

Open Government Data in Higher Education: A Multidisciplinary Innovation Teaching Experience

Iván Cantador¹[0000-0001-6663-4231], J. Ignacio Criado²[0000-0002-9184-9696],
Laura Alcaide Muñoz³[0000-0003-3885-0660],
María E. Cortés-Cediel⁴[0000-0001-8100-1336], and
Irene Liarte²[0000-0002-7885-4896]

- ¹ Escuela Politécnica Superior, Universidad Autónoma de Madrid, Campus de Cantoblanco, 28049 Madrid, Spain ivan.cantador@uam.es
- ² Facultad de Derecho, Universidad Autónoma de Madrid, Campus de Cantoblanco, 28049 Madrid, Spain {ignacio.criado,irene.liarte}@uam.es
- ³ Facultad de Ciencias Económicas y Empresariales, Universidad de Granada, Campus de Cartuja, 18071 Granada, Spain lauraam@ugr.es
- ⁴ Facultad de Ciencias Políticas y Sociología, Universidad Complutense de Madrid, Campus de Somosaguas, 28223 Pozuelo de Alarcón, Spain mcorte04@ucm.es

Abstract. In this paper, we present an innovative teaching project aimed to experiment with the use of open government data in higher education. Differently to previous studies on the topic, our project follows a multidisciplinary approach by conducting and evaluating several learning activities in degree subjects from distinct fields —political science, economics and finance, and computer engineering—, and targeting goals related to the development of professional and technological skills for understanding, accessing, analyzing and exploiting open data. Together with descriptions of these activities, which can be of inspiration for teachers having distinct backgrounds, we report results and valuable findings from a questionnaire-based study with students, providing insights about how to support effective teaching and learning using open data.

Keywords: · higher education · digital competences · interdisciplinary teaching · learning innovation · open data · open government.

1 Introduction

Nowadays, in the Big Data era, there is the need of data literacy, which entails learning how to read, process, analyze and argue with data [8]. It is related to the ability of searching and extracting knowledge from raw data [4] and, as stated by the OECD¹, it is a crucial skill that citizens have to acquire by 2030.

In the educational context, students have to be guided in going beyond the passive inspection of results returned by a search engine, and in actively querying

¹ <http://www.oecd.org/education/2030-project/teaching-and-learning/learning>

The screenshot shows the EU Open Data Portal interface. At the top, there is a navigation bar with 'Data', 'Studies', 'data.europa academy', 'News', and 'Contact'. The main header features the 'data.europa.eu' logo and a language selector set to 'English (en)'. Below the header, the 'Datasets' section is active, displaying a search bar and a filter by location map. The search results show 8,365 datasets found, categorized by 'datos.gob.es' and 'Education, culture and sport'. Two specific datasets are highlighted: 'Public Schools of Madrid' and 'Sports. Basic Municipal Sports Facilities'. Each dataset entry includes a description, update/creation dates, and available data formats (e.g., RDF/XML, JSON, CSV, RSS feed, HTML, XSL).

Fig. 1. Screenshot of the EU Open Data Portal, the official point of access to public data published by the EU institutions, agencies and other bodies. Its data collections are accessible via category menus, keyword-based search fields, and filtering criteria forms, among other mechanisms. Each collection can be provided in multiple electronic formats (e.g., CSV, XLS, XML, HTML, RDF, and JSON) and has associated metadata, such as a title, a description, and several tags.

and looking for the data that best satisfy their information needs and allow informed decision-making [10]. Researchers have indeed identified a number of factors that should be taken into consideration in schools to develop suitable knowledge management, such as acquisition, learning, dissemination and transfer methods and technological solutions [2, 14].

Among the existing Big Data sources, in this paper we focus on Open Government Data (OGD). As defined by the OECD,² OGD represents an initiative aimed to promote transparency, accountability and public value by making government data available to all. Facilitating the access and encouraging the use and free distribution of their datasets, governments foster the creation of business and the development of innovative, citizen-centric services [5].

OGD policies and portals of public administrations are thus one of the most recent trends of modernization and innovation in the public sector at international level. For instance, the EU Open Data portal (figure 1) provides access to more than 1.4 million datasets from 36 countries. In Spain, the open data portal

² <https://www.oecd.org/gov/digital-government/open-government-data.htm>

of the General State Administration³, controlled by the central government, contains over 57,000 datasets on a large number of domains, such as demography, employment, health, education, economy, culture, and natural environment, to name a few.

The openness, access, retrieval, processing and analysis of data are aspects of great interest and potential, particularly in the context of higher education [12]. Using open data in education entails a series of benefits. It allows designing practical and professional oriented courses. Serving as educational materials, it enables the development of student tasks aimed to address real-life problems [7]. This, as widely recognized in the Problem-based Learning literature, promotes the students' satisfaction and engagement during their learning process [6]. Additionally, open data have the potential to improve students' digital and data skills that are essential for future generations [15].

Despite these benefits, open data come up with a number of challenges [3]. For instance, there is an extended unfamiliarity with open data as a potential educational resource [12]. Instructors are not aware of the concept of open data and how they could integrate open data into their teaching. Even having knowledge on open data, they may have significant difficulties in finding relevant datasets [15]. Moreover, there is a generalized lack of skills; learners and educators may not have the literacies (digital and data skills) and resources (including time) to make open data useful for them [3].

The academic literature on education comprises reports of teaching experiences using open data in individual courses, e.g., databases [9], computer programming [7], and statistics [13], as well as interviews with instructors about the use of OGD in education [12]. These studies report isolated experiences, suggesting the need for further elaboration of comparative studies.

Differently to previous work, in this paper we present a multidisciplinary innovation teaching project aimed to implement and evaluate several OGD-driven learning activities for degree subjects in different fields, namely political science, economics and finance, and computer engineering. The activities do not only differ in the field of knowledge, but also in the way open data is used. Hence, we propose activities with different goals: understanding, accessing, analyzing and exploiting OGD. Moreover, the project includes an empirical study where students provided opinions and suggestions about the utility of OGD in learning, education and other contexts. For all the above, we believe our this work could be a valuable reference for educators interested in teaching with open data technologies, regardless of their background.

The remainder of the paper is organized as follows. Section 2 surveys related work on using OGD in higher education. Section 3 presents the multidisciplinary activities of our project. Next, Section 4 reports empirical results from the evaluation of such activities, and finally, Section 5 ends with some conclusions and future research lines.

³ <https://datos.gob.es>

2 Related Work

In the literature related to OGD and higher education, we can distinguish between two main lines of work: the *generation of educational resources as open data* [16] and the *use of open (government) data as educational resources in education* [1].

With respect to the use of OGD as educational resources, and focusing on **data exploitation goals**, Mazón et al. [9] performed a learning activity in a Databases course of a Computer Engineering degree where students had to propose an original scenario where different open data should be reused for a specific goal. Following a project-based learning methodology, the students had to design a relational database for managing the data in the envisioned scenario. The activity promoted a creative and entrepreneur attitude in students, as well as encouraged autonomous and lifelong learning. Besides, surveys made to students showed that reusing open data in a cooperative work context increased the students' motivation.

Also considering the exploitation of open data in the Computer Science field, Maksimenkova and Podbelskiy [7] conducted an activity in a computer programming course for 1st-year undergraduates in Software Engineering. The purpose of the activity was the implementation of computer programs able to read, process and visualize open data files. As stated by the authors, the utilization of open data was beneficial since it helped in making programming education "less artificial." However, the variety of format, complexity and quality of the open data collections was an issue that made the design and final success of the activity difficult.

Focusing on **data analysis goals**, Renuka et al. [11] designed a learning activity where Engineering students had to obtain and analyze open data to perform predefined studies related to the time evolution of urbanization factors, traffic intensity, and air quality levels in a city. The authors presented the results achieved by students, but did not report an evaluation of the activity. With similar learning goals, Rivera, Marazzi and Torres-Saavedra [13] conducted an activity for an introductory course in statistics where students had to analyze OGD collections to address real case studies, such as providing a metric of violence, analyzing demographics of hotel registrations, and studying variables related to types of properties in a country. As one of the main conclusions obtained from the activity, the authors claimed that the ubiquity of open data has the advantage of allowing teachers to easily 'locally adapt' lesson plans, integrating real data with context and purpose. They, by contrast, highlighted the lack of reliability in cases where it was not clear where the data came from.

From the literature survey, we can conclude that using open data has a number of benefits, such as increasing the students' engagement for addressing realistic problems, but entails additional efforts to instructors who have to carefully explore the data (format, complexity, quality, reliability) before using it in their classes. We observe that reported activities were done isolatedly in single subjects, and did not deal with the understanding of what open data are and how open data can be accessed. Our project, by contrast, follows a multidisciplinary

approach considering several subjects and knowledge fields, multiple types of student tasks (e.g., individual oral presentations, collaborative and problem-based assignments, final degree works) and levels of OGD use: understanding, access, analysis and exploitation.

3 Multidisciplinary Activities

Our teaching innovation project aimed at promoting a culture of use of open data among instructors and students, developing learning materials related to open data, presenting the use of open government data portals as resources for analyzing the political-administrative, economic and legal realities, and developing professional and technological skills related to the access, analysis and exploitation of open data from public administrations.

In the context of higher education and through a multidisciplinary perspective—including political and economic sciences, as well as engineering studies—we conducted several open data learning activities of different nature (i.e., writing reports, oral presentations, practical assignments, and final degree works) and methodologies (i.e., teacher-centered, student-centered, and project-based). To provide consistency, the activities were restricted to three OGD portals in Spain at national, regional and municipal level, from the governments of Spain⁴, Andalusia⁵ and Madrid⁶, respectively.

We next briefly describe the activities, grouped by their usage focus of OGD, namely understanding, access, analysis, and exploitation.

3.1 Open Government Data Understanding

A first type of activities is focused on understanding what OGD are: their goals, benefits and challenges, formats and forms of publication, applications, case studies, etc. In this sense, students could for example be requested to search and survey references, and make written reports and oral presentations, explaining what they learned on certain issues and topics.

Regarding our project, in a 2nd-3rd-year subject on Organizational Theory and Public Administration of the Law and Political Science and Public Administration Double-Degree, at the Faculty of Law of Universidad Autónoma de Madrid (UAM), a group of students participated in an immersive experience oriented to promote their understanding of the political side of OGD portals, beyond the legal dimension. After some activities led by the instructors throughout a workshop, videos, and direct content observation, the students were encouraged to strengthen their understanding of OGD portals, both connecting vision (transparency) and voice (participation) dimensions of openness in the public sector.

⁴ <https://datos.gob.es>

⁵ <https://www.juntadeandalucia.es/datosabiertos>

⁶ <https://datos.madrid.es>

Besides, for a Databases subject, a group of 3rd-year computer engineering students at the High Polytechnical School of UAM had to investigate advanced issues related to linked open data repositories (e.g., data formats and structures, storage systems, querying languages, linkage and publication schemas), focusing on the government domain. Students had to make individual seminars in the classroom, raising debates around the studied issues.

3.2 Open Government Data Access

A second type of learning activities is associated with the access to OGD, since both searching data collections and exploring the content of data files (in different formats) are tasks that require specific skills. In this context, students could be requested to find relevant data sources and collections, and extract from one or more data files particular answers or information for a given question or issue.

Within our project, in a 1st-year subject on Theory of Public Administration of the Management and Public Administration degree at the Faculty of Political Science and Sociology of Universidad Complutense de Madrid (UCM), a group of students participated in a number of workshops where they had to find, collect and present institutional information from open data portals in order to evaluate the levels of governance, transparency and accountability achieved by representative governments.

Moreover, two 4th-year computer engineering students at UAM did their final degree work on open government access problems. Each work entailed a brief survey of literature, design, implementation and evaluation of an application software, writing a report, and making an oral presentation for an academic committee. Specifically, the works consisted of a couple of intelligent dialog agents (chatbots) to assist on the access to OGD and e-participatory budgeting content through formal queries built via natural language conversations.

3.3 Open Government Data Analysis

A third type of learning activities comprises any kind of task involved in the analysis of OGD, namely the processing, integration and filtering of data, the use of statistical metrics and methods on selected data, and the application of visualization techniques and tools on data and analysis results. Students thus could be requested to perform a wide array of assignments aimed to analyze real-world phenomena.

In particular, our project encompassed a classroom assignment where the 3rd year accounting and finance students at the Faculty of Business Studies at Universidad de Granada (UGR) had to analyze main aspects of OGD platforms: data catalog, accessibility and visualization, and citizen participation. To do this, the teacher offered the students a statement explaining the above aspects, and allowed them to choose two of the three aspects. Students had to understand differences between analyzed platforms, explore real cases and that they could provide critical opinion on the state of this type of initiatives, compared to the disclosure of information in the private sector (more analyzed by them).

Also with data analysis purposes, a final degree work by a 4th-year computer engineering student at UAM aimed to develop a computer program to measure and analyze controversy on citizen debates (publicly available as open data) from an e-participation platform of Madrid City Council.

3.4 Open Government Data Exploitation

A fourth and final type of learning activities concerns the exploitation of (processed) OGD for a target purpose or application. Students could thus be requested to perform (complex) solutions to address real-world decision-making problems.

Within this category of activities, our project included a cooperative project-based activity in the Databases subject of the 3rd year of the Computer Engineering degree at UAM, where students had to design and build from scratch a relational database with OGD collected from the web. The application domain of the database built by each student team was freely chosen.

The project also comprised a final degree work at UAM, where a 4th-year computer engineering student implemented and evaluated a number of recommender systems that exploited OGD to suggest Madrid residents with citizen proposals that may be of their interest, according to both topic and location information. Hence, residents could be informed about city problems and proposed solutions. Moreover, relevant proposals may be supported by residents in the participatory budgeting processes of the city.

4 Empirical Results

Each of the classroom activities presented in Section 4 was conducted by a group of students. In all cases, after having done an activity, students were requested to voluntarily participate in our study by filling out an online questionnaire. We received feedback from 182 students (60% female, 40% male) with ages ranging from 18 to 26 years old (only 5% older than 23 years old). The questionnaire was composed of 5-point Likert scale questions aimed to evaluate the students' satisfaction with the activities and opinion about OGD. The questionnaire also had some questions for open comments and suggestions.

Figure 2 shows the distributions of scores given by the students on their perception about the utility of open data for doing academic work, complementing their study, enhancing their learning, and pursuing other purposes. Several interesting insights can be derived from them by considering the students' academic fields: computer engineering, economics and finance, and political science.

As one would expect, political science students are those that found OGD collections most useful for studying and doing academic work in their degree subjects. In this sense, some political science, and economics and finance students commented on the value of OGD, motivated by the possibility of performing data-driven sociocultural, political and economic analysis. Computer engineering students, by contrast, expressed a moderate opinion about the potential utility

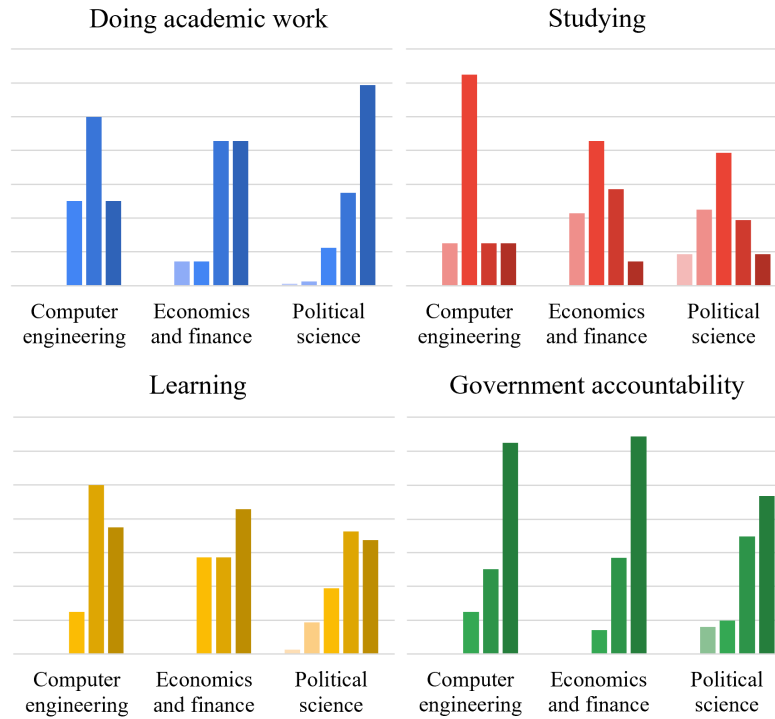


Fig. 2. Distributions of 5-point Likert scores (1 = very unuseful, 5 = very useful) given by the students about the perceived utility of using OGD for different purposes. The y-axis ranges from 0% to 70%, in intervals of 10%.

of open data for their studies. This may be due to the fact that their curriculum planning puts much more emphasis on learning algorithms and computer programming than data analysis and exploitation.

Nonetheless, interestingly, computer engineering students perceived open government as very useful for enhancing their learning. With this respect, some students mentioned the possibility of exploiting open data to build real-world databases and software applications, and to test machine learning methods on realistic prediction and classification tasks.

In general, regardless of their academic fields, students tended to give the highest score to the possibility of exploiting OGD for checking government accountability. Political science students were the most critical in this regard.

Finally, through open comments, students suggested the use of OGD for other (non-academic) purposes, such as verifying information published in news and social media, and ensuring government transparency. Here, one of the most interesting ideas that our activities raised was the opportunity of open data to tackle fake news and promote fact-checking of information, mostly in social media and algorithmic environments.

As reported through the questionnaire, 85% of the students were not aware of OGD portals before the activities. However, most of the students were satisfied or very satisfied with the activities (81%) and considered them as useful or very useful for their progress within the degree subjects (65%). This is another important conclusion of our study, as OGD portals comprise an asset of public sector organizations to engage with citizens, and promote trust and confidence in public institutions, including compulsory and higher education studies.

5 Conclusions

The academic literature recognizes the potential benefits of open (government) data as learning resources, but it also compiles a number of challenges that should be faced for their effective use; among them, the lack of teachers' awareness of the concept of open data, and how they can integrate open data in their subjects and classes.

Complementing published reports of individual experiences in particular higher education courses and surveys done with instructors, in this paper we have presented an innovative teaching project where distinct learning activities were done in subjects of different academic fields —political science, economics and finance, and computer engineering— at various levels of studying open data: understanding, access, analysis and exploitation.

Together with the categorized examples of activities that can be of inspiration for teachers having distinct backgrounds, we also have reported valuable findings obtained from a questionnaire-based multidisciplinary evaluation with students. The conducted experience showed the need for making changes on curriculum planning to develop or reinforce competencies on data processing and analysis, and for giving motivations to use open government data in daily life tasks, such as verifying the veracity of information published in news and social media, and making research works.

We plan to explore alternatives to the teacher-centered, student-centered and project-based learning methodologies followed in our project. Specifically, we want to perform new activities applying cooperative, flipped classroom, and challenge-based learning methodologies, as well as gamification mechanics (see examples given by Saddiqa et al. [15]). The activities could also be done in postgraduate studies, providing MSc and PhD students with opportunities to use open government data in their researches.

Our work may be of interest for scholars in different areas of knowledge and practitioners in the public sector. In the first case, colleagues in higher education institutions might find our cases and experiences inspirational for their own learning environments, including different areas of humanities, social and natural sciences, or engineering. In the second case, practitioners from public organizations may expand their approaches to OGD reutilization. In this context, public managers should include educational purposes in their portfolio of potential uses of their OGD portals, opening up a new area of interest among instructors and students at all levels.

Acknowledgements This work was supported by Universidad Autónoma de Madrid (D.008.21.INN) and the Spanish Ministry of Science and Innovation (PID2019-108965GB-I00).

References

1. Atenas, J., Havemann, L., Priego, E.: Open data as open educational resources: Towards transversal skills and global citizenship. *Open Praxis* **7**(4), 377–389 (2015)
2. Chu, K.W., Wang, M., Yuen, A.H.: Implementing knowledge management in school environment: Teachers' perception. *Knowledge Management & E-Learning* **3**(2), 139–152 (2011)
3. Coughlan, T.: The use of open data as a material for learning. *Educational Technology Research and Development* **68**(1), 383–411 (2020)
4. Donato, R.D., Garofalo, M., Malandrino, D., Pellegrino, M.A., Petta, A.: Education meets knowledge graphs for the knowledge management. In: 10th International Conference in Methodologies and Intelligent Systems for Technology Enhanced Learning. pp. 272–280. Springer (2020)
5. Hardy, K., Maurushat, A.: Opening up government data for Big Data analysis and public benefit. *Computer Law & Security Review* **33**(1), 30–37 (2017)
6. Hung, W., Jonassen, D.H., Liu, R., et al.: Problem-based learning. *Handbook of Research on Educational Communications and Technology* **3**(1), 485–506 (2008)
7. Maksimenkova, O., Podbelskiy, V.: On practice of using open data in construction of training and assessment tasks for programming courses. In: 10th International Conference on Computer Science & Education. pp. 233–236. IEEE (2015)
8. Mandinach, E.B., Gummer, E.S.: A systemic view of implementing data literacy in educator preparation. *Educational Researcher* **42**(1), 30–37 (2013)
9. Mazón, J.N., Lloret, E., Gómez, E., Aguilar, A., Mingot, I., Pérez, E., Quereda, L.: Reusing open data for learning database design. In: 2014 International Symposium on Computers in Education. pp. 59–64. IEEE (2014)
10. Petrides, L.A., Nodine, T.R.: Knowledge management in education: Defining the landscape (2003)
11. Renuka, T., Chitra, C., Pranasha, T.S., Dhanya, G., Shivkumar, M.: Open data usage by undergraduate students. In: 5th IEEE International Conference on MOOCs, Innovation and Technology in Education. pp. 46–51. IEEE (2017)
12. Rivas-Rebaque, B., Gértrudix-Barrio, F., de Cisneros de Britto, J.C.: La percepción del docente universitario ante el uso y valor de los datos abiertos. *Educación XX1* **22**(2), 141–163 (2019)
13. Rivera, R., Marazzi, M., Torres-Saavedra, P.A.: Incorporating open data into introductory courses in statistics. *Journal of Statistics Education* **27**(3), 198–207 (2019)
14. Rodrigues, L.L., Pai, R.: Preparation and validation of KM measurement instrument: An empirical study in educational and IT sectors. In: Knowledge Management: Nurturing Culture, Innovation, and Technology, pp. 583–593. World Scientific (2005)
15. Saddiqa, M., Magnussen, R., Larsen, B., Pedersen, J.M.: Open Data Interface (ODI) for secondary school education. *Computers & Education* **174**, 104294 (2021)
16. Zablith, F., Fernandez, M., Rowe, M.: The OU linked open data: Production and consumption. In: 8th Extended Semantic Web Conference. pp. 35–49. Springer (2011)