Introducing Open Data Concepts to STEM Students Using Real-World Open Datasets

I. Bosnić[*], I. Čavrak[*] and A. Zuiderwijk[**]
[*] University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
[**] Delft University of Technology, Delft, The Netherlands
{ivana.bosnic, igor.cavrak}@fer.hr, A.M.G.Zuiderwijk-vanEijk@tudelft.nl

***Abstract* – While open data concepts become more important in our society, education about its benefits and technical issues is still behind the practice. Students of STEM disciplines should be introduced to open data during their education. The *Open Computing* course, completely redesigned in the new Computing curriculum, introduces open data concepts, providing both the basics and advanced topics, from technical to social and legal viewpoints. Among the several educational activities, one was particularly useful for understanding the needs and implications of using open data: a synchronous group activity where students had to choose a societal issue, find and analyze two open datasets that would help gaining insight into this issue, assess interdisciplinarity approaches and stakeholders, and finally propose the added value emerging from the solution. In a short amount of time needed, this activity – which tackled multiple aspects of the problem - brought a clearer insight into the topic, building upon the conventional lectures. Students highly graded such an approach to their education, where they had to construct their knowledge by the group experience. A similar group activity appeared to be useful in the context of open data PhD training and might also be used in other disciplines and domains.**

***Keywords* – open data; education; datasets; interdisciplinary approach**

I. Introduction

Today's education provides a plethora of courses and topics to empower students with skills and knowledge needed for their future real-world endeavors. One area becoming more and more important in the society – but not yet in education – are various concepts of *openness* such as open systems, open formats, open standards, open source, open science, open education, open licensing, open hardware, and one of the currently most prevalent ones: open data. These concepts can - and should – be covered in various knowledge areas and fields, but especially in STEM education. The International Open Data Charter (Principle 6, §5), adopted by 79 national and local governments worldwide since 2015, includes a principle to "engage with schools and post-secondary education institutions to support increased open data research and to incorporate data literacy into educational curricula" [1]. However, education on open data is still behind the practice. This paper aims to be a step in that direction.

The paper structure is as follows: Section 2 presents the related work on integrating open data and education. Sections 3 and 4 give the overview and contents of the course. Section 5 presents a similar setup that was the basis for our real-world open data group assignment, described with its results in Section 6. Section 7 gives the results of students' evaluation which is additionally in Section 8.

II. Related Work

A distinction can be made among three intersections between open data and the education sector [2]. First, national and international open data regarding education at primary, secondary, vocational and higher education stages can be used by policymakers to understand developments in education better. Second, parents and students can utilize local data concerning quality levels of schools and other educational forms. Third, educators can use open data for teaching and learning activities, for example, to develop knowledge on statistics, media, science and politics. In this study, we focus on the third type of intersection between open data and education [2].

A study described in [3] collected 26 answers from academics on usage of open data in university settings, out of which only 11 answers described clear open data activities. In a newer study [4], 10 interviews with educators using open data in teaching were analyzed, with 5 courses being done at the college/university level. Detailed interviews described educational activities and ways of interacting with open data, especially in the societal context.

Although the potential integration of real-world open datasets in teaching and learning activities is not yet well understood [4], a number of studies have started to explore this topic. For example, studies on how open data can be used as an educational resource to assist educational processes at primary schools in Denmark [5],[6]. Three essential benefits of using open data in primary schools were identified: 1) enabling students to comprehend information about their local environment, 2) increasing pupils' awareness and development of digital and data skills, and 3) empowering students to come up with ideas to improve their residential areas. An additional study finds that open data can be used in school teaching to support the development of digital and data literacy skills [7]. Open data can also be integrated into teaching for university students; this was useful for students to scrutinize and comprehend the barriers and potential value of open data [8]. In addition to the learning effects, the involved students contributed to scientific research on the usability of particular open data platforms, which provided useful insights for open data platform developers and policymakers.

Besides the potential benefits of using open data as a resource for teaching and learning activities, various challenges exist [4]. Critical issues include difficulties with finding data available in the most useful format to the educator [4]. Educators also need to conduct the non-trivial task of refining and manipulating data to fit their teaching purposes [4], [9]. In case that 'live' open data is being used for educational purposes, there is a strong dependency on the data, which is risky [10]. Open datasets may be changed or updated by the data provider, affecting the validity, correctness and usability of the data for learning purposes [10]. A significant challenge is to educate the educators and increase their awareness of formats and documentation [4]. Another study also points at a lack of awareness concerning existing open datasets amongst educators [7].

III. Open Computing Course Overview

For about two decades, the *Open Computing* course is run at the University of Zagreb, Faculty of Electrical Engineering (FER). This elective course aims mainly at Computing students, studying in subfields such as Software Engineering, Computer Science or Computer Engineering. Students of all these areas could benefit from being introduced to the open concepts in several contexts. The course can be taken in the 3rd year of the undergraduate program or the 1st year of a graduate study program.

Running such a course for around twenty years involved several changes in the content as technology and society vastly changed. The course underwent through major changes in the academic year 2020/2021, coinciding with the introduction of the new study program FER-3. These changes included promoting the theory and practice of open data, being the main focus of this paper. In regards to this paper, authors from FER are course responsibles, while external author was involved in creating the open data assignment.

The course is run for 13 weeks and is awarded 5 ECTS points. Due to academic calendar constraints, it is divided into two main parts lasting 7 and 6 weeks, with the mid-term exam in-between. The number of students enrolled in the course per year is 25-50. The course consists of *(i)* lectures, *(ii)* weekly student activities and *(iii)* laboratory exercises. Due to the COVID-19 pandemic, all the activities were moved online, but we consider this mostly a benefit: it enabled us to have a more active and participatory course instead of a conventional lecture setup. The lectures were mostly pre-recorded in a shorter form, providing pure content, and giving the concrete pointers to further research. The course topics are:

- Introduction to open computing, history and examples, policies, myths and legends
- Introduction to standards, open standards
- Openness of data formats – binary formats, text formats, code pages
- Data formats – CSV, XML, JSON
- XML – introduction, structure, validation, document type definitions, DTD, XML Schema, life cycle, DOM, SAX, XSL, XPath
- JSON – introduction, structure, data types, JSON Schema, JSON-XML comparison and mapping
- Open data - *described in detail in the next section*
- Free and open – introduction to free software and open source, licenses, Creative Commons license, open source projects
- Distributed systems – introduction, models, types of communication, exchange of data, XDR – eXternal Data Representation, remote calls
- Web APIs, REST architectural style, RESTful APIs, maturity levels, OpenAPI
- Security in open systems – authentication, authorization, HTTP authentication, TLS/SSL, HTTPS, OpenSSL, OAuth, tokens (JWT), OpenID, OpenIDConnect

We particularly focus on extending the lectures with weekly activities, either individually or in groups, synchronously or asynchronously. Variations in forms of activities provide all students an opportunity to participate, regardless of their active learning preferences. Together with open-ended lectures with self-research and preparation in advance, these additional activities formed a "flipped classroom" method of carrying out the course.

Students were asked to gradually develop a web application for a particular topic that included open data in the five laboratory exercises. They had to:

i) model, create or reuse, describe, license and provide an open dataset on their topic of choice, both in a database and available for export in CSV and JSON formats

ii) create a web user interface for accessing and filtering the provided open datasets, additionally describing the metadata of the dataset using JSON Schema specification

iii) develop a high-maturity REST API for exposing their dataset to other systems, which provides easier interoperability; describe the API using the OpenAPI specification

iv) integrate their own open dataset with the external data source such as Wikipedia, and provide semantic capabilities in the form of Linked Data using JSON-LD specification and appropriate vocabularies, such as Schema.org

All tasks are obligatory and form one functional system. In addition, but not strictly related to open data, another non-obligatory laboratory exercise included involvement in an open source project of students' choice and solving a simple bug or creating a fix for an issue.

IV. Open Data in the Course – Theoretical Learning

As already stated, open data topics take around 25% of the whole course. While some elements can be found in other lectures, three weeks focus exclusively on open data from various aspects. To be more precise, the topics taught are presented in Table 1.

The pre-recorded lectures included several pointers to additional information. Each week contained a short reading list with selected webpages and examples. There was also an optional "Find out more" section.

This theoretical background gave students a solid knowledge in open data principles, but the hands-on experience was lacking. Therefore, a course group activity was created, using the following approach as its basis.

TABLE 1. OPEN DATA COURSE CONTENT

Week 1	Introduction to open data; definitions, principles and types, FAIR data comparison
	Influence of open data, ways to use it, open data interdisciplinarity, examples of open data in government and science, arguments "for" and "against" open data, stakeholders, advocacy
	The process of creating open datasets, helper tools, licensing choices, open formats and standards for open data, bulk download vs API approaches, open data portals and examples.
Week 2	Metadata - description, issues, EU recommendations. Vocabularies (Dublin Core, DCMI Metadata Terms, DCAT, Schema.org).
	Metadata - vocabulary formats: Microdata, RDFa / RDFa Lite, JSON-LD, linking data contexts using vocabularies
	Open data quality, models and frameworks for evaluation, Tim Berners-Lee's 5-star open data model, ODI Open data maturity model, Maturity model for prioritization of open data decisions, European Data portal example
Week 3	Linked data vs linked open data, linked data design principles, 5-star model (levels 4 and 5)
	Resource Description Framework – definition, usage, serialization formats, vocabularies, OWL, SPARQL
	Linked data examples – LOD cloud, Wikidata Query Service, Geonames, Dbpedia
	Best practices for publishing linked data

V. COVID-19 DATASET ASSIGNMENT – A BASIS FOR COURSE PRACTICE

As part of the Summer School Program organized by the H2020 Twinning Open Data Operational project in September 2020, we organized an assignment on "Exploring interdisciplinary approaches by using COVID-19 data". After the Summer School, participants had to be able to apply interdisciplinary approaches to study the topic of open data. We used open COVID-19 data as a concrete example that would be interesting to all course participants.

We used the following approach. First, participants received a five-minute instruction for the assignment. We provided the participants with examples of available COVID-19 data, such as data from the European Centre for Disease Prevention and Control, a European Union agency. Moreover, we discussed several considerations for using open data. Participants had to consider what data is available, from what disciplines, what the quality of this data was, what metadata is available and needed, how it can be interpreted, what conclusions can be drawn from this data, how this data compares to other datasets from other disciplines, and how policymakers can use the new insights from this data in their decision-making processes.

Second, Summer School participants from different disciplines were assigned to four interdisciplinary groups to work for 25 minutes on the assignment. They had to look for one openly available dataset concerning COVID-19. We mentioned possible data sources, including those of governmental organizations (e.g., governmental health agencies), individual researchers or groups of researchers (e.g. working at universities), and research institutes (e.g., The Netherlands Institute for Health Sciences). The dataset might cover one country or multiple countries / areas. Students created a presentation containing four slides:

i) A description of the dataset: title, URL, data provider, topic of the dataset

ii) The main characteristics of the data: what metadata is available? (e.g., about data sources, data manipulation, data interpretation and use)

iii) The interdisciplinarity: what disciplines (do you think) are involved in the collection, interpretation and use of the selected dataset?

iv) Decision-making aspects: how can this data (possibly in combination with other data) be used to help governmental policymakers address interdisciplinary COVID-19-related problems? What are the conditions and constraints for policymakers to use this dataset? (e.g., legal, societal, political, economic) Teams were interdisciplinary but could take a particular focus in defining the conditions and constraints on slide 4.

Third, in a twenty-minute session, each group briefly presented their findings (2,5 minutes), followed by 2 minutes of feedback per group. Some examples of identified datasets concerned the characteristics of COVID-19 cases in a country or region, predictions of numbers of infections over time, the effects of COVID-19 measurements implemented in a particular country, and citizen preferences for COVID-19 measurements.

Participants found the assignment surprisingly insightful. Many different datasets appeared to be available, but often these were not comparable to other datasets because of interoperability issues. Different countries register COVID-19 data in different ways. The type of data was easy to relate to. Students also indicated that the available time for the assignment was limited and more useful insights might be obtained if more than one hour would have been available. The time for presenting the findings and giving feedback was also limited.

VI. OPEN DATA IN THE COURSE – PRACTICE WITH REAL-WORLD OPEN DATASETS

A. Assignment Setup

Building upon the group activity explained in the previous section, we have decided to adapt it and include it in our course as a student group activity, which would give them a quick glimpse of real-world needs for open data and ways to solve issues using open datasets.

After the week of learning about open data, where the theoretical concepts were introduced, students had this synchronous activity during the online lectures. They were randomly divided into groups, each consisting of three team members. They were asked to think of a specific topic of using open data for some added value. Students were introduced to the five steps of the assignment to reduce their uncertainty about the task:

1. *Choosing the problem/topic* –a topic in a society where open data could be used to increase someone's quality of life
2. *Finding at least two different open datasets* for this purpose from different sources
3. *Analyzing and comparing the selected datasets* – types and formats of data, available metadata, license, quality …
4. *Problem solving and interdisciplinarity of the approach* – are there other stakeholders to be included in the process? Which areas should they cover? Why are they needed?
5. *Added value and decision making* – which added value could be obtained from these datasets? How to do it? Could there be another dataset (not necessarily available now) that could be combined for even more added value? How could this help with making decisions and policies?

The task was timed at 60 minutes. To help students with time organization, we proposed that four parts of their task take around 15 minutes each: *discussing the topic and finding the datasets*, *analyzing datasets*, *solving the problem* and *creating the presentation.* A structured presentation template was provided on Google Drive for each group, so students could focus on the important tasks. Each group had 5 minutes for solution presentations.

B. Assignment Results

In the current year, 21 of 25 students enrolled in the course took part in this group assignment, in 7 teams of 3 students. In teams of 3 everybody can equally participate, while 7 open data topics would be suitable for showing the differences in approaches. We also needed to keep the course time limits, as all teams would listen to each other's results. The teams chose the following topics:

- *Study programs on higher education institutions* – availability of academic programs in different countries as open dataset, for comparison by future students and easier mobility
- *Kindergartens in Rijeka and Zagreb (cities in Croatia)* – detailed descriptions of kindergartens: the number of free places, waiting list capacity, need for new kindergartens in dense areas.
- *Train transport comparison* – visualizing often used train routes for higher maintenance priority, analysis of density for adding/removing lines
- *Museums in Zagreb* – the list of museums with maps, popularity among visitors, prices, more detailed museum descriptions, navigation
- *The map of cultural tourism* – open datasets containing registers of art organizations, cultural goods and theaters, their locations, contacts, possible connections to tourist organizations
- *Hospitals* – contact details, location, number of (free) beds, the list of departments/services
- *Wifi hotspots* – detailed information, location, technical characteristics, possible connection to tourism and public gathering organizers

Here are several general observations on the assignment results from the teaching staff.

(i) The task of comparing two similar datasets, often from different countries or cities, enabled students to see the shortcomings of the dataset with lesser quality and question some of the dataset design decisions; both opportunities seem helpful for developing critical thinking and evaluation.

(ii) Students had the opportunity to discuss technical issues, such as dataset formats or the difference between bulk download and APIs with filtering and other functionalities. Seeing these differences in their research could be beneficial for their future technical decisions.

(iii) Students observed the need for regular dataset updates, as some were not updated for 2+ years. For some ideas, such as the number of free beds in hospitals, or train delays, real-time updates – or anything close to it – should be available.

(iv) Our students rarely think about interdisciplinarity. Although quick and straightforward, this task provided a way to think about other stakeholders that should be included in problem solving or in a better solution reach, thus creating additional learning value.

(v) It seems that students did not have a clear picture of the difference between the value of data and the value of creating the application to use the data. In some cases, the topic quickly moved to the application level, neglecting the data aspect. This is expected, as students did not have previous experience with data handling value, while constantly observing the benefits coming from various applications.

VII. Students' Evaluation

After the assignment was finished, a short evaluation was performed to gather more information about the specific assignment experience, but also about the acceptance of such interactive ways of learning in a flipped classroom instead of conventional lectures. A short anonymous form was prepared for students in *Moodle*.

Out of 21 students who participated in this task, 17 students answered the questions in our evaluation. We did not have any demographics questions, as this was designed as a quick evaluation, and as students are very homogenous in most aspects (same academic year, similar age, etc.).

Here we provide a summary of the evaluation results grouped according to the aspects that were evaluated. Most questions had some of the answers proposed, together with the *Other* text box for a free-text form.

A. Topic of Choice and Open Datasets Characteristics

Several questions dealt with how students handled the specific tasks: their process of choosing the topic, finding and analyzing the datasets. Here is a brief overview:

*How did you **choose the topic** to work on? Shortly describe the process of choosing the topic/issue.*

The majority (82 %) reported that they first *searched for suitable datasets*; out of these results, they decided on the topic. 18% wrote that they *had an idea for the issue, then they searched for appropriate datasets*. Almost all students were searching *open data portals* (mainly the one from the Croatian government) to obtain ideas. Some reported that they *compared the different portals to find similar datasets* and thus propose an issue that would be easy to solve. A few students (probably members of the same group) mentioned that one of them had the idea, based on the example presented in the assignment instructions.

*How did you find the **first/second dataset**? How **complicated was it to analyze** the first/second dataset?*

Students found the first dataset mostly on the portals, 59% on the Croatian portal, and 29% on some international portals. They had a slightly different approach for the second dataset: 35% found it on the Croatian portal, 41% on an international portal, and 18% using Google.

Students had to grade the analysis with grade 1-5, 1 being "*very complicated, it was hard to find almost all needed data of the datasets*", and 5 being "*very simple, it was easy to find almost all needed data of the datasets".*

The first dataset analysis was assessed as quite easy and with an average grade of 3,88. The second dataset analysis was harder, with an average grade of 3,5 and almost 20% of students grading it with grade 1 or 2.

B. Distribution of Time Used in the Assignment

As the assignment was limited to 60 minutes, we posed a set of evaluation questions regarding the time needed for the task and its distribution among subtasks. The first question was: ***How much time** (in minutes, out of 60 minutes) did you **spend on the subtasks**?*

Discussion about the topic and searching for the datasets – 16 minutes (average answer value)

Analysis and comparison of datasets – 15

Problem solving, value adding, making decisions – 14

Presentation preparation – 12

Figure 1 depicts that although there was some expected dissipation, we can conclude that our advice – dividing the assignment into four 15-minute tasks – was realistic.

*Briefly describe **what took the most time** during the dataset **analysis phase**?*

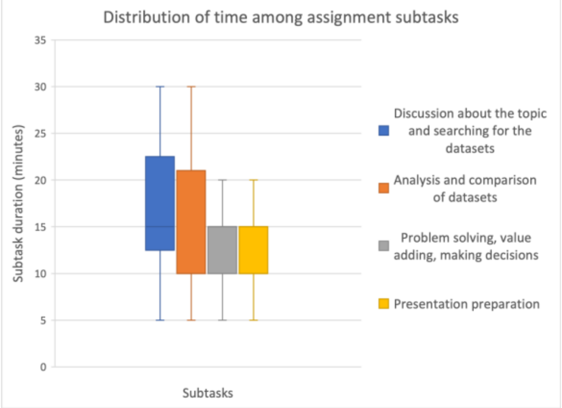

Figure 1. Subtask duration

Students mentioned the following elements: metadata analysis, dataset similarity analysis, solving technical issues with opening the datasets, proposing the added value, even translating the dataset in foreign language.

*Was there **enough time** for the assignment?*

A majority of students, 71 %, considers the time limit to be *appropriate*, 24 % finds there was *too little time*, and 6 % thinks there was *too much time*. The time limit they proposed was 70, 75 and 90 minutes.

C. Students' Self-Assessment of Assignment Aspects

In the evaluation, we also wanted students to assess their own solutions with the following questions:

*In your opinion, **how feasible is your proposal** for the improvement of society using open data?*

This self-reflecting question yielded interesting results: the majority, 59%, thinks the proposal is probably feasible, 6% believes it is surely feasible. Additional 29% thinks it is equally (un)feasible, while 6% thinks it is probably not feasible. No one assessed their idea as surely unfeasible.

*How **useful** was this assignment for **understanding the concept** of open data and its usage?*

The majority, 65 %, finds the assignment *useful* (grade 4), additional 18 % think it is *very useful* (grade 5). Others find it neutral. There were no negative answers.

D. Questions Related to the Course and Assignments

In the last part, we wanted to know how students reacted to this assignment, what do they think about group/personal assignments during the lectures, and how many similar tasks would be appropriate for this course:

*How are you satisfied with the fact that the assignment was done in **groups**, compared to a **personal** one?*

Students are mostly either *very satisfied* (47%) or *satisfied* (41%) with the group work setting, the others having no preference.

*How many "real-world" **synchronous group** assignments like this should be in this course?*

This question brought an unexpected answer, with an average of 4 group assignments during the lectures, and 40% of students asking for even 5 group assignments during the 13-week lecture period.

How many "real-world" ***synchronous personal*** *assignments like this should be in this course?*

Most students (35%) proposed two personal assignments; additional 35% asked for 0 or 1 assignment.

How would you – at this moment (5 weeks of the course) – rate how ***interesting*** *is the course?*

The majority, 59% gave the grade 4, additional 24% grade 5, and 18% grade 3. There were no students who found it not interesting.

How would you – at this moment (5 weeks of the course) – rate how ***useful*** *is the course?*

Most students find the course *very useful* (41%) or *useful* (35%), and 24% gave a neutral grade. There were no negative answers.

Did you ***prepare*** *before the group assignment?*

The majority (53%) both *read the presentations* and *watched the pre-recorded lectures*. 24% just *read the presentations* while the other 24% *did not prepare* at all.

VIII. Discussion

The evaluation brought a few points to consider:

Students' preparation for assignments. FER students are unaccustomed to the "flipped classroom" approach, with most of the lectures not including interaction and consequently, students not preparing beforehand. Although we announced the assignment well in advance, 24 % of our students did not prepare at all. This might be a problem for group work, as these students are not ready to participate in a deeper level of assignments. We plan to put a stronger emphasis on the students' need to substantially prepare for this interactive in-classroom tasks.

Societal elements in the assignment. Our students are very technology-oriented and lack interdisciplinarity in their approaches to problem definition and solving. The inclusion of non-technical factors that constitute societal context, such as *stakeholders* or *added value for society*, brought unexpected issues for some of them. However, this approach proved beneficial, and should be more frequently used in the study program, facilitating the understanding of the topics under study.

The choice of topics/issues to solve. Most students first explored open data portals to get the idea on available datasets, and out of this information, they obtained the inspiration for the issue to solve. While this is a legitimate approach, we would prefer the other way around: finding the problem and then seeing whether open data could help solve it. This might be harder but could bring students a deeper understanding of open data concepts. The previous point of a higher societal context could be beneficial here.

The number of assignments. The results show that students would like to have ~6 synchronous assignments during the 13-week lecture period, roughly taking place every second lecture week. On the other hand, later in the course, some students commented that they were always "solving mini-tasks" instead of "learning the course content." While we can argue that "mini-tasks" indeed are the right way to learn the course content, we should be cautious not to overload students with such activities. We should both analyze their time spent and advise to consider these mini-assignments as real learning opportunities.

Feasibility of ideas. Students self-assessed their proposal as quite feasible. These positive reflections confirm their understanding of open data value and its importance for society, fulfilling an important course goal.

IX. Conclusion

Instead of a usual conclusion, we would like to finish with one of the anonymous student comments that resembles the point of this real-world assignment and the global aim of the course: "*I think I have just now realized how 'openness' is omnipresent.*" This comment fits our assignment and course objectives: introduce future engineers to open data aspects before they enter the "real world". We plan to continue improving the course in this direction.

Acknowledgment

We would like to thank all students of Open Computing course, especially those who enrolled in the year 2020/21, for helping us to create a better course. We would like to thank prof. Mario Žagar, who introduced the concept of Open computing to our Faculty and started this course many years ago. This research is part of the Twinning Open Data Operational project that has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 857592.

References

[1] Open Data Charter, *International Open Data Charter*, no. September. 2015.

[2] T. Davies, S. B. Walker, M. Rubinstein, and F. L. Perini, *The State of Open Data: Histories and Horizons*. 2019.

[3] J. Atenas, L. Havemann, and E. Priego, "Open Data as Open Educational Resources: Towards Transversal Skills and Global Citizenship," *Open Praxis*, vol. 7, no. 4, 2015, doi: 10.5944/openpraxis.7.4.233.

[4] T. Coughlan, "The use of open data as a material for learning," *Educational Technology Research and Development*, vol. 68, no. 1, 2020, doi: 10.1007/s11423-019-09706-y.

[5] R. Magnussen, A. E.-12th C. S. Collaborative, and undefined 2017, "Student re-design of deprived neighborhoods in Minecraft: Game-assisted community-driven urban development," *vbn.aau.dk*, Accessed: Feb. 20, 2021. [Online].

[6] M. Saddiqa, L. Rasmussen, R. Magnussen, B. Larsen, and J. M. Pedersen, "Bringing open data into Danish schools and its potential impact on school pupils," 2019, doi: 10.1145/3306446.3340821.

[7] J. Atenas and L. Havemann, "Open data as open educational resources: case studies of emerging practice," *Open Education Working Group*. 2015.

[8] N. V. Diniz, M. Marchiori, and A. M. G. Zuiderwijk, "Open data infrastructures: The design of an infrastructure to enhance the coordination of open data use," *Conexão - Comunicação e Cultura*, vol. 9, no. 17, 2014, doi: 10.1037/015514.

[9] M. Love, C. Boisvert, E. Uruchurtu, and I. Ibbotson, "Nifty with data: Can a business intelligence analysis sourced from open data form a nifty assignment?," in *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, 2016, vol. 11-13-July-2016, doi: 10.1145/2899415.2899431.

[10] I. Dunwell, R. Dixon, K. C. M. Bul, M. Hendrix, P. M. Kato, and A. Ascolese, "Translating open data to educational minigames," 2016, doi: 10.1109/SMAP.2016.7753400.