Visualising energy: teaching data literacy in schools

Annika Wolff
The Open University
Walton Hall, MK7 6AA, UK
annika.wolff@open.ac.uk

Gerd Kortuem
The Open University
Walton Hall, MK7 6AA, UK
Gerd.kortuem@open.ac.uk

Abstract
As data sets become increasingly complex and pervasive, the importance of citizens to achieve a certain level of data literacy is more important. Citizens need to understand not only how they are contributing data but how this is being used. Data literate citizens have more opportunities for understanding cities through data and informing data driven urban innovations. Current practices around teaching data in schools still focus on using small, personally collected datasets and in teaching graph or chart based visualization. This is a long way away from the types of data and visualisations that are increasingly encountered in daily life. This paper proposes to teach data literacy in schools. Of particular interest in this paper is the idea to engage students with complex data sets to get them thinking about how to produce novel visualisations of this data. Examples are given in which a class of Year 7 and Year 9 students in the U.K. are tasked with creating visualisations of data related to their home energy consumption.

Author Keywords
Data literacy, visualisation

ACM Classification Keywords
H.5 Information Interfaces and Presentation

Introduction

Data literacy is becoming a big topic. The Employer Insights: skills survey 2015 found that 27% of the biggest employers (> 250 employees) now use data analytics to support their work [9]. But businesses also report that they cannot hire the data scientists they need to analyse their growing datasets and ensure their competitiveness [1]. The House of Lords Select Committee on Digital Skills released a report in February 2015 [5] in which they suggest that digital skills, in particular data skills necessary to protect oneself online, are now 'necessary life skills'. They further argue the importance of ensuring that the majority of the population achieve an appropriate level of digital literacy to be full participants of the increasingly digital future. In a similar vein, the emerging topic of human-data interaction (HDI) [3] investigates how the combination of personal data with other sources and lack of transparency in the applied algorithms affects individuals ability to understand how their data is used, or misused. In addition to data owned and used by companies, Open Data initiatives are making data more freely available. A proposed use for open data is to drive bottom-up citizen-led smart city innovation, as well as provide additional resources for commercial enterprise [7][8], though it is hard to find evidence of this data being used outside organised hack events, by participants with above-average data skills. Taken as a whole, this points to a need to raise the level of data literacy within society.

This paper explores the potential for teaching data literacy in schools in order to raise the data literacy of future generations.

Defining and measuring data literacy

Several definitions of data literacy exist. For example, Vahey [10] defines data literacy as "the ability to:

- formulate and answer questions using data as part of evidence-based thinking;
- use appropriate data, tools, and representations to support this thinking;
- interpret information from data;
- develop and evaluate data-based inferences and explanations; and
- use data to solve real problems and communicate their solutions."

According to Carlson et al. [2], data literacy includes "understanding what data mean, including:

- how to read graphs and charts appropriately
- draw correct conclusions from data, and
- recognize when data are being used in misleading or inappropriate ways".

Data Literate students must "be able to access, assess, manipulate, summarize, and present data."

Jeane Harris [4] defines data literacy as "competence in finding, manipulating, managing, and interpreting data, including not just numbers but also text and images."

Implicit within these definitions is often the notion that data is somehow interpreted in context, i.e. connected to real life, or vice versa that real life is reflected through data. Commonly the definitions focus on the interpretation and presentation of data, going beyond the simple skills needed to use appropriate tools, or to clean a data set. Currently, there are no metrics defined for measuring data literacy of a person or in identifying when data literacy skills are being exhibited. But if we take the idea of data literacy relating to this connection of data to the real world, then we can begin to look for evidence of this real-world connection in the

ways that the data is interpreted, visualized and communicated.

Examples from a classroom

The MK:Smart project (www.mksmart.org) is aggregating a number of data sets into a common data hub for developing smart city applications for energy, transport and water. This data is also being used as a resource for teaching data literacy in schools as part of an ongoing initiative to create an online platform called the *Urban Data School* [11].

The focus of the classroom trials is to introduce a data set to the students and to challenge them to firstly interpret the data, then secondly to contextualize the data with their own data collection tasks and thirdly to get creative in designing novel visualisations of big data sets. Currently, the students have worked on paper but future classroom sessions will investigate the possibility to develop tangible visualisations of data to see what effect this has on students creativity and conceptual understanding of the relationship between the data and the visualization.

The data that students have been working with is Smart Meter home energy data. Students first look at a standard line graph showing whole house electricity consumption for one day for a single house. Students are encouraged to ask questions from this snapshot of data before they are given access to the whole data set, at the 'smart' appliance level for four houses across a 1 year period. Students test their questions against the data and continue on a cycle of questioning and answering. Next, students collect their own home energy data by "being a Smart Meter. In the next session, students are tasked with creating a novel visualization of home energy consumption, drawing on the knowledge from the previous session and their homework. The purpose of the visualization is to communicate differences in how appliances use energy and in how much they are used in the home.

Students from two schools, one class of Year 7 (11/12 years old) and one class of Year 9 (13/14 years old) have taken part in these sessions. Students showed good evidence that they understood the relation between energy data and the real world, both through their interesting choice of questions and observations from the data and visualisations that they interpreted and also through the visualisations of data which they produced. Examples of questions were:

1. How much energy is used in a 5 bedroom house
2. How much energy do we use per year?
3. How much energy does the average family use? How often do power cuts happen?
1. Why do people use less energy during the middle of the day?

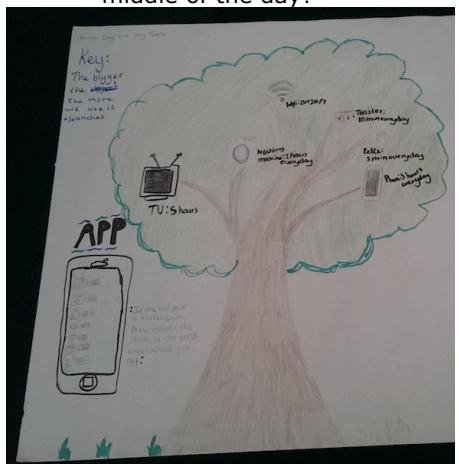

Figure 1. Visualisation of energy data by Year 7 student

Examples of visualizations can be seen in figures 1 and 2. It should be noted that students needed some amount of prompting to start thinking of how to produce visualisations that did not follow the formats they were familiar with, such as drawing a graph or a table.

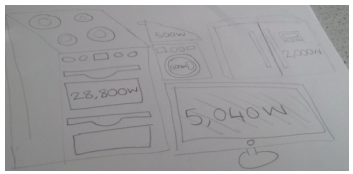

Figure 2. Visualisation of energy data by Year 9 student. Appliances are drawn based on how much power they consume and how often they are used in an average week.

Conclusions and Future work

Students demonstrate some level of data literacy skills when undertaking a task with complex smart meter energy data. Students can ask good questions of the data and can sometimes begin to think about visualisations that go beyond the standard chart or graph that they are commonly taught in school. However, students do need some prompting and support to begin to draw these types of visualisations. What will be of interest in the future is to discover what is needed to help students to get more creative in their visualisations, whether tangible visualisations help students to conceptually bridge the gap between data and the real world. In this study, students interpreted only standard charts and graphs. A further question is whether students are able to easily interpret more complex visualisations of data and if not, what support is needed to help them.

References

1. H. Bakhshi, J. Mateos-Garcia and A. Whitby. 2014. Model Workers: How leading companies are recruiting and managing data talent. Retrieved May 8, 2015 from Nesta: http://www.nesta.org.uk/publications/model-workers-how-leading-companies-are-recruiting-and-managing-data-talent
2. Jake Carlson, Michael Fosmire, Chris Miller and Megan Sapp Nelson. 2011. Determining Data Information Literacy Needs: A Study of Students and Research Faculty. Libraries Faculty and Staff Scholarship and Research. Paper 23.
3. H. Haddadi, H. Howard, A. Chaudhry, J. Crowcroft, A. Madhavapeddy and R. Mortier, 2015. Personal Data: Thinking Inside the Box. Accepted paper at *Critical Alternatives 2015*, Aarhus, Denmark
4. Jeanne Harris. 2013. Data Is Useless Without the Skills to Analyze It. *Harvard Business Review*.
5. House of Lords Select Committee on Digital Skills. 2015. Make or Break: The UK's Digital Future. Retrieved May 8, 2015 from: http://www.publications.parliament.uk/pa/ld201415/ldselect/lddigital/111/111.pdf *Accessed on 8th May 2014*
6. H. Maycotte. 2014. Data literacy - What it is and why none of us have it. Retrieved May 8, 2015 from Forbes: http://www.forbes.com/sites/homaycotte/2014/10/28/data-literacy-what-it-is-and-why-none-of-us-have-it/
7. D. McAuley, H. Rahemtulla, J. Goulding and C. Souch. 2014. How Open Data, data literacy and Linked Data will revolutionise higher education. Retrieved May 8, 2015 from Pearson blue skies: http://pearsonblueskies.com/2011/how-open-data-data-literacy-and-linked-data-will-revolutionise-higher-education/
8. A. Ojo, E. Curry and F. Zeleti. 2015. A Tale of Open Data Innovations in Five Smart Cities. *International Conference on System Sciences*, Hawaii
9. Tech Partnership. *Employer insights: skills survey 2015*. Retrieved May 12, 2015 from: https://www.thetechpartnership.com/globalassets/pdfs/tec_employer_skill_survey_web.pdf
10. P. Vahey, L. Yarnall, C. Patton, D. Zalles and K. Swan, K. 2006. Mathematizing Middle School: Results From a Cross-Disciplinary Study of Data Literacy. *American Educational Research Association Annual Conference*, San Francisco, CA
11. A. Wolff and G. Kortuem. 2015. Urban Data in the primary classroom: bringing data literacy to the UK curriculum. Data Literacy Workshop at ACM Web Science Conference 2015, Oxford

Author Bio

Annika Wolff is a researcher whose interests lie at the intersection between big data, machine and human learning. She is currently working within a smart city project to use complex urban data sets for teaching data literacy in schools. She is also interested in using urban data for revealing cultural narratives across a city. Other research interests include learning analytics and data visualisations for tutors and learners and the use of games and narratives to motivate learning.