



PROJECT MUSE®

019 Data Literacy

Published by

Walker, B. and Tim Davies.

The State of Open Data: Histories and Horizons.

African Minds, 2019.

Project MUSE. <https://muse.jhu.edu/book/67112>.



➔ For additional information about this book

<https://muse.jhu.edu/book/67112>



This work is licensed under a Creative Commons Attribution 4.0 International License.

[172.69.6.66] Project MUSE (2025-04-04 23:16 GMT)

019

Data literacy

Mariel Garcia Montes and Dirk Slater

Key points

- The release of open data has outpaced open data use. A key cause has been low levels of data literacy among individuals and organisations who could benefit from open data.
- There is a lack of evidence on data literacy baselines, a dearth of systemic interventions to build data literacy, and limited research on what works. Existing initiatives face funding or organisational challenges in working out how to scale an effective training response.
- When organisations have access to a range of data skills, they can be more effective in making use of open data. Understanding “data as a team sport” can help move from a perspective on individual data literacy to an understanding of organisational data maturity.
- Realising the full potential of open data will require continued work on building wide-ranging capacities to create, understand, analyse, and use data.

Introduction

One of the central premises of open data for development is the potential for transformation that can be realised when people use open data. Open data proponents have often made assumptions about the interest and abilities of citizens as potential end users; however, programmes to increase the amount of open data available rarely contemplate the resources needed to promote the use of that data once released. Accounting for the gap between data release and actual use is one of the goals of data literacy advocates.

What is (open) data literacy?

Contemporary academic literature puts forward many literacies needed for the modern world (numeracy, information literacy, statistical literacy, technology literacy, etc.), yet, until recently, data literacy was primarily discussed in the context of the skills needed by students and researchers. Frank et al. (2016) argue that the rise of the web and, in particular, open data, has helped to change this, putting data literacy on the agenda for a far more extensive range of organisations and individuals.¹ They also note that many different definitions of data literacy have been proposed.

School of Data,² an international network of individuals and organisations promoting skills for the effective use of data, interviewed civil society practitioners in 2016³ about their understanding of data literacy. The research focused mainly on the abilities they strive to develop among their audiences through their work (such as being able to use data to advance one's goals or to know how to find information). Practitioners expressed resistance to descriptions of data literacy they felt were widespread but of limited use. They argued that data literacy should not be seen as the result of a linear learning process nor as a binary in which one is either illiterate or literate. In addition, data literacy should not be promoted solely as an individual capacity without accounting for groups and communities.

The Data-Pop Alliance⁴ has proposed that data literacy as a term does not adequately account for the need to adapt and update our understanding of the role of (open) data in the modern world, particularly as machine learning and other technologies change the use of data. They suggest that we should not refer to a sub-type of literacy (i.e. data literacy) but rather to the broader idea of “literacy in the age of data”.⁵ Gray et al.⁶ build on this idea by introducing the concept of “data infrastructure literacy”, calling for efforts that “make space for collective inquiry, experimentation, imagination and intervention around data in educational programmes and beyond, including how data infrastructures can be challenged, contested, reshaped and repurposed to align with interests and publics other than those originally intended”. Open data literacy is not only about working with data. It also involves strategic efforts to shape the data environment within which we work.

While acknowledging the ongoing discussions around how best to define or conceptualise data literacy, this chapter considers “data literacy” primarily from the perspective of the pedagogies and organisational processes used to promote and acquire the skills needed to utilise open data. For the purposes of this chapter, we start from a working definition from Bhargava and D'Ignazio,⁷ who describe data literacy as “the ability to read, work with, analyze, and argue with data”.

Matthews⁸ offers a catalogue of core competencies (see Figure 1) discussed across different sources on data literacy, highlighting the mix of capacities required related to data consumption, data creation, and data ethics, as well as a broad range of common competencies needed to manage data. Work on “open” data literacy has tended to focus more specifically on skills relating to data discovery and acquisition, data cleaning, and working with data in open data formats (CSV, JSON, etc.), using a particular set of common open data tools, as well as data visualisation, presentation, and storytelling. There is a further challenge for open data literacy in that the landscape of data availability, digital tools, analytical methods, and linkability between datasets

is rapidly evolving. Moreover, understanding the significance of the term “open” when applied to data literacy implies an additional set of competencies and skills that focus on interoperability. Work on open data literacy in general aims to promote an understanding of, and commitment to, openness.

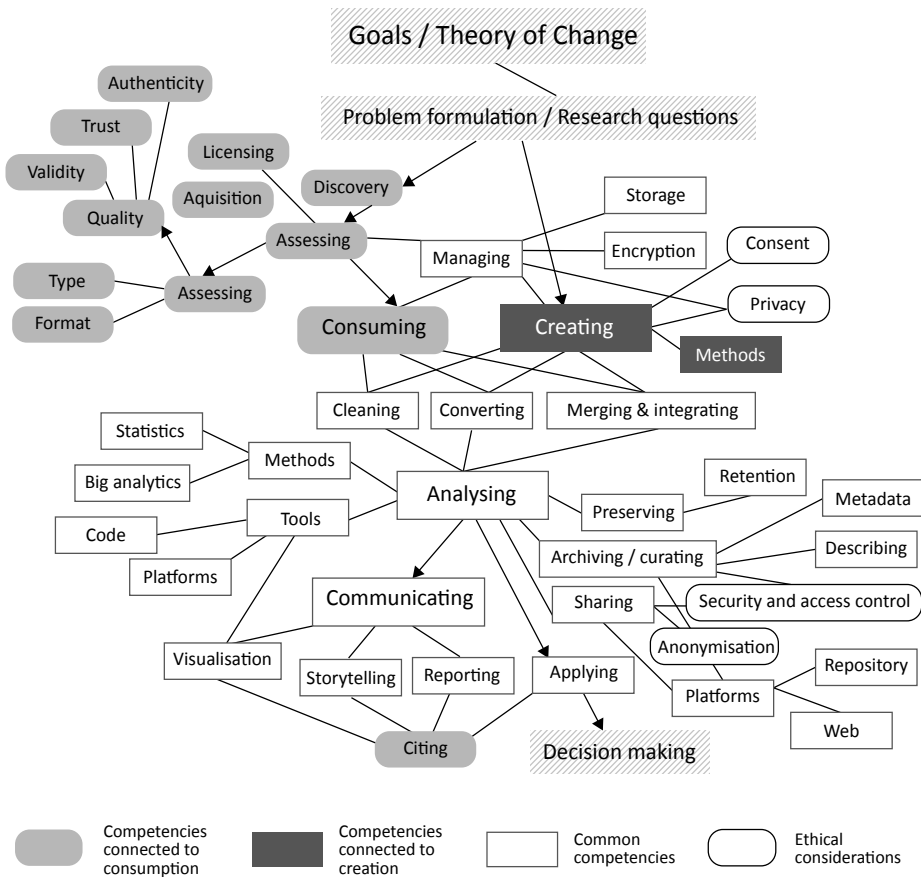


Figure 1: Generalised conceptual model of data literacy activities. Starting point Koltay (2015)⁹. Source: <http://ci-journal.org/index.php/ciej/article/view/1348/1222>

As we will see in the following sections, while initiatives focused on the broad topic of data literacy have developed their practice over the last decade, a lack of definitional clarity has, at times, frustrated efforts to establish baselines and to measure progress, and limited the development of an evidence base on what works.

Building open data literacy around the world

Over the last decade, as attention has shifted from just securing the release of “raw data” to the actual use of data, a variety of initiatives focused on spreading skills and methodologies related to data use have been implemented. Civil society organisations, as well as governments and international organisations, have promoted several data literacy capacity-building models, ranging from short-term efforts, such as workshops, community events, and datathons, to medium-term efforts, such as multi-week “lab” models and training programmes, and further, to long-term initiatives, such as fellowship programmes, research initiatives, and other longer-term projects.¹⁰

Data is a team sport

In 2017, the “Data is a team sport” podcast broadcast a series of conversations with data literacy practitioners to capture lessons learned and examine how methodologies are shifting and adapting in the wake of an ever-evolving data literacy ecosystem. As political and economic forces become more adept at using data, we enter a new era of data literacy where just being able to understand information is not enough.¹¹

Organisations rely on different frameworks and pedagogies for their short-, medium-, and long-term efforts. For the School of Data, working with real data is central to the “data expedition” learning model. They note that on an expedition, learners may “go in circles, get lost, make mistakes, and sometimes, not reach [their] goal. But that’s fine! This is the best way to get familiar with working with data.”¹² Data expeditions are structured along a data pipeline, an approach focused on “working with data from beginning to end.”¹³ Moreover, participants adopt different roles based on their skills to work in teams that work through the pipeline.

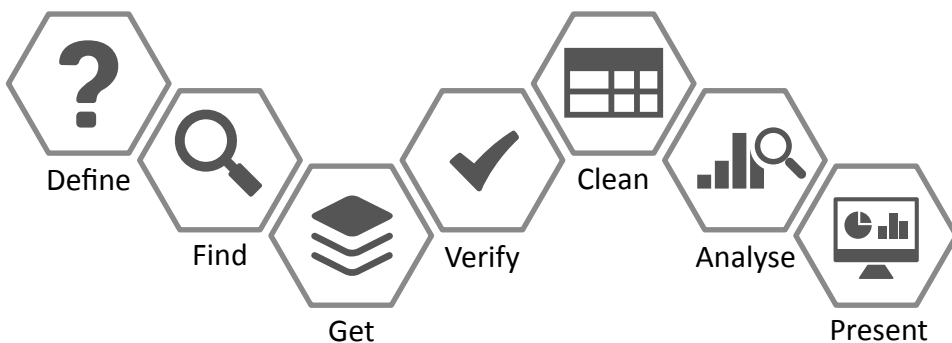


Figure 2: The School of Data’s data pipeline
Source: <https://schoolofdata.org/methodology/>

The recognition that much real-world open data is low quality, and that working with it may require quite advanced skills, has fed into the design of Global Integrity’s “Open Data Treasure Hunt” approach, which is also informed by the data expedition model. In an open data treasure hunt, participants are encouraged to assess the quality and integrity of data and identify how to put it to use. The hunt establishes realistic expectations of data use and identifies data issues that can be sent to the data creators and curators to make the data more accessible to work with in the future.¹⁴

The Open Data Institute (ODI), founded in London in 2012 and now distributed worldwide through a franchised “nodes” model, has emphasised creating training oriented toward organisations and governments on creation and publication, as well as the use of data. The ODI training menu has expanded from courses initiated in 2013, such as “Open Data in a Day” focused on basic concepts, including data formats and licences, and the three-day “Open Data in Practice” (2013) course that covers data publication, discovery, and business models, through to the “Open Data Science” course launched in 2016 and “Introduction to Data Ethics” launched in 2018. ODI courses have reached over 17 000 people, including approximately 2 000 attending in-person public courses, 5 500 via e-learning, and a further 4 750 face-to-face and 5 000 online through in-house courses.¹⁵

The School of Data and ODI have also emphasised a “Training the Trainers” model, including School of Data fellowships, as well as specific work in Tanzania that sought to develop in-country capacity both to understand open data and to deliver practical training to others.¹⁶ This model relates to curriculum development for open data educators, an area where the work of the Investigative Reporters and Editors Network is notable.¹⁷

Building open data capacity in Tanzania¹⁸

ODI and the World Bank have collaborated with the Tanzania Open Data Initiative to improve data publication and use in the country. Through in-person visits to Tanzania, they have trained 222 people and reached another 127. Their efforts have also resulted in four e-learning modules.

DataBasic¹⁹ is an online learning platform created by Rahul Bhargava and Catherine D’Ignazio, data literacy researchers and facilitators at the Center for Civic Media at the MIT Media Lab. Their scan of existing online data tools led them to the conclusion that available tools were “designed for users, not learners, and privilege the production of quick visuals at the expense of supporting a learning process”. Instead, they set out to design tools that respond to pedagogical goals, rather than being designed around particular data outputs. Their designs draw on Paulo Freire’s empowerment-focused pedagogy with its emphasis on critical thinking and consciousness, creating tools that are “focused, guided, inviting, and expandable”.²⁰

It is notable that, in absolute terms, these generalist data literacy-building efforts have reached a relatively small audience (in the tens rather than hundreds of thousands of people trained). Market research by data analysis firm Qlik in 2017 and 2018, looking at data skills within businesses, put levels of data literacy in Europe at just 17% of workers²¹ and in Asia²² and America at around 20%²³ with the majority of workers surveyed interested in improving their skills. While

the methodology for these studies is unpublished and the results should be approached critically, they are indicative of a significant gap between the potential to use data and the skills to use it. This raises important questions about what it would take to fill these skills gaps. One promising approach has been to work with open data literacy as a capacity of communities and organisations rather than one of individuals.

DataKind UK²⁴ and Data Orchard,²⁵ charities based in the United Kingdom that advocate for the use of data and research for good, have championed an approach that looks at the use of data well beyond individual data literacy. Aimed at the organisational level, “data maturity” is defined by the ability to manage data-driven projects. This approach to achieving organisational data literacy recognises that workshops or training will succeed in raising awareness, but do not necessarily support the deeper engagement or skills building needed for an organisation to achieve self-sufficiency in running data projects. To help organisations better understand their progression toward data maturity, DataKind UK and Data Orchard have developed a framework that looks at seven key themes and plots out a five-stage journey to reach data maturity,²⁶ while also matching organisations with mentors who can provide support in implementing data projects.

Also aimed at the organisational level, projects like Internews’ fellowships, which have been embedded in newsrooms for several weeks in places like Palestine²⁷ and Kenya,²⁸ offer a longer-term model for building capacity by changing the processes of organisations to promote the use of data in everyday work. Another initiative using the fellowship model to promote data literacy within a team, the Caribbean Open Institute fellowship in Ayitic, aimed to help young women to find employment in the digital economy in Haiti.²⁹

Three other responses to the practical challenge of building data literacy are worth noting. First, a focus on infomediaries, actors who analyse and create outputs based on open data to make it more accessible and useful for end-users.³⁰ Data journalists, in particular, have been playing this role in the open data and data literacy ecosystem by finding and presenting relevant data to citizens as in stories like the Panama Papers,³¹ which used vast quantities of data and had a global impact, as well as more local efforts that seek to uncover facts and explore causes of societal problems.³² Networks like Exposing the Invisible³³ and Visualizing Impact³⁴ have included data literacy components in their work to build capacity for investigative journalism.

Second, we have seen other initiatives that seek to embed data literacy components within wider programmes. For example, Publish What You Pay (PWYP) ran a “Data Extractors”³⁵ programme that provided support for individuals embedded in PWYP chapters to “dig for data” and find new ways to use it. The International Federation of the Red Cross (IFRC) has developed an in-house data literacy program for staff called “The Data Playbook” that is a collection of “social” learning exercises meant to drive a deeper understanding of the role of data in supporting effective humanitarian responses.³⁶

Third, additional initiatives have been developed to address ethical considerations around the use of data. The Responsible Data Forum³⁷ is an initiative stewarded by The Engine Room that works to “identify the unintended consequences of using data in this kind of work, and bring people together to create solutions”. The Responsible Data Forum has explored the implications of the publication and use of data in the fields of public health, violence, and humanitarian response.

What works? Where are the gaps?

The critical question then is whether all these interventions are helping to close the gap between the potential benefits of open data and their realisation. Although funders have made significant investments in data literacy-related interventions, often as part of other programmes, the empirical evidence on what works when it comes to open data literacy is surprisingly scant. We were not able to locate any published pre- and post-intervention data from data literacy-building programmes, or established measurement tools to monitor progress. While the literature describing pedagogic principles suggests some context-specific best practices and highlights challenges in improving data literacy, the lack of a coherent and generally accepted definition of data literacy and requisite skill set leaves us without a real quantification of progress on open data literacy. Without better measurement, it will remain challenging to secure ongoing investment instead of ad hoc programming.

A study by the School of Data that interviewed organisations involved in data literacy building noted that while effective programmes require long-term engagement, the majority of organisations are only engaged in short-term efforts (workshops that last from two hours to ten days, community events, and datathons). Funding for medium-term training and longer-term initiatives is rare.³⁸

A podcast episode for Data is a Team Sport, entitled “Government Priorities and Incentives”,³⁹ recently explored the challenges facing governments in producing open data. Ania Calderon from the Open Data Charter and Tamara Puhovski, an open government consultant based in Croatia, reflected

that for governments to develop and maintain open data programs, there needs to be an ecosystem of data literate actors. This includes knowledgeable civil servants and incentivised elected officials being held to account by a critical thinking citizenry supported by smart, open data advocates. Once elected officials are elected, it's too late to educate and motivate them to push for open data programs as they have too many other priorities and pressures. A level of data literacy is required before they are elected to provide enough knowledge, motivation, and commitment.

They also maintained that the connection between the existence of open data to data literacy efforts interwoven into the very fabric of society is not enough.

A government's incentivisation for open data can't be based on budgetary or monetary savings alone. They need to be motivated to use data to improve the effectiveness of their programs. Access to government produced open data is critical for healthy functioning democracies, but the government's ability to release open data is heavily dependent on their capacities to produce and work with data. There is currently not enough technical support for most public officials tasked with implementing open data projects.

This reality is confirmed by the World Bank, one of the international organisations that supports governments in the development of national statistical systems. The World Bank acknowledges that its own current support models have focused more on data production rather than data sharing and use, with only 27 of the 201 projects evaluated including support for activities to build capacity for data use.⁴⁰

In understanding the gaps in open data literacy building, it is also important to pay attention to issues of language, geography, and gender. Many resources and tools are available only in English, and networks like Open Heroines⁴¹ have drawn attention to gender-based disparities in the open data movement, which may have an impact on who has access to data literacy opportunities.

Efforts to build capacity within marginalised communities to help them collect and use data to prove injustices that are occurring within their communities have had varying degrees of success. A multi-year project initiated by Tactical Technology Collective supported sex workers in collecting and using data to prove the impact of police violence in their communities in Kolkata, India and Phnom Penh, Cambodia.⁴² The benefits for the advocacy groups involved included the development of skills applied to the effective collection and use of data for advocacy, new skills in analysing data to deliver a clearer understanding of the threats faced by the community, and significant improvement to their ability to conduct advocacy for their human rights. Since the data they collected had a focus on violence against women, they received requests from other advocacy groups to make use of the data. Community-based groups need to be supported to produce and use open data to strengthen their advocacy efforts. If more public data was open, more grassroots groups could learn to make use of it to contextualise their advocacy activities across the social justice spectrum.

Turning to qualitative evidence on what works, we do have some clues as to where future efforts should be focused. A group of researchers at Dalhousie University examined data literacy efforts in higher education in Canada and proposed a set of best practices in a knowledge synthesis report. They recommend a focus on hands-on learning in workshops and labs, module-based learning, project-based learning, the inclusion of real-world data that is relevant to the students' interests, and an engaging context, as well as integrating data literacy teaching into existing subjects that make use of some element of data.⁴³

The future of open data literacy

Gray et al. have illustrated⁴⁴ how the rise of open data has played an essential part in increasing awareness of data in general and in bringing the history and context of datasets into question. In addition, as awareness of machine learning and the use of data by corporations has grown, new themes have arisen in the data discourse, including issues of privacy and the ethical use of data.

The Data-Pop Alliance proposes that the only way to harness the potential of open data through data literacy is by conceptualising data literacy “as a significant means and metrics for social inclusion”.⁴⁵ They also point to big data as one of the areas where the power dynamics of data are most evident, especially concerning personal data. Even though open data advocates appeal for the clear distinction between public-interest data (government or corporate) and

personal information, initiatives like the Responsible Data Forum are pointing to the difficulty of sustaining that distinction in practice. As a result, it may be less and less sustainable to treat open data literacy in isolation from wider critical data literacy building.

In writing this chapter, we have focused on documented data literacy efforts; however, it is important to recognise that efforts are currently underway around the world, and the Open Data Barometer's (ODB) visualisation of support for use helps make sense of the status of data literacy at a global scale. The ODB asks "To what extent is training about open data available for individuals or businesses who want to increase their technical skills or develop businesses to use (open) data?" In Figure 3, the map used to visualise results indicates scores between 0 (no training available) to 10 (widespread access to high-quality training).

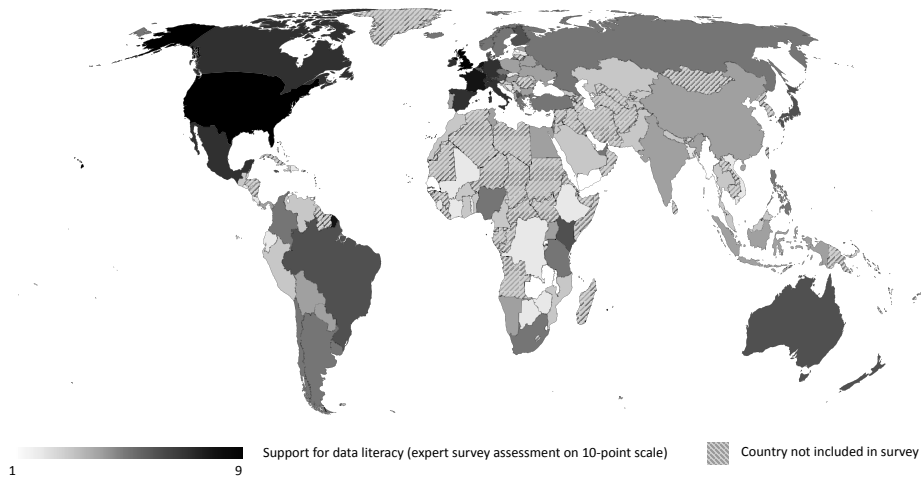


Figure 3: Support for use of data all over the world

Source: https://opendatabarometer.org/?_year=2017&indicator=ODB

Conclusion

The success of open data efforts is heavily dependent on the existence of an ecosystem of actors focused on driving the use of data through all aspects of society. There are strong traditions within civil society to build and learn from existing efforts. Through the development of networks that are openly sharing research and learnings regarding effective and less effective practices, we are developing a deeper understanding of how to achieve data literacy. However, we need to put a greater focus on understanding how class, gender, and race impact access to data and training on data use or we risk data literacy activities becoming another means of increasing inequality.

As data literacy practitioners ourselves (the authors of this chapter), we advise organisational or institutional data literacy programmes to integrate long-term capacity building, social learning exercises, awareness-raising on the collaborative nature of data projects, support for "Open Data Literacy" networks of practice, and a greater focus on the value of openness.

Long-term capacity building should provide mentoring and counselling which aim for self-sufficiency in running data-driven projects, rather than continuing to fund isolated events with the hope of increasing the number of beneficiaries through single interventions. Internews' newslab model provides a good example in this regard.

Organisations should use social learning exercises and activities to bring individuals, teams, and projects together to discuss their contexts and challenges to understand how the data will be used versus focusing entirely on technical skills. DataBasic⁴⁶ provides a good example of learning tools that contemplate social learning.

Open data literacy programming should raise awareness of the different skills and actors needed to undertake data projects, as well as the importance of distributing labour across teams and fostering collaboration. Individualistic efforts on data literacy do not suffice. School of Data's division of roles as part of data expeditions provides an example of the way this collaboration can be fostered through short-term efforts.

A focus should be placed on the value of openness in fighting inequality versus focusing solely on the value of data analysis. Equity must be placed at the centre of data analysis, and practitioners must actively push for reflection on the inclusion gaps in the data and the harm these gaps can bring. Communities like the Responsible Data Forum and Open Heroines foster continuing conversations that are relevant to this exercise.

Finally, to recognise and further the efforts that already exist on the ground, organisations should strengthen "Open Data Literacy" networks of practice to provide opportunities for peers to dialogue and discuss shared challenges and lessons learned.

Further reading

Bohman, S. (2015). Data visualization: An untapped potential for political participation and civic engagement. In A. Kő & E. Francesconi (Eds.), *Electronic government and the information systems perspective. Lecture Notes in Computer Science*, 9265, 302–315. Cham: Springer.

Journal of Community Informatics, 12(3) Special Issue on Data Literacy: <http://ci-journal.org/index.php/ciej/issue/view/59>

Letouzé, E., Noonan, A., Bhargava, R., Deahl, E., Sangokoya, D., & Shoup, N. (2015). *Beyond data literacy: Reinventing community engagement and empowerment in the age of data*. Cambridge, MA: Data-Pop Alliance (Harvard Humanitarian Initiative, MIT Media Lab and Overseas Development Institute) and Internews. <https://datatherapy.files.wordpress.com/2015/10/beyond-data-literacy-2015.pdf>

About the authors

Mariel Garcia Montes is a practitioner and researcher in the field of technology with the Center of Civic Media at the Massachusetts Institute of Technology (MIT). You can follow Mariel at <https://twitter.com/faeriedevilish>.

Dirk Slater is the founder of FabRiders, a consultancy focusing on utilising data and technologies to strengthen social change efforts. FabRiders leads and facilitates participatory events, designs and develops network-centric resources, and mentors capacity-building expertise. Follow Dirk at <http://www.twitter.com/fabriders>.

How to cite this chapter

Montes, M.G. & Slater, D. (2019). Issues in open data: Data literacy. In T. Davies, S. Walker, M. Rubinstein, & F. Perini (Eds.), *The state of open data: Histories and horizons* (pp. 274–286). Cape Town and Ottawa: African Minds and International Development Research Centre. <http://stateofopendata.od4d.net>



This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence. It was carried out with the aid of a grant from the International Development Research Centre, Ottawa, Canada.

Endnotes

- 1 Frank, M., Walker, J., Attard, J., & Tygel, A. (2016). Data literacy: What is it and how can we make it happen? *The Journal of Community Informatics*, 12(3). <http://ci-journal.org/index.php/ciej/article/view/1347>
- 2 <https://schoolofdata.org/>
- 3 Slater, D. & Garcia, M. (2016). Our data literacy research findings. *School of Data*, 8 January. <https://schoolofdata.org/2016/01/08/our-data-literacy-research-findings/>
- 4 <http://datapopalliance.org>
- 5 Letouzé, E., Noonan, A., Bhargava, R., Deahl, E., Sangokoya, D., & Shoup, N. (2015). *Beyond data literacy: Reinventing community engagement and empowerment in the age of data*. Cambridge, MA: Data-Pop Alliance (Harvard Humanitarian Initiative, MIT Media Lab and Overseas Development Institute) and Internews. <https://datatherapy.files.wordpress.com/2015/10/beyond-data-literacy-2015.pdf>
- 6 Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society*, 5(2), 1–13. <https://doi.org/10.1177/2053951718786316>
- 7 Bhargava, R. & D'Ignazio, C. (2015). *Designing tools and activities for data literacy learners*. Oxford, UK: Web Science – Data Literacy Workshop. <https://dam-prod.media.mit.edu/x/2016/10/20/Designing-Tools-and-Activities-for-Data-Literacy-Learners.pdf>
- 8 Matthews, P. (2016). Data literacy conceptions, community capabilities. *The Journal of Community Informatics*, 12(3). <http://ci-journal.org/index.php/ciej/article/view/1348>
- 9 Koltay, T. (2015). Data literacy: In search of a name and identity. *Journal of Documentation*, 71(2), 401–415. <https://doi.org/10.1108/JD-02-2014-0026>
- 10 Slater, D. & Garcia, M. (2016). Research results part 2: The methodologies of data literacy. *School of Data*, 14 January. <https://schoolofdata.org/2016/01/14/research-results-part-2-the-methodologies-of-data-literacy/>
- 11 <https://schoolofdata.org/teamsport/>

- 12 School of Data. (n.d.). Methodology. <https://schoolofdata.org/methodology/>
- 13 Ibid.
- 14 Global Integrity. (n.d.). What is a 'treasure hunt'? <https://www.globalintegrity.org/treasure-hunts/>
- 15 Figures from correspondence with the Open Data Institute.
- 16 ODI (Open Data Institute). (2018). Building open data capacity in Tanzania. <https://theodi.org/project/building-open-data-capacity-in-tanzania-the-story-of-two-years-of-training-and-support/>
- 17 IRE (Investigative Reporters and Editors). (2018). Educators' Center. <https://ire.org/resource-center/educators-center/>
- 18 ODI (Open Data Institute). (2018). Building open data capacity in Tanzania. <https://theodi.org/project/building-open-data-capacity-in-tanzania-the-story-of-two-years-of-training-and-support/>
- 19 <http://databasic.io>
- 20 D'Ignazio, C. & Bhargava, R. (2016). DataBasic: Design principles, tools and activities for data literacy learners. *The Journal of Community Informatics*, 12(3). <http://ci-journal.org/index.php/ciej/article/view/1294>
- 21 The Campaign for Data Equality. (n.d.). Research. <http://dataequality.org/research>
- 22 Pani, P. (2018). India pips Singapore, China in data literacy. *The Hindu Business Line*, 6 March. <https://www.thehindubusinessline.com/info-tech/india-pips-singapore-china-in-data-literacy/article22951784.ece>
- 23 The Campaign for Data Literacy. (n.d.). Research. <http://www.dataliteracy.info/research>
- 24 <http://datakind.org/chapters/datakind-uk>
- 25 <http://dataorchard.co.uk>
- 26 Data Evolution. (2017). The data maturity framework. <http://dataevolution.org.uk/the-framework/>
- 27 Constantares, E. (2014). Using data journalism to probe economics in the West Bank. *School of Data*, 27 August. <https://schoolofdata.org/2014/08/27/using-data-journalism-to-probe-economics-in-the-west-bank/>
- 28 Jooste, I. (2016). Data journalism: The new normal in Kenya. *Internews*, 15 June. <https://www.internews.org/story/data-journalism-new-normal-kenya>
- 29 <http://www.ayitic.net/en/index.html>
- 30 Janssen, M. & Zuiderwijk, A. (2014). Infomediary business models for connecting open data providers and users. *Social Science Computer Review*, 32(5), 694–711.
- 31 ICIJ (International Consortium of Investigative Journalists). (2018). The Panama Papers: Exposing the rogue offshore finance industry. <https://www.icij.org/investigations/panama-papers/>
- 32 Slater, D. (2017). Data is a team sport: Data-driven journalism. *School of Data*, 20 June. <https://schoolofdata.org/2017/06/20/data-driven-journalism/>
- 33 <http://exposingtheinvisible.org/>
- 34 <http://visualizingimpact.org>
- 35 <http://www.publishwhatyoupay.org/our-work/using-the-data/the-data-extractors/>
- 36 Slater, D. & Leson, H. (2018). Data playbook (Beta). International Federation of Red Cross and Red Crescent Societies Global Disaster Preparedness Center. <https://www.preparecenter.org/toolkit/data-playbook>
- 37 <http://responsibledata.io/>
- 38 Slater, D. & Garcia, M. (2016). Research results part 5: Improving data literacy efforts. *School of Data*, 5 February. <https://schoolofdata.org/2016/02/05/research-results-part-5-improving-data-literacy-efforts/>
- 39 Slater, D. (2017). Data is a team sport: Government priorities and incentives. *School of Data*, 13 August. <https://schoolofdata.org/2017/08/13/team-sport-government/>
- 40 World Bank. (2017). *Data for development: An evaluation of World Bank support for data and statistical capacity*. Washington, DC: Independent Evaluation Group, World Bank. <http://documents.worldbank.org/curated/en/705581506620506441/pdf/120111-WP-REVISED-PUBLIC.pdf>
- 41 <http://openheroines.org>

- 42 Ganesh, M., Martini, B., & Slater, D. (2015). Leave your potatoes at home. *TheFabBlog* [Blog post], 28 April. <https://www.fabriders.net/potatoes/>
- 43 Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., Kelley, D., Matwin, S., & Wuetherick, B. (2015). *Strategies and best practices for data literacy education: Knowledge synthesis report*. Halifax, Canada: Dalhousie University. <http://dalspace.library.dal.ca/bitstream/handle/10222/64578/Strategies%20and%20Best%20Practices%20for%20Data%20Literacy%20Education.pdf>
- 44 Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society*, 5(2), 1–13. <https://doi.org/10.1177/2053951718786316>
- 45 Letouzé, E., Noonan, A., Bhargava, R., Deahl, E., Sangokoya, D., & Shoup, N. (2015). *Beyond data literacy: Reinventing community engagement and empowerment in the age of data*. Cambridge, MA: Data-Pop Alliance (Harvard Humanitarian Initiative, MIT Media Lab and Overseas Development Institute) and Internews. <https://datatherapy.files.wordpress.com/2015/10/beyond-data-literacy-2015.pdf>
- 46 <http://databasic.io>