How to gain knowledge when data are shared?
Open Government Data from a media pedagogical perspective (draft version)

Valentin Dander
University of Innsbruck/Austria
valentin.dander@uibk.ac.at

Abstract

Numeral and big data are becoming one of the dominant codes of describing society. Public services are sharing Open Government Data (OGD) as public goods. Academic interest so far focused on political, technical and organizational implications. Educational research has been widely neglecting OGD. I argue that contemporary media pedagogy needs to consider this development in research and practice, engaging with the question how data can be turned into knowledge. Which objections emerge and which pre-conditions and capabilities are required by data users? This paper illustrates and discusses potentials of OGD users in terms of learning and subject transformation in order to draw consequences on a conceptual level. Various objectives for media pedagogy – such as media, digital and data literacy, numeracy and picturacy – will be discussed.

[key words: Open Government Data, OGD, media pedagogy, education, digital and media literacy]

1. Shared Datasets

Technological developments in recent years and decades seem to largely follow Gordon E. Moore’s law, in which he (1965) declared the bi-annual doubling of transistors on integrated circuits. This exponential increase similarly applies to digital storage capacities. Accompanied by a constant growth of ‘big data’ we’ve come to a point where the Wired magazine, in 2008, can title its special feature “The Petabyte Age: Because More Isn’t Just More — More Is Different”¹. Until now, big data remain in the hands of state, business and academic users, since using them meaningfully requires specific hardware, software and adequate skills (Manovich 2011, p.1). But data are not only shared and processed in terms of big data. Digital sharing cultures are blossoming in general. Volker Grassmuck (2012) names this trend the “Sharing Turn”, in the same breath stating that sharing can be understood as an anthropological constant of human existence. Not only are we talking about private users who are more or less intentionally leaving digital footprints and exchanging thoughts, ideas and knowledge. Since 2009 – and starting in the US and UK – more and more governments and diverse organizations are following the call by the open (government) data (OGD) movement to open

their data storages. Several democratic promises accompany this openness and are gratefully picked up by governments, such as Obama administration: “We will work together to [...] establish a system of transparency, public participation, and collaboration. Openness will strengthen our democracy.”\(^2\)

One might ask: Who are ‘we’? And what kind of ‘work’ are ‘we’ to do?

The OGD concept is implicitly thought of as a three step procedure: At first, data are shared according to principles, e.g. accessibility, free licensing, machine-readability (Sunlight Foundation 2010). In a second step, data are refined by professionals (mining, linking, programming applications, analyzing, visualizing) transformed into in-form-ation until non-expert users can start to work with it. There appears to be a divide between this mediated access to information and any idea of immediate civil participation:

> “The Internet is the public space of the modern world, and through it governments now have the opportunity to better understand the needs of their citizens and citizens may participate more fully in their government. [...] Open data promotes increased civil discourse, improved public welfare, and a more efficient use of public resources” (Open Government Working Group 2007).

We can conclude for now that data of different scale become one of the dominant codes to describe society and therefore acquire an increasingly important role in “public space” – or even more so within the public sphere. Civil society is invited to actively participate in governance by analyzing and querying this shared data. We might a skeptical remark: Not only data are shared. All citizens – by part-time ‘working’ for their governments – equally receive their accountability share. Referring to the Obama quote above, we might call this crowd sourcing governance.

### 2. Objections to OGD

Numerous objections have been expressed towards OGD. I argue that we have to take them into account before we can continue with propositions a ‘data-education’. The leading proponents of OGD seem to be widely neglecting the possible side-effects of this project. Conflicts of interest might also arise from their professional background. While democratic intentions are positioned in the foreground, this lobbyism can to some extent be understood as extending significance and value of digital industries. Hence, OGD can not only be understood as a political project. Neelie Kroes, Vice President of the European Commission, responsible for the Digital Agenda in Europe, blogged about the enormous economic potential of OGD:

> “[A]ll together, public sector information generates over 30 billion euros per year in economic activity, with services from geo-location services to weather forecasts. By opening up this resource fully, we could more than double the value of this activity – to around €70 billion. This opening up can generate tax revenues which far exceed revenue from any fees previously charged for the data” (2011; emphasis in original).

An optimistic view might regard these expectations as another argument for a winning game on all sides. The counter-argument however, has to ask which power structures arise from this. Which data are to be published in which form? Who will benefit (the most)? And which effects on governance will be visible? In consequence the basic principle of sharing can be queried, as “sharing is an act that takes place among equals – power relations cancel the spirit of sharing, which is based on free will

and generosity” (Sarikakis 2012, p.37). This complex arrangement of motivation complicates the analysis and leads simple approaches astray.

Longo (2011) adds another perspective by naming some indications that OGD functions as a “trojan horse” for New Public Management strategies by letting citizens participate as informers working on specific quests: comparing educational institutions, mapping crime scenes, finding ‘overpaid’ cultural projects. Public services are monitored by the public and held accountable, while the government can easily retreat into a moderating position. Additionally, conducting all this work is not considered as ‘work’, but ‘participation’ or ‘honorary work’. Sarikakis’ (2012, p.38) writing on economic structures in social network sites matches these thoughts:

“The deprivation of privacy of individuals strengthens the public monetary position of the industry. Again, this contradictory element of invisibility of private/public through the publicness of the user, is accompanied by the invisibility of labour/leisure […] that articulate[s] vast amounts of un-remunerated labour hours.”

Users within the OGD apparatus are finding themselves in a twofold position: enhancing participatory democracy on the one hand and creating value for undefined beneficiaries on the other. This challenge of the public/private distinction seems to go along with the “invisibility of labour/leisure” as Sarikakis puts it.

Also the OGD condition of mutual and self-observation possibly causes a normalizing effect on individuals and institutions, as Johnson (2013, p.9) argues, referring to the Integrated Postsecondary Educational Data System (IPEDS) in the United States. While Johnson introduces Foucault’s concept of a disciplinary society to describe those effects, I propose to recur to ‘governmentality’ as an analytical category to understand the phenomenon:

“The theory of the art of government was […] connected to a set of analyses and forms of knowledge which began to develop in the late sixteenth century and grew in importance during the seventeenth, and which were essentially to do with knowledge of the state, in all its different elements, dimensions and factors of power, questions which were termed precisely ‘statistics’, meaning the science of the state” (Foucault 1991, p.96).

While in the early times of governmentality these governing techniques remained limited to professionals (scientists, administrators), contemporary strategies blur the competences and responsibilities within political processes. Intensified by digital access, information on societal conditions is mirrored to the general public and necessarily provokes consequences in knowledge structures and behavior.

“[I]f the state is what it is today, this is so precisely thanks to this governmentality, which is at once internal and external to the state, since it is the tactics of government which make possible the continual definition and redefinition of what is within the competence of the state and what is not, the public versus the private” (Foucault 1991, p.103).

Quantitative, statistically processed representations of ‘facts’ predominantly operate with bell-shaped curves and hence bring forward a normalism tendency. In his study A Trial on Normalism (1997, p.452ff) Link elaborates on the subjectivizing forces of “datizing cultures” and their ensuing

---

4 Engemann (2013) together with Traue proposed the catchy term “gouvernemediality” in media context.
“curve scapes” on the three dimensions of signal, orientation and control. Thereby deviant behavior becomes less probable, while the majority is ‘pushing to the middle’. The equivalent on a state level can be found in benchmarking as a controlling technique. Not surprisingly this will be discussed as one of the main issues at the OGD conference ‘Efficient State’ 2013 in Berlin.  

Several further objections can be named. Be it the embedded injustice in rather ‘half-baked’ than raw data (Johnson 2013) or the legitimate question “who is in a position to make ‘effective use’ of this newly available data?” (Gurstein 2010b). Isn’t O(G)D just exacerbating the digital divide? Furthermore, does opening data necessarily lead to a more open government? Yu and Robinson (2012) negate and suggest keeping the two categories apart. Also, it makes a difference to wait for governments to share certain data or to have the right to demand them. While in Sweden, the United Kingdom, the United States of America and many other countries Freedom of Information is regarded as a basic democratic principle, in Austria it legally collides with an overruling, constitutional Obligation of Discretion (Austrian Constitution). A realistic view on OGD shows that data per se are no legitimate cause for euphoria. To handle all these challenges well,

“we need to learn how to read and interpret them critically, to read between the lines, to notice what is absent or omitted, to understand the gravity and implications of different figures, and so on. We should not imagine that anyone can easily understand any dataset” (Gray 2012).

Democracy to some extent requires participation beyond voting. Participation in its original sense means to ‘take or capture a part’. If this part does not mainly consist of work, but also of power to co-shape communities and societies, it can be worth to accept the ‘share’.

3. Examples for meaningful use?

Different approaches, aiming at meaningful use of OGD, are already being tested. I will sketch four short examples and then draw conclusions for media pedagogy on a conceptual level.

In June 2012 a W3C Workshop on Using Open Data took place in Brussels. One group of researchers and software developers from Dresden, Germany, presented a concept for OPEN – The Open Data Processing Engine, an enhanced Database Management System (DBMS) to enable non-expert users to identify, extract, integrate and analyze datasets. It is planned to serve as a single access point and tool for democratization of data use. Users will be invited to contribute (Braunschweig et al. 2012). Simplifying infrastructures like OPEN can indeed be helpful in educational settings. Still, even operating this instrument remains demanding to non-expert users. Also, the authors regrettably do not comment on how the DBMS will be licensed. It would be eligible to steadily extend the pool of data processing software under Free Software Definition.

---


7 Find the report on [http://www.w3.org/2012/06/pmod/report](http://www.w3.org/2012/06/pmod/report) [Accessed 2013-04-24].

A second example is the UK Data Service, “a comprehensive resource funded by the ESRC [Economic and Social Research Council; V.D.] to support researchers, teachers and policymakers who depend on high-quality social and economic data”. It is providing advice, online as well as face-to-face training and courses, guides for learning and teaching and a “wide range of secondary data” (University of Essex & University of Manchester 2012). For full access to the available datasets one must register and the accessible data are pre-processed. About 30 case studies indicate that the service is in use by university and college teachers at under- and post-graduate level. UK Data Service is using proprietary software (nesstar) for online publishing and analysis of statistical information.\(^9\)

In terms of ‘openness’ the most consequent service is the School of Data by the Open Knowledge Foundation. It proclaims to “empower civil society organizations, journalists and citizens with the skills they need to use data effectively in their efforts to create fairer and more sustainable societies” (Open Knowledge Foundation & Peer 2 Peer University 2013; emphasis in original). Various learning materials are offered online, such as a handbook, tutorials and everybody is invited to join so-called explorer missions. The latter are set up as self-organized, informal MOOCs (Massive Open Online Courses) including narrative gaming elements. As no official certification is offered, the School of Data is integrating an OpenBadges system.\(^10\) All content displayed is licensed under a Creative Commons Attribution-ShareAlike v3.0 license.

A brief pilot study I’ve conducted with three male secondary school students at the age of 15, 17 and 18 years, brought up some complementary aspects. They had about one hour of time to stroll through the Tyrolean OGD portal, whilst thinking aloud and discussing among them. I only provided the most basic information: What is an OGD portal and what can be found there? First of all, they showed and uttered great interest in the numeric data – especially those being related to their home towns. It was stated that they could be used for political argumentation, e.g. against xenophobic positions. They wouldn’t query the data for being ‘raw’ or constructed in order to fulfill certain institutional needs. While they had no problems to navigate on the portal and read the spreadsheets, they complained about a lack of human readability. Finally they started reflecting on their political education courses in school, going along with history classes in the 11th and 12th grade. The older students complained about content being too abstract which could be replaced by or intertwined with these concrete data sources.

Now, what can we learn from these four examples? Especially for an educational context it is crucial to have free or at least affordable software at hand to process data – and to find a compromise between necessary complexity and barrier-free simplicity. This can be achieved with a holistic approach as proposed by the OPEN-team, uniting all required functions within one single application, which might reduce constraints at first sight. But it also possibly reduces the user’s opportunities of looking behind the scenes and understanding how the single jigsaw pieces fit together. This largely depends on how transparently the structure itself is built. The UK Data Service chose an academic audience and restricts freedom of action to particular steps of data refinement. Secondary data is adequate to learn about quantitative methods in social science in institutional settings, working on ‘real data’. Still, therefore OGD is not needed and it won’t counteract the above mentioned objections. Registration and proprietary software pose additional deterrence to interested users. If using OGD for formal learning scenarios it might be considered to apply it to various subjects. Apart


from acquiring instrumental competences in statistics, OGD could enrich political education in schools with relevant, contemporary data on e.g. the community. The School of Data is free from institutional restrictions and can focus its activities and learning objectives on open data. This informal, mutual ‘private’ learning space seems to be more publicly relevant than the other. Who is using this service remains open. Even if everybody is invited, we might expect to find expert and semi-expert users on SoD-data expeditions.

If aiming at non-experts on a broad community basis, we need to step back further. Gurstein (2010a) lists seven “elements that are required to be in place on the end user side for the effective use of open data to take place.” I will quote those elements in a shortened version:

1. **Internet access** sufficient to support making the data available; barrier-free;

2. **Computers and software** sufficiently powerful, having sufficient time;

3. **Computer/software skills** to use the soft- and hardware;

4. **Content and formatting** – having the data available in a format such as to allow for effective use at a variety of levels of linguistic and computer literacy;

5. **Interpretation/Sense making** sufficient knowledge and skill to see what data uses make sense (and which don’t) and to add local value;

6. **Advocacy** – having supportive individual or community resources sufficient for translating data into activities for local benefit;

7. **Governance** – the required financing, legal, regulatory or policy regime, required to enable the use to which the data would be put;

To complicate this plain list, some more aspects can be added: knowing about open data and motivation to engage with it; context information on open data sets and sources (who collected it how, when, where and for which purpose?); online- or offline-communities to exchange ideas and skills. Gurstein (2010a) asks “What are the language, computer literacy, data analytic literacy levels that are required for an effective use of the ‘open data’?” But before we even try to aswer this question we have to ask which capabilities, competences and skills are required for an effective use of O(G)D.

---

4. **Conceptual remarks**

If media pedagogy wants to pick up the open data issue, the objectives it wants to envisage must be discussed. The variety of definitions for literacies, competences and education models exceeds this paper. Some of the most dominant threads shall be sketched, largely following Hug’s (2012) argumentation. The most established approach in anglophone literature and classrooms certainly is the media literacy concept which Hobbs (2011, p.12ff) enhanced to “digital and media literacy” and defines by five “elements”: access (using, finding, comprehending), analyze (critical thinking), communicate (expression), reflect (social responsibility) and act (ethical democratic citizenship). This

---

11 Some of those remarks are drawn from the *five stars of open data engagement* that have been worked out by various contributors at the UKGovCamp 2012 to address non-technical OGD principles (Davies 2012).
definition shows slight differences to the media literacy definition the NAMLE (National Association for Media Literacy Education)\(^\text{12}\) offers. The NAMLE sees media education on a predominantly knowledge based level, while media literacy merely refers to practical abilities. The European Charter for Media Literacy\(^\text{13}\) only adds concern on harmful content. As NAMLE is calling for a “wider set of literacy skills” an obvious confusion of terms arises. What do literacy and skills mean in this context? Kress (2004, p.21ff) critically addresses misleading compounds including ‘literacy’ – and there have been several within recent years: visual literacy, information or computer literacy, critical or political literacy, statistical literacy and so on. Data analytic literacy, as mentioned above by Gurstein, and data literacy are just two more in a long list.

Two aspects appear problematic at this point. Until now, there has not been any widely accepted definition for data literacy (if one exists at all). This makes it hard to answer the questions, why data literacy should be separated from media literacy and which particular components it contains if it does not contain those being named above by Hobbs. Then, secondly, if we take Kress’ critical remarks on literacies into account, we ought to think of more fundamental skills and abilities to describe what data use requires: literacy in its literal meaning, numeracy and signing. To stress the powerful democratic and socio-political character of mathematics, Skovsmose (1998) elaborated mathemacy. This is going beyond more basic mathematical operations that are being addressed in numeracy. ‘Signing’ might be misleading, compared to alternatives like Heffernan’s picturacy (2002). In turn, his concept of picturacy focuses on comprehension of photography and paintings. “Visuacy” (Hug 2012) can avoid such restrictions and include the latter as well as matters of design, graphics and visualizations.\(^\text{14}\) Yet, considering complex and interconnected datascapes, all these capabilities are needed to effectively understand all steps of accessing, understanding, processing, displaying, encoding and decoding. These multimodal conditions can not only be found in this context, but generally are fundamental for digital environments (Kress 2004). Here, we have to question the value of categories like literacy, mathemacy or visuacy beyond an analytical one. Regarding the interdependency of these capabilities, introducing another cross-sectional subject to school would hardly lead to groundbreaking changes. The German term ‘Bildung’ refers to an even more essential and abstract, subject-oriented category. Medienbildung, designates the process of transforming relations to one self and the surrounding through mediatic experience, aiming at orientation in its widest sense (Koller et al. 2007; Jörissen & Marotzki 2009). Hence it can be helpful as an umbrella term.

Although the School of Data is labeled ‘school’, it in fact scrutinizes educational institutions in offering informal, project-based ‘data expeditions’ in disrespect of the classic learner/teacher distinction.

“As to institutionalized education, large parts of it can be taken as examples demonstrating how much literacy-based forms of the communicative stabilization of learning cultures can restrict the probing of creative, conceptual and critical-reflexive scopes” (Hug 2012, p.123).

These remarks raise the question whether the institutionalization of data-education is desirable at all. Foucault’s governmentality concept sharpens our perspective on power structures within governmental conditions. On the one hand this calls for response to and reflection on governmental practices and apparatuses which can be situated in formalized educational contexts, as well. On the

---


\(^\text{14}\) The visual display of quantitative information (Tufte 2006) confronts us with the problem of representation in a different way than e.g. realist art or war photography.
other hand the role of teachers within governmental dispositives has to be dealt with carefully. An educational conception towards ‘governing students not to be governed (that much)’ within formal, obligatory education can too easily act out what it pretends to counteract. Informal settings however bear the risk to fortify social injustice and privilege – if largely used by well-educated citizens and semi-experts, as assumed. A compromise between public and privatized approaches might be opened with local community projects (Gurstein 2010b). This way, its educational character goes along with its concrete political perspective.

5. Conclusions

Online sharing practices result in immense quantities of data. Various public services are continuously sharing their data storages as open government data. In consequence quantitative data becomes increasingly significant as a representative code in digital environments.

While O(G)D proponents stress the potentials of OGD as a means to expanding transparency and participation, numerous objections have been expressed. OGD can be read as a political as well as an economic endeavor. A political economy perspective allows revealing un-remunerated labor behind participation. This double bind applies to its advocates and users as well as to those releasing data. Hence, OGD can be understood as an amplifier of New Public Management in public services. Thereby citizens can be established as controllers and quality managers, while undergoing a normalizing shift that emerges from ‘curve scapes’ and comparison. Hence, participation may be understood as self-management and self-governance under the conditions of governmentality and governmediality.

Media pedagogy needs to consider these complex and ambiguous dynamics and find adequate response in a theoretical and practical dimension. ‘Literacy’ compounds are misleading, as they either remain attached to a linguistic perspective or blur the denotation of ‘literacy’. Effective use of OGD requires more than linguistic symbols. Numbers and visual forms are crucial codes in equal measure within multimodal digital and even more so in data environments. They therefore demand specific skills. Literacy, mathemacy and visuacy can serve as core competences for handling data, while Medienbildung is proposed to embrace the others towards a process- and subject-oriented understanding of ‘learning’. While single skills might be taught in public institutions, their interdependence and the political implications of OGD challenge formal education. Meanwhile informal learning projects are experimenting with community-based, self-organized learning arrangements. Here, empirical research is needed to better understand who is learning or teaching what through OGD in which context and for which purpose.
References


Grassmuck, V., 2012. The Sharing Turn: Why we are generally nice and have a good chance to cooperate our way out of the mess we have gotten ourselves into. In W. Sützl et al., eds. *Media, knowledge and education: cultures and ethics of sharing*. Innsbruck: Innsbruck University Press, pp. 17–34.


University of Essex & University of Manchester, 2012. UK Data Service. UK Data Service. Available at: http://ukdataservice.ac.uk/ [Accessed April 24, 2013].