The open library and its enemies

Based on a paper presented at the 37th UKSG Annual Conference, Harrogate, April 2014.

It also owes much to research and thinking done for a lunchtime lecture organized by Digital Repository of Ireland in Dublin in September 2013, to long conversations about the role of libraries with Sandy Mahal of the Reading Agency, and to 30 years spent online.

In the age of electronics an open society – one in which questions can be asked, where critical thinking is not just permitted but encouraged and where investigation rather than ideology is used to seek out the truth about the world (the open society according to Karl Popper[1]) – has also to be an open *data* society because reusable, structured data has become the main machine for doing the heavy lifting of moving knowledge around, just as books move ideas around.

An open library is one that embodies the Open Definition[2] in its working practices, one that is available and accessible, open to all participants, and which offers services that can be freely shared. Such institutions are pillars of any open data society.

Open data and open libraries

BILL THOMPSON

Journalist, writer and commentator
BBC and freelance

The open data movement is predicated on the view that certain data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other legal or technological mechanisms of control. It is most succinctly expressed in the Open Definition, which states that 'a piece of data or content is open if anyone is free to use, reuse, and redistribute it – subject only, at most, to the requirement to attribute and/or share-alike'[3].

This is commonly interpreted around three separate axes. The first concerns *availability and access:* the data must be available as a whole and at no more than a reasonable reproduction cost, preferably by downloading over the internet. The data must also be available in a convenient and modifiable form. Second, it covers *reuse and redistribution:* the data must be provided under terms that permit reuse and redistribution including the intermixing with other datasets. The data must be machine-readable. Finally, the Open Definition mandates *universal participation:* everyone must be able to use, reuse and redistribute – there should be no discrimination against fields of endeavour or against persons or groups. This means that 'non-commercial' restrictions that would prevent 'commercial' use, or restrictions of use for certain purposes (e.g. only in education), are not allowed.

> "The library has always offered access to the private space of the mind …"

If open data is data that can be freely used, reused and redistributed by anyone then an *'open library'* is one that embodies the Open Definition in its operations – a library that is not just accessible but permeable and machine-readable, a library that offers itself as more than a passive catalogue of holdings but is a node on the network and a bridge between real and virtual, online and offline.

The point of having open libraries is not to replace physical spaces with online ones or physical books with electronic editions, it is to allow online and physical spaces to intersect so that there is no need to distinguish between them: to create, in libraries and other shared spaces, a sharing between real and virtual, a liminal space with its own affordances for readers[4]. The library has always offered access to the private space of the mind, and we see

it as readers sit engrossed in books, sometimes dropping to the floor beside a shelf because the effort of moving to a seat would require too much detachment from the scenes inside their head conjured up by the words.

Now libraries can provide access to the shared space of the network. This is more complicated, because it requires more than shelves, seats and light to read by, but no less important. We have made this space, many of us already live significant parts of our lives in it, and an open library can be a gateway to it. At the same time each individual library retains the unique property of being the only place that is the place it is – to paraphrase Walter Benjamin's comment about works of art in the age of mechanical reproduction[5].

"A call for open libraries is not a call … for us to abandon the physical library"

Even if a room in a library is a gateway to an infinitely replicable virtual space we must still acknowledge that the combination of physical and virtual space is itself as unique as the physical space, and allow for the local reality to intrude, with its own characteristics. A call for open libraries is not a call for every library to be or feel the same, or for us to abandon the physical library as the crossing-point between real and virtual.

Channelling Popper

Libraries that find ways to become crossing-points are an important component of the 'open data society', a society that is firmly grounded in access to and use of open data, a society that is transparent, open to engagement with new ways of thinking.

The term 'open data society' is a play on the formulation that philosopher Karl Popper originated in his book *The Open Society and Its Enemies*[6], written during the Second World War and first published in 1945. For Popper, an open society was not a description of a political system but rather an approach to what a society considered possible – an epistemological rather than political question.

In Popper's view an *open* society is one that is open to challenge and open to different points of view instead of being grounded in unchallengeable authority, whether religiously derived or imposed by a political ideology. This view comes from his philosophical work and his own theory of knowledge, since if knowledge is provisional and fallible this implies that society must be open to alternative points of view because the 'facts' on which it appears to be based may themselves be found to be false. An open society allows cultural and religious pluralism; by contrast closed societies are grounded in claims to certainty and an imposition of a particular version of reality, where freedom of thought is dangerous and must be suppressed and only certain forms of intellectual exploration are permissible.

When he wrote *The Open Society and Its Enemies*, Popper believed that the social sciences had failed to grasp the significance and the nature of fascism and communism because they could not understand how those types of societies understood the world. He argued that totalitarianism forced knowledge to become *political* and that this made critical thinking impossible and led directly to the destruction of knowledge in totalitarian countries, and he criticized philosophers like Plato, Hegel and Marx who he thought had laid the framework for totalitarianism.

The age of electronics

We live in an age of electronics, where many aspects of daily life are shaped – for good or ill – by the capabilities of machines that rely on the flow and detection of tiny electric currents and the opening and closing of silicon-based switches. The things these technologies can do are truly astonishing, and their application has transformed the lives of us all – not just those who have easy access to the latest shiny toys but even those who live in poverty and may never themselves hold a mobile phone or computer or share information over the internet.

As a result, conversations around openness are closely linked to conversations about the internet, not merely because the net has over the last 30 years been one of the principal channels through which ideas of openness have permeated the technology world and influenced politics and popular culture, but also because it is hard to imagine open data thinking having the impact it has had without a channel that provides easy access to that data and the results of its use, and the net *is* that channel.

The types of knowledge that open data makes possible, which an open internet makes shareable and which an open library makes accessible support the sorts of open society that Popper was concerned with, so that any society that fully embraces the open data and open knowledge manifestos would find it difficult to be closed in the Popperian sense.

That doesn't mean it would be a good society, or nice to live there, or that it would not be evil. It just would be hard for it to be closed and remain closed.

"… conversations around openness are closely linked to conversations about the internet …"

We can't rely on the internet

An open data society is what happens when Karl Popper's vision of the open society meets the internet, although its emergence and success are far from guaranteed not least because many players have an investment in closed data, closed networks and closed thinking.

However, the experiment that is the open data society started because of the largely unanticipated consequences of the global adoption of a set of technologies that were built around an assumption of openness without any real concern for the broader impact[7]. Those technologies are the ones that have given us today's internet, and continue to develop.

Today's internet is a vast, unregulated, worldwide experiment in openness, but the experiment does not come without risk. The push towards open data and the desire to build structures of scholarship, regulation and governance on top of the assumption that data will be open is one aspect, but it may well prove to be the most significant since it creates a real possibility that we will refactor modern society and find a way to build social structures on a new set of assumptions, just as the Enlightenment replaced religious catechism with the results of scientific investigation in large parts of the world 500 years ago.

That does not mean we can necessarily predict how the technology will develop. Popper argued against 'historicism', the idea that there was a flow to history and that there were core beliefs that could not be challenged. It is just as important to avoid technological essentialism, and accept that a programme or a data set has no essential values and no essential qualities. Asking 'what is this data for?' is as useful as asking 'what is a table?' Open data is a tool through which political power can be exercised in various traditional and non-traditional ways, and at the same time it defines a contested zone where politics is done.

"Today's internet is a vast, unregulated, worldwide experiment in openness …"

But we cannot simply pull down the walls to the unimpeded flow of information and expect no consequences. No technology exists in a vacuum, and the growing use of powerful digital computers connected by an ever-faster and ever more pervasive network offering gateways to vast amounts of structured data requires us to ask hard questions about the ways they will be used to shape society.

Those whose businesses rely on limiting people's ability to copy and modify songs or images or video – the 'content industries' – find it hard to cope with the openness, but so do those who want to manage the free flow of information for reasons that are not simply commercial, such as the doctors who keep our medical records or the companies storing personal e-mails, or those who make money by marshalling academic papers and selling subscriptions.

This is the conflict that lies at the heart of the open data society. It is not a technological issue and will not be solved by technology. It is at its core an issue of epistemology,

a question of how we know the world, a question that comes before we ask how that knowledge can be applied and used.

The legal, regulatory, political and financial frameworks that define modern society are not necessarily amenable to the emergence of a working open data society. Openness is fragile, open data doubly so, and the open society always subject to challenge from those who would lock down application programming interfaces (APIs)[8] or impose rigid ideologies.

"Openness is fragile, open data doubly so …"

If we want to live in an open data society then we have to build it, and if we want open libraries then we have to build them, too. Which means that libraries – and librarians – will inevitably end up taking sides in the conflict between those who believe the first part of Stewart Brand's famous epithet:

Information wants to be free[9]

And those who prefer the less well-known second half:

Information wants to be expensive[10].

References and notes

1. For background on Popper see: http://plato.stanford.edu/entries/popper/ (accessed 9 September 2014).
2. The Open Definition notes that 'A piece of data or content is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and/or share-alike' . See more at: http://opendefinition.org/ (accessed 9 September 2014).
3. The Open Definition , ref. 2.
4. I use the term loosely to cover anyone who takes advantage of the services libraries have to offer, printed, multimedia or other.
5. Benjamin, W, Hannah Arendt, ed. *"The Work of Art in the Age of Mechanical Reproduction", Illuminations*, ed. Arendt, H, 1968, London, Fontana. pp. 214–218. The full text can be found at http://www.marxists.org/reference/subject/philosophy/works/ge/benjamin.htm (accessed 9 September 2014).
6. Popper, K, *The Open Society and Its Enemies* (2 Volumes), 1945, London, Routledge. For background to Popper and his beliefs, see: http://www.iep.utm.edu/popp-pol/ (accessed 9 September 2014).
7. See for example the analysis in Hafner, K and Lyon, M, *Where Wizards Stay Up Late: The Origins of the Internet*, 1996, New York , Simon & Schuster.
8. An API is the way a computer program accesses a data set either locally or over the network.
9. The earliest recorded occurrence of the expression was at the first Hackers Conference in 1984. Brand told Apple Computer co-founder Steve Wozniak: *"On the one hand information wants to be expensive, because it's so valuable. The right information in the right place just changes your life. On the other hand, information wants to be free, because the cost of getting it out is getting lower and lower all the time. So you have these two fighting against each other."*
10. Or 'controlled'. The two are equivalent in many online settings.

Bill Thompson
BBC
E-mail: bill.thompson@bbc.co.uk | Twitter: @billt

ORCID iD: http://orcid.org/0000-0003-4402-5296

To cite this article:

Thompson, B, The open library and its enemies, *Insights*, 2014, 27(3), 229–232; DOI: http://dx.doi.org/10.1629/2048-7754.172