Journal of Media Literacy Education, 12(3), 102-119, 2020
https://doi.org/10.23860/JMLE-2020-12-3-9
ISSN: 2167-8715

Data literacy on the road: Setting up a large-scale data literacy initiative in the DataBuzz project

Tom Seymoens

iMEC-SMIT - Vrije Universiteit Brussel, Belgium

Leo Van Audenhove

iMEC-SMIT - Vrije Universiteit Brussel, Belgium

University of the Western Cape & Flemish Knowledge Centre on Digital and Media Literacy, Belgium

Wendy Van den Broeck

iMEC-SMIT - Vrije Universiteit Brussel, Belgium

Ilse Mariën

iMEC-SMIT - Vrije Universiteit Brussel, Belgium

OPEN ACCESS

Voices from the Field - Peer-reviewed article

Citation: Seymoens, T., Van Audenhove, L., Van den Broeck, W., & Mariën, I (2020). Data literacy on the road: Setting up a large-scale data literacy initiative in the DataBuzz project. *Journal of Media Literacy Education, 12*(3), 102-119. https://doi.org/10.23860/JMLE-2020-12-3-9

Corresponding Author:
Tom Seymoens
tom.seymoens@vub.be

Received: October 22, 2020
Accepted: November 11, 2020
Published: December 14, 2020

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Editorial Board

ABSTRACT

This paper presents *the DataBuzz Project*. *DataBuzz* is a high-tech, mobile educational lab, which is housed in a 13-meter electric bus. Its specific goal is to increase the data literacy of different segments of society in the Brussels region through inclusive and participatory games and workshops. In this paper, we will explore how to carry out practical data literacy initiatives geared to the general public. We discuss the different interactive workshops, which have been specifically developed for *DataBuzz*. We highlight the background, design choices, and execution of this large-scale data literacy initiative. We describe the factors that need to be taken into account to reach successful execution for such an ambitious project and the actions undertaken to become a long-term, sustainable solution. Throughout the article, we use the Data Literacy Competence Model as an analytical lens to analyse individual projects on data literacy and *DataBuzz* as an integrated project.

Keywords: *data literacy, education, interactive workshops, vulnerable social groups*.

INTRODUCTION

In this special issue, several authors have argued that in our "age of data" (Barassi, 2019; Bhargava et al., 2015; Mojsilovic, 2014) data literacy is becoming more important. However, how do we reach a widespread adoption of data literacy? How do individuals become more attentive, critical, informed and empowered stakeholders in relation to data, algorithms, and artificial intelligence in an increasingly mediated environment (Markham, 2020)? The answer to this question is not a simple one, as opinions differ on what it means to be data literate. The majority of circulating definitions seem to lean toward an instrumental framing, which stresses the importance of the required technical, computational and statistical competencies to become a data-literate individual (Gray et al., 2018; Raffaghelli & Stewart, 2020). Increasingly, definitions are developed that take a more critical stance, enabling the analysis of technological and political hegemonies that influence the datafication processes and outcomes (Gray et al., 2018; Markham, 2020; Raffaghelli & Stewart, 2020). In this article, we start from the Data Literacy Competence Model (DLCM), developed by the Flemish Knowledge Centre for Digital and Media Literacy (see Figure 1). This model comprises two major competence clusters *using data* and *understanding data* (Mediawijs, 2020). Competences refer to the knowledge, skills and attitudes that allow individuals to act adequately in a given situation (Mediawijs, 2020a). The competence clusters are defined in more detail as:

- *Using data*, or the knowledge, skills and attitudes to use data actively and creatively, namely:
  - interpreting: being able to read a graph, a table, a list of data and understand what they mean;
  - navigating: finding your way through a collection of different types of data and ways they were processed and being able to extract the message or what you need from them;
  - collecting: being able to set up a process to collect raw data and organise an analysis of it; and
  - presenting: being able to present and visualise the results of a data analysis in a targeted manner, tailored to an audience.
- *Understanding data*, or the knowledge, skills and attitudes to critically and consciously assess the role of data, namely:
  - observing: being able to observe how data is communicated and used;
  - analysing: being able to analyse the individual and social consequences of the way in which data is communicated and used;
  - evaluating: being able to evaluate whether those consequences are harmful or constructive; and
  - reflecting: being able to reflect on how the way in which you and others communicate and use data should be adjusted to minimise the harmful consequences.

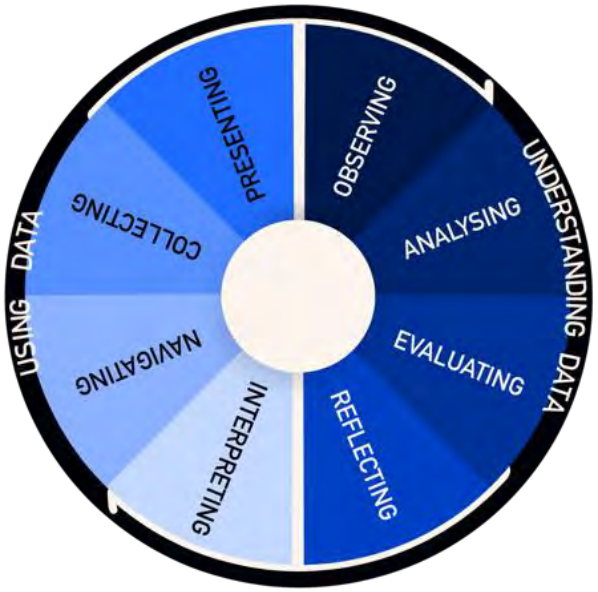

Figure 1. *Data Literacy Competence Model*

The cluster of competences for *using data* is more practice-oriented. However, this cluster breaks with the view, promoted by a lot of data literacy researchers, that data literacy starts with identifying a problem to which data analysis is the answer. The model of the Knowledge Centre follows the levels of literacy in relation to data. It starts with the questions: Can I read data? Can I navigate different types of data? Can I organize data in order to analytically understand them? Can I collect and work on existing and new data? Can I present and communicate those data? Therefore, the sub-competences include the following: 1) interpreting; 2) navigating; 3) collecting; and 4) presenting.

The cluster of competences for *understanding data* is more oriented towards critically and consciously

understanding the role of data in society, personal life, etc. This focus is often neglected or underdeveloped, especially by those authors who have a background in statistics, data analytics, computer sciences or coding. These competences are closely linked to reflections on the role of data in the social sciences. The competence model of the Knowledge Centre highlights the importance of bringing both clusters of competences into one overarching model.

In this paper, we will explore how to carry out practical data literacy initiatives geared to the general public. First, we provide an overview of the limited number of cases we found for the implementation of projects on data literacy worldwide. Second, we present *the DataBuzz Project*, a high-tech, mobile educational lab designed and used to reinforce data literacy in Belgium. Third, we discuss the different interactive workshops, which have been specifically developed for *DataBuzz* and tailored to different segments of society. We highlight the background, design choices, and execution of this large-scale data literacy initiative. We describe the factors that need to be taken into account to reach successful execution for such an ambitious project and the actions undertaken to become a more long-term, sustainable solution. Throughout the article, we use the DLCM as an analytical lens to analyse individual projects on data literacy and *DataBuzz* as an integrated project.

Initiatives on data literacy

In the field of media and digital literacy there is a large body of academic literature reporting on specific initiatives and projects with a focus on fostering these literacies. As data literacy is in its infancy, there are not many accounts of projects in this field. Figure 2 displays a list of six relevant projects that we identified and analysed in detail. We will not report on this in depth. Rather, for each project we describe the set-up, the target audience, the theme, and the targeted competences. We also map these projects on the DLCM (see Figure 2 and Table 1).

We can draw several insights from Figure 2 and Table 1. First, the majority of the projects target a rather restricted audience of students. Second, most initiatives are limited to workshops or exhibitions. Third, initiatives tend to either focus on *using* or *understanding data*. Initiatives rarely cover the full spectrum of competencies in the DLCM. Most projects seem to focus on competences in the cluster of *using data*. This finding is in line with the observation that much of the literature on data literacy is highly utilitarian and inspired by the scientific staged research method: Identify, Collect, Evaluate, Analyse, Interpret, Present, Protect (see for example, Oceans of Data Institute, 2016). Fourth, most initiatives are small scale, stand alone, and non-recurring. In terms of experimenting and learning these initiatives are highly useful. However, they likely do not achieve a long-term, sustainable impact on society.

Figure 2. *Data literacy initiatives in academic literature, mapped on the DLCM*

THE DATABUZZ PROJECT

Initial conceptualisation: DataBuzz for Brussels' students

The Minister of Finance of the Brussels Capital Region and head of the Flemish Community Commission (VGC) responsible for the Flemish Schools in Brussels initiated *the DataBuzz Project* on the 23rd of April 2019. In a schoolyard in one of the suburbs of Brussels, a class of girls and boys aged 12-13 years participated in the *Escape the DataBuzz!* workshop by solving data related quizzes, riddles, and questions on privacy and protection of personal data. This was the official start of a mobile education lab, housed in a 13-meter electric bus, which brings data literacy courses and materials to the different municipalities and schools in Brussels (see Figure 3).

Table 1. *Data literacy initiatives*

Publication	Set-Up	Audience	Theme	DLCM
Argast & Zvyagintseva (2016)	Workshop and hackathon	The general public	Open data	*Using Data*: Interpreting, Navigating *Understanding Data*: Observing, Analysing, Evaluating
Deahl (2014)	Weeklong summer camp (Festival of Code)	Young programmers (younger than 18)	Coding	*Using Data*: Interpreting, Navigating, Collecting, Presenting
Deahl (2014); Williams et al. (2014)	High school math class pilot test (City Digits)	High school students (aged 17-18)	Local community data	*Using Data*: Interpreting, Navigating, Collecting, Presenting
Wolff et al. (2019)	Primary and Secondary school curricula (Urban Data School)	Children (aged 10-14	Smart city data	*Using Data*: Interpreting, Navigating, Collecting, Presenting
Bhargava et al. (2016)	Workshop at a Secondary School	Students (aged 16-21)	Local community data	*Using Data*: Interpreting, Navigating, Collecting, Presenting
Markham (2020)	Arts-based installations at events	The general public	Personal data impacts Data ethics	*Understanding Data*: Observing, Analysing, Evaluating, Reflecting

Concretely, *DataBuzz* will be visiting schools and playgrounds in Brussels for at least the next two years. During their visit to the bus, participants have the chance to experiment with novel technologies, such as virtual reality glasses, face recognition software, and sensors. This program allows them to witness the possibilities of these innovations and understand their underlying principles. The project currently offers seven different thematic and interactive workshops for groups of kids between 10-18 years old. These workshops focus on *understanding data* as well as *using data* in a hands-on educational environment. In the next section, we will present and analyse each of these workshops.

Figure 3. *The DataBuzz*

This project was conceptualised and initiated by the research group, imec-SMIT from the Vrije Universiteit Brussel (VUB); has been largely financed by the VGC; and is executed in close collaboration with imec, an international innovation hub in nanoelectronics and digital technologies. For the VUB and imec, this venture is in line with their research interest in smart education, as well as their societal role to promote awareness and garner interest in innovation and technology for the broader public. The VGC, in turn, is charged with the governance of the Flemish-speaking schools in the Brussels-Capital Region. This is a challenging responsibility due to the high level of diversity in socio-economic and cultural background of the school-age population (VGC, 2019). In addition to these initial partners, more and more parties collaborate directly and indirectly in *DataBuzz*. The bus company Keolis manages and drives the bus; the Knowledge Centre for Digital and Media Literacy, Gluon and BILD co-developed education content; and the telecom provider Telenet provides high speed datalinks to the bus.

The ultimate aim of *DataBuzz* is to strengthen data literacy in youth through partnership education programs with their schools. Recent research shows that data literacy skills are unevenly distributed among the Belgium population, largely following socio-economic and cultural boundaries. In addition, Belgian youth in general lag behind the European average with regard to data competences (Brotcorne & Mariën, 2020; European Commission, 2018). It is quite possible that this issue is even more true for students in the Brussels Capital Region, which is characterized by large inequalities and segregation between areas and schools.

Brussels schools are segregated between 'elite schools' and so-called 'problem schools' (Wayens et al., 2013). This trend can become a self-perpetuating cycle, where social inequalities lead to educational inequalities, which in turn reinforce the former, and so on and so forth (Rea et al., 2009).

The result is an educational system that is not adequately equipped to widely develop data competences across a diverse school population. Few schools possess the personal and financial resources to develop actions in this field. The development of data literacy could take place in extra-curricular initiatives; however, this again can lead to imbalances. For example, people from socially vulnerable segments of society rarely come into contact with activities organised outside of the school.

DataBuzz offers a temporary solution to some of the challenges we have identified thus far. First, it guarantees that we can reach all of Brussels' youth, as it is compulsory for youth between the ages of six and 18 to attend school. Second, this project allows all schools to gain equal access to expensive resources, such as the most modern infrastructure and technologies, which would otherwise be unattainable. Third, the course materials are developed by a dedicated team, which disseminates them free of charge to the Brussels' schools. Fourth, materials can be locally adjusted to student and teacher preferences and context. Fifth, the deployment of the available material with a large number of students and over a longer period of time allows for iterative content development.

DataBuzz as a platform

The upfront investment and running costs of a mobile educational lab with advanced technology was very high. In addition to the purchase of an electric bus, the bus needed a complete new and tailor-made interior, the technology needed to be purchased and installed, a team needed to be hired; a communication plan devised and executed; the lesson plans and educational material needed to be developed; an operational planning for the school visits needed to be set up, etc. As soon as *DataBuzz* was up and running for school visits, we started to see it as a platform upon which other initiatives could also be built by looking for additional project opportunities and partnerships. In this way, we hoped to ensure that *DataBuzz* could become a successful, long-term, sustainable project, operating on a large scale with maximum impact. These additional collaborations focus on: (a) linking *DataBuzz* with other existing initiatives in the data literacy domain; (b) upgrading the existing workshops; (c) developing/integrating new educational material; (d) targeting new segments of society; and (e) acquiring new research funds in relation to projects for *DataBuzz*.

One such initiative was the *Smart Spaces* workshop (see later for a description), which was developed together with Gluon, a Brussels-based organisation that focusses on teaching digital and critical thinking skills to young people through visual arts program. Another project is *Everyone Data Literate!* that focuses on the creation and execution of lesson plans for adult education. A different target group is reached through the 13 Centres for Basic Education (CBE) and 49 Adult Education Centres (CVO) throughout Flanders. Four new lesson plans were created, each highlighting a different aspect of the data society. The scope of two of these workshops is explained in the next section. One of the additional target groups we tried to reach were incarcerated inmates, as media, digital and data literacy is a huge problem amongst this often-forgotten population, but unfortunately the project was not funded.

An operational effort

Although we could have used a larger bus, which could host more students at the same time, we opted to use a 13-meter model in order to be able to access the Brussels' school grounds and municipalities without too much difficulty. The 13-meter bus was divided into six different zones (see Figure 4). The various zones represent different facets of the learning trajectory. This was taken into careful consideration during the design. The first zone, known as 'the bus stop,' was where the students had to wait before they could enter the bus. The bus has an awning to shelter passengers waiting to get on board in case of bad weather. There were two doors, one for entering and another for exiting the bus. Students with restricted mobility could enter and leave the bus by making use of its ramp. The second zone, the briefing zone, was used to explain the course of action of the workshops. To do this in a tranquil environment, there were nine seats placed around two tables. The instructors and participants could make use of two touch screens to clarify the design. Zones 3 and 4 were intended for break-out sessions and personal reflection. The zones were smaller in size and had only one seat. The fifth zone was the largest in the bus and was used for activities that required collaboration. It also possessed the largest screen, a 65-inch touch screen,

intended for extensive interaction. The last zone was where students were given the chance to present their work to the tutor and each other at the end of each session.

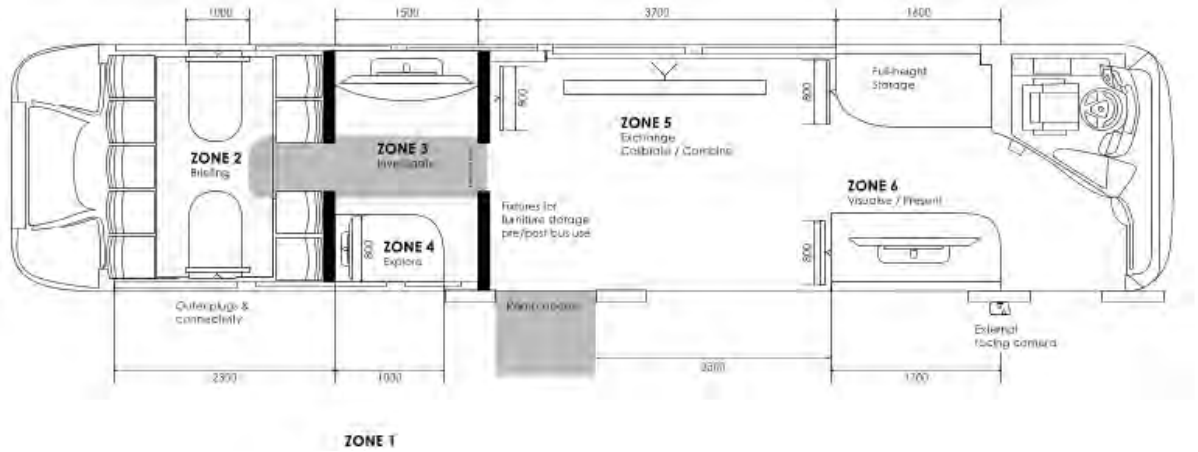

Figure 4. *Six zones inside the DataBuzz*

The bus was equipped with technical equipment as well, including six high-definition touch screens. It had a storage place, which could hold 10 laptops, 30 smartphones, and a set of virtual reality glasses (see Figure 5). In line with the smart education program and the research objectives of the project, the bus was equipped with behavioural analytics software. Together with its hardware components (cameras, smart glasses, ECG equipment, etc.) this enabled eye tracking, galvanic skin response, heart rate activity, and facial expressions. It is important to note that if data was collected, the participants were always made fully aware of its purpose through an informed consent form prior participating in the activity. In the event that the participants were minors – which was a frequent occurrence – three different consent forms were created and these would be signed by the children, their parents, and their teachers prior to participation in the program. A server was installed in the bus for secure data storage. As the main target audience were students between the age of 10 and 18, the bus was designed to be kid-friendly and robust materials were used for furnishings and floors.

The DataBuzz workshops

As of this time, seven workshops have been created for two target groups: (1) Students between the ages of 10 and 18 attending Flemish-speaking schools in Brussels; and (2) adult-learners with an emphasis on functionally illiterate and low-skilled people who have received an education through one of the 13 Flemish Centres for Basic Education (CBE) or one of the 49 Adult Education Centres (CVO). In this section, we discuss the set-up, the target audience, the theme, and the targeted competences of each workshop.

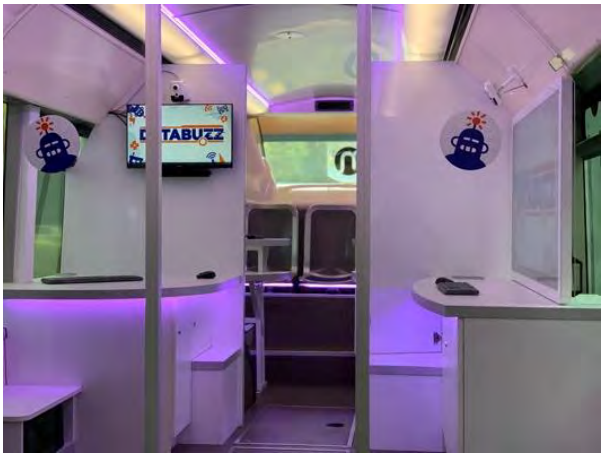

Figure 5. *Interior of the DataBuzz*

'Escape the DataBuzz!': 10- to 13- year-old students

A recent report on the use of digital media by Flemish teenagers between the ages of 12 and 18 has revealed that the majority of students (53%) find it important that there is legislation that protects their personal information. However, only 27% of the respondents were aware of the existence of the general data protection regulation (GDPR). In addition, six out of 10 students mentioned online privacy was rarely a topic in class, although they felt that their teachers were in a position to educate them on this matter (Apestaartjaren, 2020). Schooling initiatives on privacy and personal data control were important, as kids between the ages of eight and 11 struggled to identify

the related risks and did not have a set of internalized privacy rules (Stoilova et al., 2019). Against this background, *Escape the DataBuzz!*, a workshop on online privacy and personal data control, is designed for pupils who attend the last two years of primary school or the first year of secondary school (see Table 2 and Figure 7). The workshop is styled as an Escape Game. Past experimentation has shown that this set-up has a positive effect on student motivation and willingness to learn (Borrego et al., 2017). The workshop was structured as follows. Ten participants were divided into two groups before entering the bus. After the last person had boarded the bus, the doors were closed. The participants discovered that in order to exit the bus, they would need to find a way to get in touch with the bus driver. Throughout the workshop, they were given clues and riddles related to online privacy (see Figure 6). The first team that was able to contact the bus driver won the game.

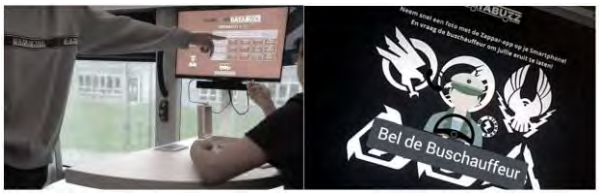

Figure 6. *Escape the DataBuzz!*

At the start, the students are instructed to freely explore the bus and look for hints. This way, they stumble upon a switched-off smartphone and an unsecured box. Inside the box, the participants find various types of information (e.g., a report card, a driver's license, a profile picture, etc.). When the participants figure out that they have to try to unlock the smartphone, eight different pieces of information appear on the screens on the bus. Four of them are regarded as 'personal information' that can be used to identify an individual person. If they can discern which ones are and are not personal information, they are presented with a code that enables them to unlock the smartphone.

Next, the students are asked to study the contents of the box again. As soon as they realize that every piece of information inside the box is personal information, they are shown that the box has to be secured. The students then search for a lock that is placed somewhere on the bus. The lock is hidden very well, and after a short while of unsuccessful searching, the screens on the bus present a hint. On their unlocked smartphone, an application is installed that can help them locate the lock through Bluetooth. In order to use it to safely secure the box, the participants need a code. Another puzzle is presented to them, which teaches them about 'contextual integrity' (i.e., how the appropriateness of sharing personal information depends on the social context in which it is released). This game educates the pupils on how certain information (e.g., their location) is safe to share with one stakeholder (e.g., their parents), but not with just anybody. Through the game a code is formed that enables them to lock the box with personal information.

Figure 7. *Escape the DataBuzz!, mapped on the DLCM*

Table 2. *Escape the DataBuzz!*

Name	Set-Up	Audience	Theme	DLCM
Escape the DataBuzz!	Escape game 50 minutes	Third cycle primary school (aged 10-12) First year secondary school (aged 12-13)	Online Privacy Personal Data Protection	*Using Data*: Interpreting *Understanding Data*: Observing, Analysing, Evaluating, Reflecting

While locks are used to safely store physical information and documents offline, digital information is protected by passwords. The students are given eight rules to help them create a strong one. They then receive a list of potential passwords and, armed with the set of rules, they are asked to identify the strongest one. This

password can be used to unlock a laptop that is handed to them. Once they are successfully logged onto the laptop, the students have to finish a quiz that teaches them about the existence of the GDPR, the legislation behind the reuse of personal information, the dangers of sharing their passwords, photo rights, and secure data storage. If all questions are answered correctly, a symbol appears.

The students can scan this symbol with the Zappar-application[1] on the smartphone. This makes it possible for them to call the bus driver to help them escape.

'How smart is my phone?': 12- to 14- year-old students

The average age at which Flemish teenagers between the ages of 12 and 18 have acquired their first smartphone is currently at 11 years and one month. This is almost one year earlier than studies from two years ago, which indicates that the average age of first smartphone possession is rapidly dropping. In addition to this, 45% of the teenagers have the feeling that they spend too much time on their smartphones (Apestaartjaren, 2020). 36% state to use their smartphone more than 4 hours on average schooldays.

With *How Smart is my Phone?*[2], *DataBuzz* brings two workshops to students between the ages of 11 and 14, which allows them to critically reflect upon their screen time and develop strategies for a digital balance (see Table 3 and Figure 9).

Three weeks before the arrival of *DataBuzz,* the students are asked to install a research application, developed at the University of Ghent (see Figure 8). This application tracks and analyses their smartphone use.

At the start of the first workshop, the students are asked to make an estimation of their smartphone use. They indicate on a form, for example, their perceived screen time; the number of notifications they receive on a daily basis; and the time they spend on social media. Once this task is performed, they compare their perceived smartphone use (the form data) with their actual use (the data from the application).

Through a set of follow-up questions, the students reflect upon the differences and the findings that strike them most odd.

Figure 8. *Mobile DNA*

Subsequently, the students are divided into groups of four. They have to record a vlog with a smartphone or tablet in which they react to one of the suggested topics, such as digital balance, the endless scroll, and the joy of missing out. This exercise facilitates a conscious contemplation on their current smartphone usage.

Figure 9. *How smart is my phone?, mapped on the DLCM*

In the last session of this first workshop, the students are again divided in smaller teams. They are asked to come up with one challenge, relating to a balanced smartphone use, that they and their peers have to adhere to in the next couple of days. Examples include "*Leave your smartphone out of your bedroom*" or "*Disable all notifications*". After they devise their challenge, they create an infographic to explain to their peers how this relates to a more ideal smartphone use. The workshop

[1] https://www.zappar.com/getzappar/

[2] This workshop is created by Mediawijs, a partner of the DataBuzz project.

ends with a discussion on the feasibility of the challenges and the students' takeaways of the session.

Several days later, *DataBuzz* returns for the second workshop. The students divide themselves into different teams, and each team receives an information sheet. The sheet displays figures about the smartphone use of teenagers, how digital media affect their life and how well-drafted school regulations could bring more balance. After discussing these findings, the different groups are brought back together, and the students have to explain to each other what they just learned.

In the last part, the students play a boardgame. They have to roll the dice. Each number is linked to a specific location or setting (e.g., "*in the hallway*", "*in the classroom*", or "*while eating*"). The students then have to pick a card with a statement that refers to the setting (e.g., "*It disturbs me that a friend is checking Instagram when we're eating together*"). Through a guided discussion on each of the statements, the students decide amongst themselves on a number of strategies and rules that they wish to see included in school regulations.

Table 3. *How smart is my phone?*

Name	Set-Up	Audience	Theme	DLCM
How smart is my phone?	Two workshops 2x100 minutes	First cycle secondary school (aged 12-14)	Educated smartphone use Digital balance	*Using Data*: Interpreting, Navigating, Presenting *Understanding Data*: Observing, Analysing, Evaluating, Reflecting

'Breaking News!': 14- to 16- year-old students

Twenty-nine percent of the Flemish students between the ages of 12 and 18 sometimes feel overwhelmed about the amount of news they have access to, and about one in five are confused about the quality of news and information they receive (Apestaartjaren, 2020). Studies also indicate that adolescents find it challenging to evaluate the quality of information presented in news reports (Ku et al., 2019).

In this context, *DataBuzz* has developed *Breaking News!*, a course divided into two workshops of 50 minutes that teaches pupils age 12-14: (a) how journalists make use and present data sources in their reports; (b) how they can correctly interpret presented evidence; and (c) how they can tell a story themselves based on personally collected data (see Table 4 and Figure 11).

The first workshop starts with a couple of exercises on the interpretation of graphical representations of data. The students are divided into small groups and are asked to discuss different charts that have been used in news reports. They learn how different visualisations of the same data have a great effect on the message that is conveyed.

In a second section, the students are led to the website of the Belgian Statistical Office.[3] The website allows visitors to search for, play around with, and visualize data that stems from the students' environment. Through a number of exercises on the website, the students learn how to compare data of their municipality (e.g., population increase, the number of road accident victims, etc.) with the rest of Belgium.

Figure 10. *A design tool for data visualization (Canva or Piktochart)*

The first workshop ends with the students setting up a small-scale journalistic investigation. They choose a topic and research question and learn how to collect data from their peers that enables them to formulate an answer. Together with the supervisor of *DataBuzz*, they learn how to program an online questionnaire with the proper questions that relate to their research question. This first workshop concludes with the students distributing the survey to their peers.

Several days later, when all data has been collected, the *DataBuzz* returns for the second part of the

[3] https://www.statbeljunior.be/

'*Breaking News!*' course. This session starts with extracting the gathered data from the online platform, cleaning the data in a spreadsheet, and summarizing the results in a pivot table. The supervisor helps the students to come up with a valid interpretation of the data.

In the second part of the program, the students learn how to use a simple graphical design tool (see Figure 10). In this way, they develop an infographic to visualize their findings in a visual-pleasing manner. This exercise teaches the students how to make a story based on data. The workshop ends with a group presentation of the infographics they designed.

Figure 11. *Breaking News! mapped on the DLCM*

Table 4. *Breaking News!*

Name	Set-Up	Audience	Theme	DLCM
Breaking News!	Two workshops 2x50 minutes	Second cycle secondary school (aged 14-16)	Data Journalism	*Using Data*: Interpreting, Navigating, Collecting, Presenting *Understanding Data*: Observing, Analysing, Evaluating, Reflecting

'Smart Spaces': 16- to 18- year-old students

Many of the educational initiatives around data literacy in a school setting focus on student interaction with small, personally collected data sets. We adhered to a similar approach in *Breaking News!*. Although beneficial, these approaches should be complemented with initiatives that deal with larger and more complex data sets, as research has shown that youth do not understand how larger data sets are collected; do not have the skills to interpret them correctly; and do not understand how to relate the information contained therein to their own experiences (Wolff et al., 2019). Against this background, *Smart Spaces* is a two-part lesson plan that is centred around complex environmental data and developed for students who are entering their last two years of secondary school[4]. The course not only offers these students the skills to interpret and analyse data sets; it also increases their understanding of how data uncovers hidden layers within their environment and how it can be a useful tool for social engagement and policy action (see Table 5 and Figure 13).

The course starts with an introduction to the ways large-scale data collection can lead to meaningful citizen participation. First, a number of citizen science and smart city projects are briefly showcased to the participants. Next, in order to enlighten the students as to how data is collected in these types of projects, they are divided into smaller teams, each of which receives a box with a raspberry pi (a single board computer, see Figure 12); various sensors (e.g. a sound, motion, temperature, humidity, and distance sensor); a laptop and a guidebook that enables them to build a station for different purposes. Students are guided in the process of building and testing the sensors in their environment. The data appears on their screens in real-time. Together with the supervisor of the session, the students analyse the data collected by the sensors. In the third section, they receive an introduction to the importance of data

[4] This workshop is created by Gluon, a partner of the DataBuzz project.

visualization and how it can be applied to summarize, interpret, and communicate the most important outcomes. This first session ends with the installation of a prefabricated sensor station in the environment of the school facilities. The attached components measure air quality, noise levels, light pollution and traffic.

Figure 12. *Raspberry Pi, used in Smart Spaces*

Several days later, when the sensor station has collected sufficient data, a second session is organised. After a short synopsis of the first session, the students are directed to Modlab website[5]. By making use of this platform, the participants can delve into the collected environmental data in their school environment; compare it to collected data in different regions of Brussels; and view the results. Together with the supervisor and a prefabricated chatbot that asks the participants dedicated questions, the students learn how to interpret the results and how they can be used for societal action and to inform policy.

Figure 13. *Smart Spaces, mapped on the DLCM*

Table 5. *Smart Spaces*

Name	Set-Up	Audience	Theme	DLCM
Smart Spaces	Two workshops 2x100 minutes	Third cycle secondary school (aged 16-18)	Environmental data	*Using Data*: Interpreting, Navigating, Collecting, Presenting *Understanding Data*: Observing, Analysing, Evaluating, Reflecting

'AI-AI-AI?': 16- to 18- year-old Students

Flemish media regularly publish articles about the increasing influence of artificial intelligence (AI) on our society. Some reports focus their attention on the positive achievements of AI; for instance, in the health care sector (Dierickx, 2020; Van de Weghe, 2020); the financial sector (Van Geyte, 2020); and in education (Grymonprez & Maenhout, 2020). Others shed light on potential threats such as those with regard to privacy (Vanhecke & Deckmyn, 2019); job security (Baert, 2020); and discrimination (Evers, 2020). Educational initiatives can help people to form a well-informed opinion from the different views on AI that reach them through the media. For this purpose, *DataBuzz* has developed a workshop *AI-AI-AI?* The course is divided in three sections and has a duration of 90 minutes. It is tailored to students in the third cycle of secondary school between the ages of 16 and 18. The workshop is offered to smaller groups of four to six students (see Table 6 and Figure 15).

In the first part of the course, the participants learn how views on AI have developed over the years. Three different phases are discussed: (1) procedural AI; (2) machine learning; and (3) deep learning. Procedural AI is explained through the following exercise. The students take on the role of system developer and have to specify the instructions for preparing a sandwich with chocolate spread. Another student plays the machine and has to follow the written down specification ad verbatim. This exercise reveals some of the pitfalls of procedural AI. For example, the rules have to be made explicit upfront, and the system will only work for the specified use case.

[5] http://www.modlab.be

Next, the students learn about the principle of machine learning (ML). In this subset of AI, a system is taught by allowing it to learn from input data as opposed to providing it with the correct rules. The students learn of the mechanisms of ML by playing around with Teachable Machine (see Figure 14). The website allows its visitors to create a model that can make a distinction between different classes of images, sounds, and videos. The students learn: (a) the importance of great amounts of quality input data to make their model more efficient; and (b) how skewed input leads to skewed output.

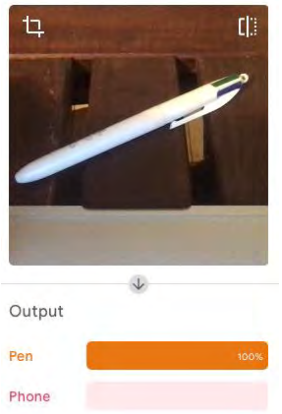

Figure 14. *Teachable Machine*

The students then learn about the philosophy behind deep learning. They learn how this differs from traditional ML, as it does not require human supervision to train its model and it can find patterns in unlabelled data sources. To test their newly acquired knowledge of the three branches of AI, they play a quiz in teams in which they have to determine the used subset of AI in different situations.

The second part of the workshop deals with exploring the impact of AI on everyday life. Four application domains are discussed: media, mobility, healthcare, and music. To explain the use of AI in media platforms, the students study their Facebook timeline, Google search results, and Netflix recommendations on their own devices and analyse how and why they differ from each other. With regard to mobility, students are introduced to the example of Tesla's full, self-driving car. They learn about the possibilities and issues created by this innovation through a guided discussion. For the third application domain, we turn to healthcare. The students learn about the potential power of pattern recognition through the example of identifying malignant breast cancer. Lastly, the participants are asked to listen attentively to two songs. One is composed entirely by an AI system to sound similar to the musical style of The Beatles. The other song is an actual song from a musical group from the 60's. After listening to both songs, they have to determine which song is AI-driven.

In the third and last section of the workshop, the group is confronted with three challenges that AI systems impose on society. First, we return to the case of autonomous vehicles and discuss the relevance of the Trolley Problem. Next, they receive some examples of discriminatory practices that arise from data-based decision making when the input data is flawed. Finally, they learn about deepfake technology; how it is used to create synthetic media; and its effects on the dissemination of disinformation and hoaxes.

Figure 15. *AI-AI-AI?, mapped on the DLCM*

Table 6. *AI-AI-AI?*

Name	Set-Up	Audience	Theme	DLCM
AI-AI-AI?	A Workshop 90 minutes	Third cycle secondary school (aged 16-18)	Artificial Intelligence	*Understanding Data*: Observing, Analysing, Evaluating, Reflecting

'Bye Bye Bias!': Adult learners / 16- to 18- year-old students

Earlier in this article, we referred to a survey of 158 lecturers from facilities that offer adult education, which highlighted the interest in a lesson plan about why AI and algorithms were not always free from bias and the subsequent impact on society. Inspired by the results from this survey, we decided to develop the *Bye Bye Bias!* workshop (see Table 7 and Figure 17). The course was divided into three sections with a total duration of 60 minutes. While initially target towards adult learners, the workshop is also deployed in the third cycle of secondary school.

The first section of the workshop is based on the contents of the *AI-AI-AI?* course. Playing the game of *Quick, Draw!*[6], the participants receive an introduction to the principles of AI. In this manner, the students figure out how systems are trained through a multitude of examples (e.g., thousands of drawings of cats).

After this short introduction to the principles of ML, the emphasis of the workshop shifts to the ways bias can influence the outcomes of these systems. An initial exercise gives the participants the opportunity to reflect upon their own prejudices. They play a game in which they link pictures of real people with their job description. After this exercise, the students then perform a Google Image search for 'Five white teenagers' and 'Five black teenagers' (see Figure 16).

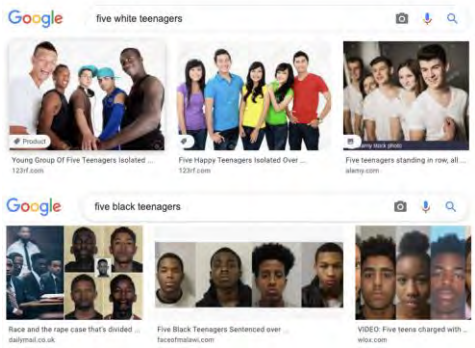

Figure 16. *Discrimination through automated systems*

They then analyse the images they have found in small groups. Each group compares the images they found and contemplates on the different outcome. Based on these two tasks – a reflection of their own prejudices and those of 'machines') they explore the principle of "*Garbage In, Garbage Out*". This refers to the concept that if a machine is trained on a data set that is skewed towards a particular belief, it's future output will replicate these preconceptions. In this way, the participants learn that AI systems can draw attention to the prejudices that exist or have existed in society.

In the third section of the course, the students learn from examples of how this propensity for prejudice can have adverse effects on population segments of society. They are confronted with recent examples from the use of AI-based systems, including face recognition in public places; screening of resumes by HR departments; and assessing creditworthiness by financial institutions. A number of initiatives from the industry, as well as from academia, that have been designed to tackle these harmful effects are briefly presented at the end of the workshop.

Figure 17. *Bye Bye Bias!, Mapped on DLCM*

Table 7. *Bye Bye Bias!*

Name	Set-Up	Audience	Theme	DLCM
Bye Bye Bias!	A workshop 60 minutes	Adult learners Third cycle secondary school (aged 16-18)	Discrimination in algorithms	*Understanding Data*: Observing, Analysing, Evaluating, Reflecting

[6] https://quickdraw.withgoogle.com/

'Cookies and Bubbles': Adult learners / 16- to 18-year-old students

After our survey, we created a third workshop called *Cookies and Bubbles* in connection with the lecturers at the facilities for Adult Education. The course explains how personal data is used by various organisations to deliver targeted content to their users. The focus lies on the personalization of advertising and news content (see Table 8 and Figure 18).

The course is offered to both adult learners and to students in the third cycle of secondary education. The course is divided into two main sections. At the start of the workshop, the participants are divided into smaller groups of five or six students. The students first receive a short, theoretical introduction on the principles behind targeted advertising and personalized news. They learn that everyone leaves digital traces behind through their interaction with platforms, applications, and digital technologies. These traces get picked up by cookies and are used to create a profile about the users. Advertisers and news agencies can target these profiles to deliver content that is in line with their inferred taste and preferences.

Next, the students play three simple games that teach them about the existence of three different types of cookies: (1) functional cookies that support basic features of a platform, such as remembering the contents of the shopping cart in an online shop; (2) analytical cookies that are used to analyse the user actions on platforms and ameliorate their future experience; and (3) tracking cookies that trace and store users steps and behaviour over different platforms and allow the inference of a digital profile that can be targeted with personalized advertising and news. The students are then asked to reflect in groups on why platforms have to ask permission for the use of tracking cookies but not for functional or analytical cookies. In a guided discussion with the instructors, they learn about the benefits of cookies (e.g., the required filtering of information), as well as their drawbacks (e.g., in relation to privacy). The first section ends with five practical tips that the students can apply to make their interactions with digital platforms relatively anonymous.

In the second section of the course, we explore how personalization of news and advertising has an impact on how the world is presented to us. To enable the participants to understand this concept, they have to use their own devices to search for specific terms on Google. They then come together in groups and compare their results. They share which items are presented on the first page and in which order. This way, the participants discover that some of the results are unique to them. They are asked to link these findings with their newly acquired knowledge of cookies from the first section of the course and are asked to reflect on the possible side-effects of personalization of information. The students become acquainted with the filter bubble and the challenges it produces. They learn about how online platform algorithms decide which information reaches us and ensure the information is most in line with our preferences.

Figure 18. *Cookies and Bubbles, Mapped on the DLCM*

Students learn that we have a responsibility to critically question our own perspectives and information-seeking behaviour. Through an exercise, it is revealed that our own beliefs have a great influence on how we perceive reality (see Figure 19). There are a number of simple rules the students can follow that can help them to escape the online bubble as well. The session concludes with the presentation of four tips that will support the students with this undertaking.

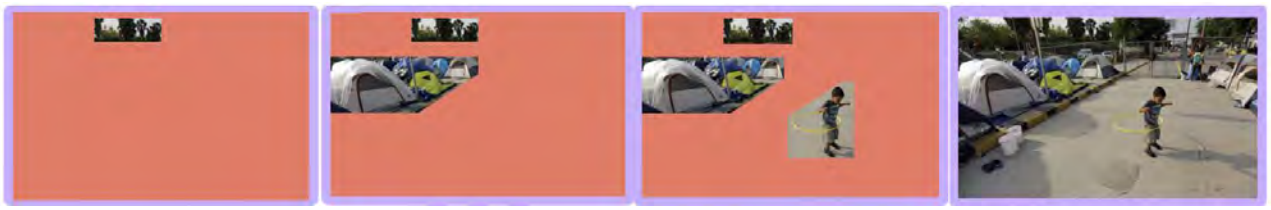

Figure 19. *Confronting students with their own belief system*

Table 8. *Cookies and Bubbles*

Name	Set-Up	Audience	Theme	DLCM
Cookies and Bubbles	A workshop 60 minutes	Adult learners Third cycle secondary school (aged 16-18)	Content personalization	*Understanding Data*: Observing, Analysing, Evaluating, Reflecting

CONCLUSION

In this article, we used the DLCM to map the competences that data literacy projects try to stimulate. In Section 2, we applied the DLCM to initiatives in the field of data literacy, which we found in the literature. We came to the conclusion that the main focus of these initiatives was on using rather than understanding data, a phenomenon inspired mainly by the instrumental framing of data literacy in these initiatives. We clearly took a different approach with DataBuzz, as was seen in Figure 20, which provided an integrated view on the different workshops and projects of DataBuzz. From this mapping it became clear that all of the initiatives in DataBuzz focus on understanding data, as we see this as crucial competencies to develop from an early age onwards. In four of the initiatives, we also focused on at least one competence in the using data cluster, but always in combination with competencies focused on understanding data. Two of our initiatives focused on all competences of the DLCM: Breaking News! (14-16 year olds) and Smart Spaces (16-18 year olds). These are the only two projects that focused on collecting data, which we consider to be one of the most advanced and difficult competences in the model. It is no accident that both projects were designed for older students. We did reflect on the question of whether we should invest in additional efforts to better cover the using data competences in the other workshops as well.

In conclusion, we would like to raise some questions related to the DLCM. First, as with all competence models this model is ideal typical. Not all citizens will achieve high levels of competencies for all sub-competences. It is also not feasible to shift the entire population to high levels of literacy on all the sub-competences. The goal should be to augment the general level of data literacy through educational initiatives. In our view, this means a focus on both clusters of competencies. Second, many scholars who focus on using data presume that being able to handle data will automatically result in better understanding data. We do not make this presumption. Students can be adept at programming algorithms or handling data without understanding their social consequences. Inversely, citizens might be well aware of the role of data and algorithms in society, but lack numerical and statistical skills to independently use data.

An important question is whether the inverse relation is possible. Does a focus on understanding data lead to a higher interest in using data? In other words, can our DataBuzz stimulate an interest in data, statistics, technology, etc. For us, this implies four things. First, the link between these clusters of competences needs to be better understood. Second, we need educational programs that focus on both clusters of competences. Third, these programs need to be developed by a multidisciplinary team, which covers different areas of expertise. Fourth, we need to evaluate the learning outcomes of these trajectories in order to fully understand their impact and contribution to increased data literacy in society. In the course of the next years we hope to find some answers to these questions through the research complementing the DataBuzz.

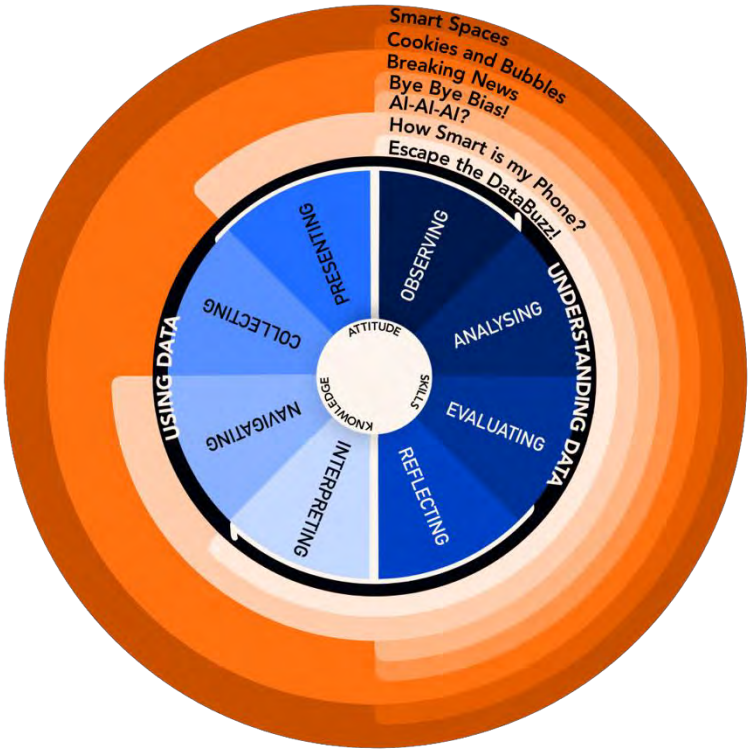

Figure 20. *Workshops of the DataBuzz, mapped on the DLCM*

ACKNOWLEDGEMENTS

We would like to thank the Flemish Community Commission (VGC), the Digital Belgium Skills Fund (DBSF), imec's Smart Education Research (Edulab) and the Vrije Universiteit Brussel (VUB) for financial support in setting up this data literacy initiative.

REFERENCES

Apestaartjaren. (2020). *De digitale leefwereld van jongeren.* https://www.apestaartjaren.be

Argast, A., & Zvyagintseva, L. (2016). Data literacy projects in Canada: Field notes from the Open Data Institute, Toronto node. *The Journal of Community Informatics*, *12*(3), 223-231. http://www.ci-journal.net/index.php/ciej/article/view/1302%0Apapers3://publication/uuid/108219B1-76DE-4AAA-8D04-011F8817501A

Baert, S. (2020). *Wie niet geschoold is, is gezien*. De Tijd. https://www.tijd.be/opinie/algemeen/wie-niet-geschoold-is-is-gezien/10195178.html

Barassi, V. (2019). Datafied citizens in the age of coerced digital participation. *Sociological Research Online*, *24*(3), 414-429. https://doi.org/10.1177/1360780419857734

Bhargava, R., Deahl, E., Letouzé, E., Noonan, A., Sangokoya, D., & Shoup, N. (2015). *Beyond data literacy: Reinventing community engagement and empowerment in the age of data*. Data-Pop Alliance. https://datapopalliance.org/item/beyond-data-literacy-reinventing-community-engagement-and-empowerment-in-the-age-of-data/

Bhargava, R., Kadouaki, R., Bhargava, E., Guilherme, C., & D'Ignazio, C. (2016). Data Murals: Using the arts to build data literacy. *The Journal of Community Informatics*, *12*(3), 197-216. http://ci-journal.org/index.php/ciej/article/view/1276

Borrego, C., Fernández, C., Blanes, I., & Robles, S. (2017). Room escape at class: Escape games activities to facilitate the motivation and learning in computer science. *Journal of Technology and Science Education*, *7*(2), 162-171. https://doi.org/10.3926/jotse.247

Brotcorne, P., & Mariën, I. (2020). *Barometer van de Digitale Inclusie*. Koning Boudewijnstichting. https://www.kbs-frb.be/nl/Activities/Publications/2020/2020_08_24_CF

Deahl, E. S. (2014). Better the data you know: Developing youth Data Literacy in schools and informal learning environments. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2445621

Dierickx, L. (2020). *Artificiële Intelligentie kan detectie borstkanker verbeteren.* De Tijd. https://www.tijd.be/ondernemen/farma-biotech/artificiele-intelligentie-kan-opsporing-borstkanker-verbeteren/10195357.html

European Commission. (2018). *Education and Training Monitor 2018*. https://ec.europa.eu/education/sites/education/files/document-library-docs/volume-1-2018-education-and-training-monitor-country-analysis.pdf

Evers, F. (2020). *AI 'discrimineert' bij het zoeken van een job.* De Morgen. https://www.demorgen.be/nieuws/artificiele-intelligentie-discrimineert-op-de-arbeidsmarkt~b645cdfb/

Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society*, *5*(2). 1-13. https://doi.org/https://doi.org/10.1177/2053951718786316

Grymonprez, S., & Maenhout, K. (2020). *Mainstream onderwijs is dood*. De Standaard. https://www.standaard.be/cnt/dmf20200626_05002940

Ku, K. Y. L., Kong, Q., Song, Y., Deng, L., Kang, Y., & Hu, A. (2019). What predicts adolescents' critical thinking about real-life news? The roles of social media news consumption and news media literacy. *Thinking Skills and Creativity*, *33*(January), 100570. https://doi.org/10.1016/j.tsc.2019.05.004

Markham, A. N. (2020). Taking Data Literacy to the streets: Critical pedagogy in the public sphere. *Qualitative Inquiry*, *26*(2), 227-237. https://doi.org/10.1177/1077800419859024

Mediawijs. (2020). *DataWisdom? Data Literacy? What does it encompass?*. https://en.mediawijs.be/node/6382

Mediawijs. (2020a) Dossier: Het Mediawijs Competentiemodel. https://mediawijs.be/dossiers/dossier-mediawijs-competentiemodel/dossier-mediawijs-competentiemodel-waaraan-werk-je

Mojsilovic, A. (2014). *The age of data and opportunities.* IBM. https://www.ibmbigdatahub.com/blog/age-data-and-opportunities

Oceans of Data Institute. (2016). *Building Global Interest in Data Literacy: A Dialogue*. http://oceansofdata.org/sites/oceansofdata.org/files/ODI%20Data%20Literacy%20Report_0.pdf

Raffaghelli, J. E., & Stewart, B. (2020). Centering complexity in 'educators' data literacy' to support future practices in faculty development: a systematic review of the literature. *Teaching in Higher Education*, *25*(4), 435-455. https://doi.org/10.1080/13562517.2019.1696301

Rea, A., Nagels, C., & Christiaens, J. (2009). Young people in Brussels: social inequality and cultural diversity. *Brussels Studies*, 1-11. https://doi.org/10.4000/brussels.957

Stoilova, M., Livingstone, S., & Nandagiri, R. (2019). *Children's data and privacy online: Growing up in a digital age. Research Findings.* LSE. https://www.lse.ac.uk/my-privacy-uk/Assets/Documents/Childrens-data-and-privacy-online-report-for-web.pdf

Van de Weghe, T. (2020). *BEKIJK: hoe de coronacrisis artificiële intelligentie in een stroomversnelling gebracht heeft*. VRT NWS. https://www.vrt.be/vrtnws/nl/2020/06/04/ai-in-coronatijd/

Van Geyte, J. (2020). *Banken bereiden zich voor op coronagolf aan wanbetalingen: "Tegen het jaareinde komen de echte problemen boven".* Het Laatste Nieuws. https://www.hln.be/economie/banken-bereiden-zich-voor-op-coronagolf-aan-wanbetalingen-tegen-het-jaareinde-komen-de-echte-problemen-boven~a0f49b48/

Vanhecke, N., & Deckmyn, D. (2019, September 21). *Politie moet stoppen met test gezichtsherkenning*. De Standaard. https://www.standaard.be/cnt/dmf20190920_04619293

VGC. (2019). *De evolutie van het Nederlandstalig Onderwijs in Brussel in cijfers*. https://www.vgc.be/sites/vgc/files/2019-06/1819_website_evolutie_NOB_samengevoegd.pdf

Wayens, B., Janssens, R., & Vaesen, J. (2013). Education in Brussels: Complex crisis management. *Brussels Studies*, *70*. https://doi.org/10.4000/brussels.1183

Williams, S., Deahl, E., Rubel, L., & Lim, V. (2014). City Digits: Local Lotto: Developing youth data literacy by investigating the lottery. *Journal of Digital and Media Literacy*, 1-16. http://www.jodml.org/2014/12/15/city-digits-local-lotto-developing-youth-data-literacy-by-investigating-the-lottery/

Wolff, A., Wermelinger, M., & Petre, M. (2019). Exploring design principles for data literacy

activities to support children's inquiries from complex data. *International Journal of Human Computer Studies*, *129*(February 2018), 41-54. https://doi.org/10.1016/j.ijhcs.2019.03.006