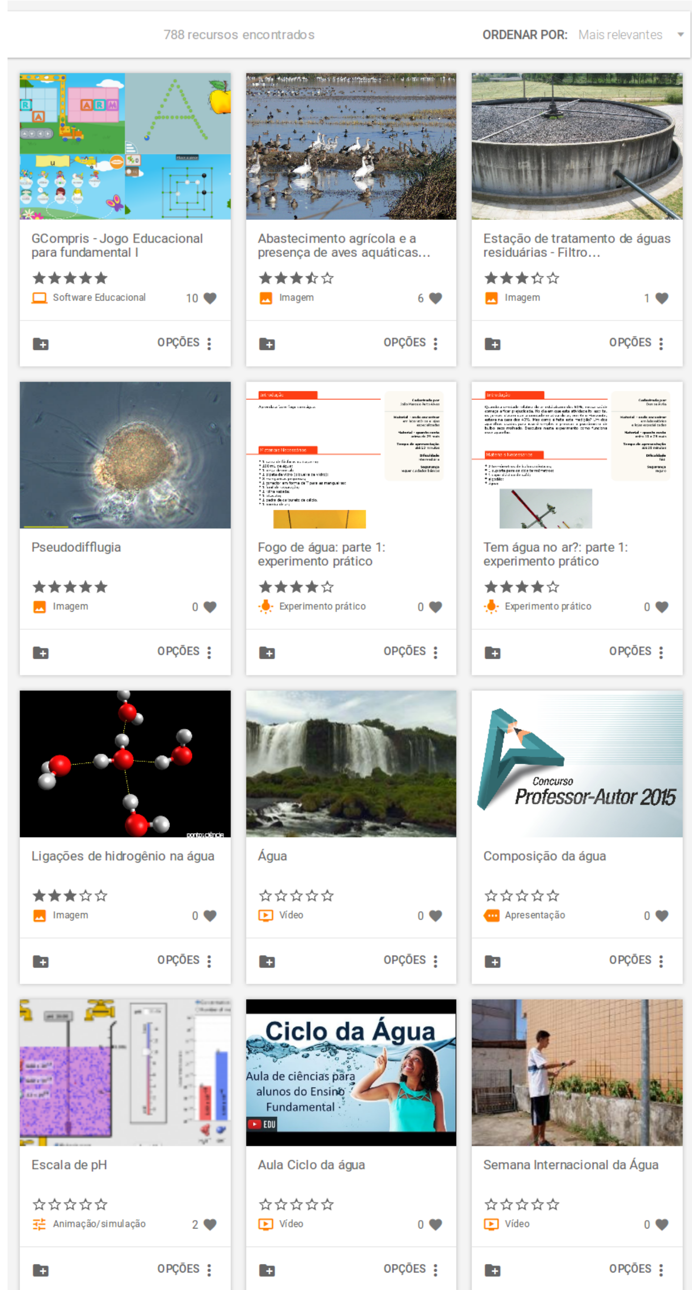

Fig. 6. Results from search using only Elasticsearch score multiplied by the collaborative ranking.

REFERENCES

[1] N. Butcher, *A basic guide to open educational resources (OER)*. Commonwealth of Learning (COL);, 2015.

[2] J. Atenas and L. Havemann, "Quality assurance in the open: an evaluation of oer repositories," *INNOQUAL-International Journal for Innovation and Quality in Learning*, vol. 1, no. 2, pp. 22–34, 2013.

[3] W. L. Pontes, R. M. França, A. P. M. Costa, and B. P., "Filtragens de recomendação de objetos de aprendizagem: uma revisão sistemática

do cbie," *Brazilian Symposium on Computers in Education*, vol. 25, p. 549–558, 2014.

[4] J. Atenas, L. Havemann, and E. Priego, "Opening teaching landscapes: The importance of quality assurance in the delivery of open educational resources," *Open Praxis*, vol. 6, no. 1, pp. 29–43, 2014.

[5] T. Segaran, *Programming Collective Intelligence*. O'Reilly Media, 2008, p. 2.

[6] C. Marlow, M. Naaman, D. Boyd, and M. Davis, "Position paper, tagging, taxonomy, flickr, article, toread," in *In Collaborative Web Tagging Workshop at WWW'06*. Citeseer, 2006.

[7] R. T. Fielding, "Architectural styles and the design of network-based software architectures," Ph.D. dissertation, University of California, Irvine, 2000.

[8] M. Massé, *Rest API Design Rulebook*. O'Reilly Media, 2011, p. 5.

[9] V. Hollister, "What is dspace?" 2011. [Online]. Available: https://wiki.duraspace.org/pages/viewpage.action?pageId=25467341

[10] M. Diggory and B. Luyten, "Storage layer," 2016. [Online]. Available: https://wiki.duraspace.org/display/DSDOC5x/Storage+Layer

[11] O. A. Initiative, "Open archives initiative protocol for metadata harvesting." [Online]. Available: https://www.openarchives.org/pmh/

[12] Elasticsearch, "The heart of the elastic stack." [Online]. Available: https://www.elastic.co/products/elasticsearch

[13] R. Zadeh, "Elasticsearch as a nosql database," 2013. [Online]. Available: https://www.elastic.co/blog/found-elasticsearch-as-nosql

[14] T. P. G. D. Group, "About." [Online]. Available: https://www.postgresql.org/about/

[15] R. Vuorikari and X. Ochoa, "Exploratory analysis of the main characteristics of tags and tagging of educational resources in a multi-lingual context," *Journal of Digital Information*, vol. 10, no. 2, 2009.

[16] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich, *Recommender Systems: An Introduction*. Cambridge University Press, 2010.

[17] S. Chan, T. Stone, K. P. Szeto, and K. H. Chan, "Predictionio: A distributed machine learning server for practical software development," in *Proceedings of the 22nd ACM International Conference on Information & Knowledge Management*, ser. CIKM '13. New York, NY, USA: ACM, 2013, pp. 2493–2496.

[18] T. Yu, O. J. Mengshoel, A. Jude, E. Feller, J. Forgeat, and N. Radia, "Incremental learning for matrix factorization in recommender systems," in *2016 IEEE International Conference on Big Data (Big Data)*, Dec 2016, pp. 1056–1063.

[19] "Building a correlated cross-occurrence (cco) recommenders with the mahout cli." [Online]. Available: http://mahout.apache.org/users/recommender/intro-cooccurrence-spark.html

[20] C. C. Aggarwal, *Recommender Systems - The Textbook*. Springer, 2016, p. 24.

[21] M. A. H. Pimentel, I. B. Sant'Anna, and M. D. D. Fabro, "Searching and ranking educational resources based on terms clustering," in *20^{th} International Conference on Enterprise Information Systems*, Madeira - Portugal, March 2018.