

## Supporting Learning Activities in the Health Sector through Open Data

Marco Alfano, Giovanni Fulantelli, Calogero Mondì, Giuseppe Rinella, Davide Taibi

**Abstract:** *The growing availability of Open Data for educational purposes is playing a key role in shaping and influencing the development of the learning processes. At the same time, Open Data have a particular potential in the health sector so that the major healthcare organizations, universities, colleges, and medical schools are in the process of creating high quality 'Open' education & training resources and tools. However, in this growing world of open health resources for educational purposes, the task of finding the right material can become very difficult mainly when using a general purpose search engine like google. Thus, to overcome this problem and facilitate the search and learning activities in the health domain, we are in the process of building a system that collects open health datasets, applications and projects/initiatives and creates a single entry point to those resources so greatly facilitating the task of a teacher/student in finding what he/she is exactly looking for. In this paper we present the benefit of using Health Open Data for educational purposes together with the basics details of the system we are developing and one practical example of usage.*

**Key words:** *Open Health Data, e-Learning, Open Education, Linked Data in Education, Open data.*

### INTRODUCTION

Open Data is publicly available data that can be universally accessed, used, and redistributed free of charge. It is changing the way governments, private companies, associations and citizens in general use data to understand and elaborate on different areas such as financial, energy, health, education, and more.

In the educational context the “open” approaches have played a key role in shaping and influencing the development of the learning process [5]. The availability of ‘Open Educational Resources’ (OER), before, and more recently the diffusion of ‘Massive Open Online Courses’ (MOOC), delivered by universities all over the world, have increased the opportunity for learning. At the same time, Open Data presents a particular potential in the health sector. By releasing health data to patients and, on an anonymized basis, to researchers and citizens in general, governments and healthcare organizations are using the power of greater openness of data to improve the quality of care, lower healthcare costs, improve research and learning and facilitate patient choices [12]. As a consequence, the recognized value of Open Data for education, on one hand, and the growing availability of Open Data for Health, on the other hand, has lead the major healthcare provider organizations, universities, colleges, and medical schools to create high quality open education & training resources and tools [7].

In this huge amount of open resources that can be used for educational purposes, finding the right material can become a very difficult task, mainly when using general-purpose search engines like Google. For this reason, concerning the educational Open Data domain, an explorer to detect the topic coverage for the datasets in the educational domain has been developed [14] together with an approach to detect the educational relevance of the resources contained in Linked Data datasets [15]. With regards to the educational health domain, some web search methodologies for different types users that classify web pages on the basis of their level of health information and used language have been developed [1], [2].

Combining these different approaches and applying them to Open Health Data, we have built a system that collects open datasets, applications and projects/initiatives in the field of health and creates a single entry point to those resources so greatly facilitating the task of a teacher/student in finding what he/she is exactly looking for.

The paper is organized as follows. The second section describes the Open Data for education in general and the health sector in particular. The third section describes the principles and architectural details of a system for helping a learner to easily find the open

resources he/she needs in the health field. The fourth section describes some practical use of our system and the final section presents some conclusions and future work.

## OPEN DATA FOR EDUCATION IN THE HEALTH SECTOR

The Open Data movement, initially developed in the government sector, has fostered the openness of data in several sectors, including the educational field. In particular, Open Data in education have been introduced in:

- *Datasets*, organizational datasets related to courses, teachers and institutions, as well as educational content;
- *Applications*, datasets are exploited as a data layer of web and mobile applications to support learning activities;
- *Initiatives*, in many cases the Open Data applications for education are developed as a result of specific initiatives such as projects, conferences and challenges.

With respect to open datasets, in the last few years, more and more universities have published datasets, in an open format, containing the course information, the buildings where the courses are delivered, the people belonging to the university structure with different roles (e.g., lecturers, professors and management staff). Moreover, the development of the Linked Data cloud has offered new opportunities to interlink educational resources and the *Web of Data* [4] that contains several datasets providing suitable content in the educational context. The availability of open datasets has fostered the development of web and mobile applications. As an example, a mobile application exploiting the Open Data published by the Open University in UK to provide location-based services to students and academics is presented in [3]. Another example is the MeLOD (Mobile Environment for Learning with Linked Open Data) platform, designed to support, through the use of mobile devices and Open Data, the informal learning experiences that take place during the visit of a city [6].

Among the projects promoting the use of Open Data for educational purposes, it is worthy to mention the Linked-Up project<sup>4</sup>. It is a European funded project carried out with the aim to push the exploitation of the vast amount of public Open Data available on the Web especially by educational institutions and organizations. Among the results of the project, a Linked Education cloud has been developed as a collection of open datasets suitable for educational purposes.

Finally, the use of Open Health Data for learning purposes has brought many advantages to different user typologies, among which:

- Patients, who can better understand their health conditions and make informed choices among the healthcare options available to them;
- Researchers, who have a lot of raw data available and can perform richer and more effective researches;
- Healthcare organizations, which facilitate patient education (and then patient empowerment) while, at the same, improving transparency and customer satisfaction;
- Healthcare companies, that are facilitated in their research activities and can increase their innovation level.

An example in this field is the best practice network m-educator [10], funded by the European commission, that puts together representatives of the above typologies to discuss about innovative approaches to discover, retrieve, share and re-use medical educational content. Among its results, this network has developed a dataset containing learning resources for medical education in open format, and the m-educator3.0 platform<sup>5</sup>, that collects a set of applications exploiting Open Data for education in the health sector.

---

<sup>4</sup> <http://linkedup-project.eu/>

<sup>5</sup> <http://www.meducator3.net/>

### AN OPEN HEALTH DATA PLATFORM FOR E-LEARNING

We have developed an Open Health Data (OHD) platform with the aim of supporting learning activities based on the huge amount of data published in open format. The platform is organized into three main layers:

- *Datasets*, that contains Open Data in the health sector;
- *Applications*, that includes the description of applications exploiting the Open Data as a data layer to support learning activities;
- *Initiatives*, that contains information about the main relevant initiatives (such as projects, workshops, conferences and challenges) related to Open Data for health education.

This three-layer structure presents an increasingly granularity and reflects the main findings on the use of Open Health Data for education as stated above. In fact, datasets are at the basis of the pyramid and they provide the data layer for the development of applications, that, in turn, are in most cases developed as a result of projects or challenges promoting the use of Open Data in education. The authors have created the metadata structure for each layer with the aim of collecting those information and enable better searching and filtering. The metadata structure for each layer is reported in Table 1.

Table 2. Metadata for the three layers

<b>Health Datasets</b>	<b>Health Applications</b>	<b>Health Initiatives</b>
<b>Name; Description</b> <b>Institution ;Country; URL</b> <b>Format (CSV, Excel, Json, ...)</b> <b>License (e.g. CC, CC0, ODL, ...)</b> <b>Last update; Language</b> <b>Number of downloads</b> <b>Tag; Keywords; Location</b> <b>Geographical coverage</b>	<b>Name</b> <b>Description</b> <b>Type</b> <b>Coordinator</b> <b>Language</b> <b>URL</b> <b>Tag</b> <b>Keywords</b>	<b>Name</b> <b>Type (workshop/conference, project, etc.)</b> <b>Description</b> <b>URL</b> <b>Organizer</b> <b>Country</b> <b>Date</b> <b>Key words</b>

Fig. 1.a shows the mock-up the OHD platform structured in three sections: Dataset, Applications and Initiatives. A user enters the sections of the website and can filter the datasets, initiatives and applications according to his/her preferences. In particular, Fig. 1b shows the section related to datasets where a user can choose some filters to refine his/her search as, for example, country, language, format, license, and institution. Moreover, a user can choose to search by tags and to order datasets by relevance, titles, date, number of visits, etc.



Figure 1.a. OHD homepage

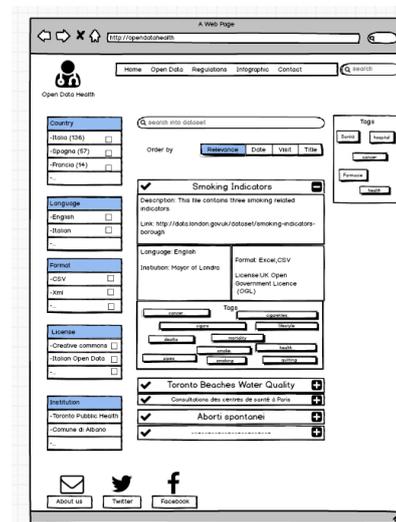


Figure 1.b. Datasets Layer

At present, the platform contains around 160 datasets, with the majority of them in CSV format (around 48%). In order to categorize the datasets, both tags declared by the owner of the datasets and categories identified by the authors have been adopted. The applications section contains about 30 “good practices” grouped in web and mobile applications (most of them have been collected from the Government of Canada<sup>6</sup> website).

### A USE CASE SCENARIO OF USING THE OHD PLATFORM FOR LEARNING

Learning approaches based on ‘Inquiry Based Learning’ (IBL) as well as ‘Simulation Learning Environment’, can benefit from the availability of our OHD platform. In fact, in an IBL scenario, students can access the huge amount of datasets to get real data about the topic that is under investigation. When it comes to simulation based learning environment, the datasets can offer a knowledge base composed by real data, to support simulations efficiently.

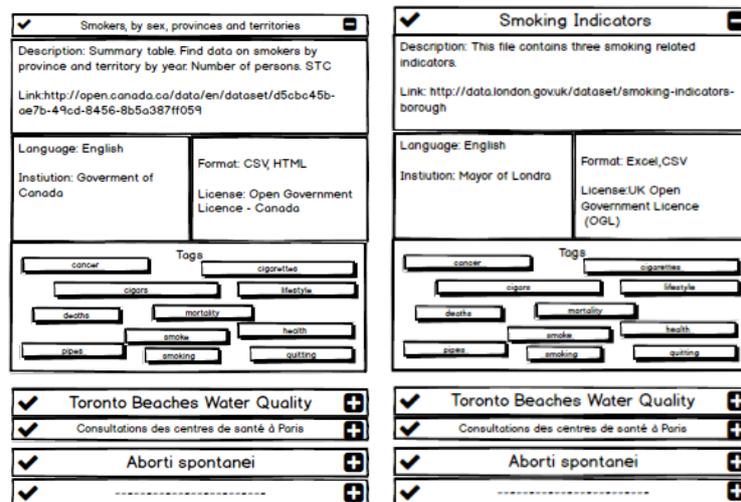


Figure 2. Search results

As a practical example of use of our OHD platform we can assume that Margaret, a 4th year student of high school, must do a research in the health sector. In particular, the student focuses on the percentage of smokers in different countries. The student uses our platform, clicks on the Datasets section and chooses to make a search using the “cancer”, “cigarettes”, “deaths”, “lifestyle”, “mortality”, and “smoke” tags.

The platform provides her with a list of datasets and she chooses those of Canada<sup>7</sup> and London<sup>8</sup> (the latter dataset contains smoking information on the whole United Kingdom), as shown in Fig. 2. Margaret accesses the two institutional portals where the datasets are stored, and downloads the datasets she is looking for. Finally, she performs an empirical analysis by comparing the data of the two datasets and, in particular, the smoker’s rates in order to get the trends in both countries. Finally, she combines the results and gets a chart as the one presented in Fig. 3. This example shows how our platform simplifies and accelerates the student work.

6 <http://open.canada.ca/en/apps>

7 <http://open.canada.ca/data/en/dataset/36c8bbcf-2b10-427f-82fb-4e0462e81051>

8 <http://data.london.gov.uk/dataset/smoking-indicators-borough>

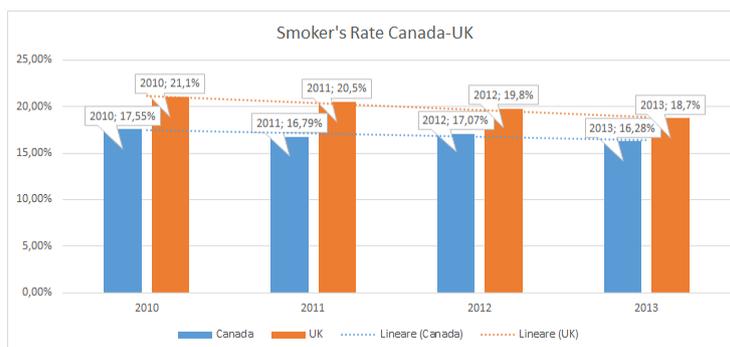


Figure 3. Smoker's rate in Canada and UK.

## CONCLUSIONS AND FUTURE WORK

This paper has presented the benefit of using Health Open Data for educational purposes together with the basics details of a system that collects open datasets, applications and projects/initiatives in the field of health and creates a single entry point to those resources so greatly facilitating the task of a teacher/student in finding what he/she is exactly looking for. We are in the process of implementing such a system and to collect more datasets, applications and initiatives/projects.

Further learning scenarios, exploiting at full the potentiality of our OHD platform, will be investigated. In particular, learning activities based on simulation would take advantages from the availability of data in an open format coming from real contexts. For this reason, we plan to integrate the OHD platform with simulation environments to support learning activities.

## REFERENCES

- [1] Alfano, M., Lenzitti, B., and Lo Bosco, G., "A web search methodology for health consumers", Proc. of ACM International Conference on Computer Systems and Technologies (CompSysTech'14), 27-28 June 2014, Ruse, pp. 150-157, 2014.
- [2] Alfano, M., Lenzitti, B., and Lo Bosco, G., Fiding the best web medical content for different learner categories. Procs. of 2nd International Conference on Medical Education Informatics (MEI 2015), 2015.
- [3] d'Aquin, M., Zablith, F. and Motta, E. (2011). wayOU - linked data-based social location tracking in a large, distributed organisation. In: 8th Extended Semantic Web Conference (ESWC 2011), 29 May - 2 June 2011, Heraklion, Greece.
- [4] Dietze, S., Yu, H., Giordano, D., Kaldoudi, E., Dovrolis, N., Taibi, D. 2012. Linked education: interlinking educational resources and the web of data. In: Proceedings of the 27th ACM Symposium on Applied Computing, Special Track on Semantic Web and Applications. SAC'12, ACM, New York, NY, USA, 2012.
- [5] Fulantelli, G., Gentile, M., Taibi, D., Allegra, M. 2008. The Open Learning Object model to promote Open Educational Resources. JIME.
- [6] Fulantelli G., Taibi D., and Arrigo M. 2013. A semantic approach to mobile learning analytics. In Proc. of the First International Conference on Technological Ecosystem for Enhancing Multiculturality (TEEM '13). ACM, New York, USA, 287-292.
- [7] Groen, P. 'Open' Education & Training Resources for Healthcare. <http://www.openhealthnews.com/blogs/groenpj/2012-09-09/open-education-training-resources-healthcare>. 2012.
- [8] Journal of open health data, <http://openhealthdata.metajnl.com/>
- [9] Linked Open Data for Learning Object Discovery: Adaptive e-Learning Systems, <http://www.computer.org/csdl/proceedings/incos/2011/4579/00/4579a060-abs.html>

[10] Mitsopoulou, E., Taibi, D., Giordano, D., Dietze, S., Yu, H. Q., Bamidis, P., Bratsas, C. and Woodham, L. 2011. Connecting Medical Educational Resources to the Linked Data Cloud: the mEducator RDF Schema, Store and API, in *Linked Learning 2011*, CEUR-WS, Vol. 717, 2011.

[11] Noveck B. S., "Open Data: The Democratic Imperative," *Crooked Timber*, July 7, 2012, accessed February 21, 2014, <http://crookedtimber.org/2012/07/05/open-data-the-democratic-imperative/>.

[12] The Open Data era in health and social care,  
<http://images.thegovlab.org/wordpress/wp-content/uploads/2014/06/nhs-full-report.pdf>

[13] Yu, H. Q., Dietze, S., Li, N., Pedrinaci, C., Taibi, D., Dovrolis, N., Stefanut, T., Kaldoudi, E., Domingue, J., A Linked Data-driven & Service-oriented Architecture for Sharing Educational Resources, in *Proc. of Linked Learning 2011: the 1st International Workshop on eLearning Approaches for the Linked Data Age*, CEUR-Vol 717, 2011.

[14] Taibi D., Dietze S., Fetahu, B., Fulantelli G. Exploring type-specific topic profiles of datasets: a demo for educational linked data. In *Poster and system demonstration proceedings of the International Semantic Web Conference, ISWC, 2014*

[15] Taibi, D., Fulantelli, G., Dietze, S., Fetahu, B. Evaluating relevance of educational resources of Social & Semantic Web, 8th European Conference on Technology-enhanced Learning (ECTEL'13), Paphos, Cyprus, 17-21 September, 2013.

#### **ABOUT THE AUTHORS**

Marco Alfano, PhD, Anghelos Centre on Communication Studies and Dipartimento di Matematica e Informatica, University of Palermo, Palermo, Italy, Phone: +39 091341791, E-mail: [marco.alfano@anghelos.org](mailto:marco.alfano@anghelos.org).

Dr. Giovanni Fulantelli, Institute for Educational Technologies, National Research Council of Italy, Phone: +39 091 6809 220, E-mail: [giovanni.fulantelli@itd.cnr.it](mailto:giovanni.fulantelli@itd.cnr.it).

Dr. Calogero Mondì, Institute for Educational Technologies, National Research Council of Italy, E-mail: [carlo.mondi84@gmail.com](mailto:carlo.mondi84@gmail.com)

Dr. Giuseppe Rinella, Institute for Educational Technologies, National Research Council of Italy, E-mail: [giuseppe.rinella@libero.it](mailto:giuseppe.rinella@libero.it)

Dr. Davide Taibi, Institute for Educational Technologies, National Research Council of Italy, Phone: +39 091 6809 216, E-mail: [davide.taibi@itd.cnr.it](mailto:davide.taibi@itd.cnr.it)

**The paper has been reviewed.**