Privacy: the lost companion of open scholarship

On #DataPrivacyDay Camille Regnault and Kevin Sanders explore how privacy awareness interacts with openness in research and scholarship in some perhaps unexpected ways…

It may seem counter-intuitive to see privacy as remotely relevant to the discourse of open practice in research, and indeed many see privacy as inherently antithetical to the motivations of the open access movement.

Where 'privacy interests' present impediments to disseminating the results of important but sensitive research, restrict discoverability (through the use of proprietary formats which limit the utility of data mining or indexing services), run counter to the ethos of publicly-funded research or delay the reproducibility of that research (through toll-access journals and lengthy embargos), it is difficult to extol the virtues of privacy and save face.

Julie Cohen (2013), an advocate of privacy herself, admits:

The list of privacy's counterweights is long and growing. The recent additions of social media, mobile platforms, cloud computing, data mining, and predictive analytics now threaten to tip the scales entirely, placing privacy in permanent opposition to the progress of knowledge.

Despite the clear fact that privacy, speech, intellectual freedom and the Internet do not always neatly dovetail, it is important to recognise that they are also not in perpetual conflict either; it makes little sense to frame privacy and openness as a strictly zero-sum game.

In this blog post, we pitch our tent of open scholarship and open research in a liminal space between these obverse positions, and we explore how privacy and openness can positively converge to advance research and scholarship for a range of stakeholders.

For developing critical perspectives ('intellectual privacy')

Although privacy concerns have perhaps primarily been focused outside of the academy, the expanding discourse has raised some interesting questions around whether qualifications on openness can ever be reasonably applied in this domain. Openness, after all, often evokes ideals such as altruism, efficiency, academic integrity, wider participation, and innovation. However, there are also potential downsides which necessitate consideration in order to maximise the benefits of the former.

As openness facilitates exchanges with publics beyond the academy, there is a potential exposure to bad agents wishing to further aggressive stances through the veil of online 'anonymity', particularly via social media platforms.

There is also a vulnerability to a form of always-on disclosure which has the capacity to undermine the validity of research (if the data management plan is poorly defined and confidentiality is breached).

We are also increasingly aware of how corporate or state actors deploy digital profiling and blur the boundaries between perceptions of personal and academic interests and spaces online.

The Royal Society stated words to this effect in their 2012 report on Open Science:

A commitment to open science does not imply openness to everything, to anyone or for any purpose. Open science should be bounded by considerations of quality, legitimate commercial interests, privacy and security.

In some respects therefore open practice requires the consideration of privacy in order to function as intended. This is why Neil Richards (2015) maintains that privacy is a much under-valued part of intellectual engagement which allows for critical reflection and the formulation of new responses which take time.

Julie Cohen (2013) has likewise suggested that privacy, in very specific contexts, can provide:

breathing room to engage in the processes of boundary management that enable and constitute self-development… It enables individuals both to maintain relational ties and to develop critical perspectives on the world around them.

Though inherently flawed, one way in which intellectual privacy routinely operates in the chain of scholarly communication is through the mechanism of double-blind peer review. The double-blind mechanism is theoretically able to conceal both the identity of the author and reviewer, and thus intends to guard against bias. (Indeed, the BMJ also offer 'triple blind peer review' where the handling editor, reviewer and author are anonymous to each other).

This is significant where multiple studies have shown that reviewers have been influenced by the gender, ethnicity or academic standing of authors from "less prestigious" institutions.

On the other hand, many journals have been recognised for offering the opportunity to submit their work to an open peer review system, extending possibilities for further engagement and revision as well as building in accountability, and granting authors a right of reply.

The point here is not that we should systematically choose open peer review over private or vice versa, but that we should maintain practices that steer practitioners towards making informed decisions regarding the choices put before them.

For integrity in research data management (RDM)

Privacy considerations are arguably more cut and dried when it comes to outlining data management plans (DMPs) which are increasingly required by funders and other stakeholders to ensure that the results of the research are preserved and to maximise their value and impact through open data provision where appropriate.

The Wellcome Trust's requirement for an 'Outputs Management Plan' addresses its commitment to 'creating an environment that enables and incentivises researchers to maximise the value of their research outputs, including data, software and materials' for the public good.

The charity however recognises 'that in some circumstances, controls and limits on sharing are necessary – for example, to protect the confidentiality and privacy of research participants'.

A key part of the guidance on 'Access procedures for data' for example explores the various ways in which conditional access might need to be considered as a mitigation strategy 'where a study involves identifiable data about research participants'.

Privacy measures in these contexts could include controlled access for limited groups or graded access 'where less sensitive data is made readily available, and more sensitive datasets have a more stringent assessment'. This is particularly common to the field of genomics where encrypted access or 'differential privacy' may be used to protect publicly identifiable information present within large data sets; something which the anonymization of data alone, cannot currently achieve.

It could similarly comprise sharing research in publicly available subject and institutional repositories (green access) as 'closed' or 'dark deposits'. The latter allow the output to be made discoverable through rich description, use of standards, and controlled vocabularies (metadata) whilst depositors can provide a contact email address to facilitate access requests through a framework of peer trust.

For more information about data management plans, the Digital Curation Centre's (DCC) DMPonline has proformas based around the generic requirements for the research councils and major funders, which can help to frame DMPs with some disciplinary orientation.

You can also contact us directly using the address below.

For equity and security of access

We're seeing an increasing number of websites using encryption to protect data in transit on the open web, however it's still far from being the default for many Internet services and a startling number of reputable academic websites and online resources remain woefully insecure.

HTTPS ensures websites and academic blogs adhere to common security standards and therefore lends a degree of integrity to academic profiles as well as some security assurances. HTTPS creates an encrypted connection and establishes trust by verifying that you are communicating directly to the intended server and ensuring that only that server can interpret what has been sent.

HTTPS is often trivialised as 'geek speak' but is an important factor to consider in open scholarship as the encrypted connection mitigates the potential man-in-the-middle attacks that can misrepresent content served over HTTP. This provision of integrity is invisible to many, but is significant when it comes to developing your network and maximising engagement (and we would argue when online in general), particularly where opportunities for collaboration and even employment are likely to arise.

To increase your own use of encryption and to help protect your data, we would recommend installing the HTTPS Everywhere browser plugin that tells websites to use encryption where the host have implemented the technology, but may not have configured it properly. Popular blogging websites such as Wordpress also enable HTTPS.

This year, the UWL Repository celebrates its 1st birthday of being HTTPS compliant. However, as we have acknowledged in a previous blog post, 'the scholarly commons is only as accessible as it is permitted to be on the clear-net, as there are many powerful stakeholders that have the ability to suppress access and thus censor scholars and other publics from accessing the published results of academic research and scholarship'.

This is why in 2018, Library Services took the further step of making the UWL Repository accessible from within the Tor network as an onion service:

Having repositories available as onion services is of significant benefit for those accessing the material from, for instance, oppressive geopolitical contexts. Onion services offer not only enhanced privacy for users, but also help to circumvent censorship. Some governments and regimes routinely deny access to clear-net websites deemed obscene or a threat to national security. Providing an onion service of the repository not only protects those that may suffer enhanced digital surveillance for challenging social constructs or social relations (which can have a severely chilling effect on intellectual freedom), but also on entire geographical areas that are locked out of accessing publicly accessible content on the clear-net.

These actions reinforce what Vayena and Gasser (2016) concluded in their study regarding the tensions between openness and privacy in the field of genomics: 'privacy and openness are rich, complex, and related norms. It is overly simplistic to think of them as static, one dimensional, or as inherently antithetical'.

You can read more about the project and find support for Onion services here. The Onion address for the UWL Repository can be accessed over the Tor network using the following link: https://6dtdxvvrug3v6g6d.onion

Recommended tools:

- Register for an ORCID: https://orcid.org/register
- Research Data Management support from the Digital Curation Centre: https://dmponline.dcc.ac.uk/
- Install the HTTPS Everywhere browser extension: https://www.eff.org/https-everywhere
- Let's Encrypt: https://letsencrypt.org/
- Introduction to using TOR services https://www.torproject.org/about/overview.html.en

For all enquiries relating to open access, research data management, green access, etc. contact us: library[at]uwl.ac.uk

Bibliography:

Benefits and drawbacks of double-blind peer review. (n.d.). Retrieved 27 January 2019, from https://www.exordo.com/blog/double-blind-peer-review/

BMJ. (n.d.). The peer review process. Retrieved 27 January 2019, from https://authors.bmj.com/after-submitting/peer-review-process/

Cohen, J. (2013, May 20). What Privacy is For. Retrieved 25 January 2019, from https://harvardlawreview.org/2013/05/what-privacy-is-for/

Differential Privacy. (n.d.). Retrieved 22 January 2019, from https://privacytools.seas.harvard.edu/differential-privacy

Erlich, Y., Williams, J. B., Glazer, D., Yocum, K., Farahany, N., Olson, M., … Kain, R. C. (2014). Redefining Genomic Privacy: Trust and Empowerment. *PLOS Biology*, *12*(11), e1001983. https://doi.org/10.1371/journal.pbio.1001983

Gajda, A. (2016). Academic Freedom, the Presumption of Openness, and Privacy. *Revue Internationale Des Gouvernements Ouverts*, *2*(0), 151–164. Retrieved from http://ojs.imodev.org/index.php/RIGO/article/view/14

Kuehn, B. M. (2017). Rooting out bias. *ELife*, *6*, e32014. https://doi.org/10.7554/eLife.32014

Makula, A. (2017). "Is it like academia.edu?": Faculty perceptions and usage of academic social networking sites and implications for librarians and institutional repositories. Journal of New Librarianship, 2(1), 2479. https://doi.org/10.21173/newlibs/2/1

P, C. (2016, June 3). Social media and the student experience — a reflection on open communication in Higher Education. Retrieved 25 January 2019, from https://medium.com/open-knowledge-in-he/social-media-and-the-student-experience-a-reflection-on-open-communication-in-higher-education-84a871e659f1

Richards, N. (2015, January 7). How encryption protects our intellectual privacy (and why you should care). *Wired UK*. Retrieved from https://www.wired.co.uk/article/encryption-intellectual-privacy

Ross-Hellauer, T. (2017). What is open peer review? A systematic review. *F1000Research*, *6*, 588. https://doi.org/10.12688/f1000research.11369.2

The Royal Society. (2012). Science as an open enterprise (Summary), 10. Retrieved from https://royalsociety.org/~/media/policy/projects/sape/2012-06-20-saoe-summary.pdf

Sanders, K. (n.d.). Bag of onions: growing bulbs of intellectual freedom from academic libraries. Retrieved 27 January 2019, from https://uwlopenaccess.edublogs.org/2018/03/05/bag-of-onions-ggrowing-bulbs-of-intellectual-freedom-from-academic-libraries/

Shore, P. (2017, August 30). Open Science — can we be too open? Retrieved 25 January 2019, from https://medium.com/open-knowledge-in-he/open-science-can-we-be-too-open-bc4f15aebf57

Single-Blind Vs. Double-Blind Peer Review. (2018, February 13). Retrieved 27 January 2019, from https://www.enago.com/academy/double-blind-peer-review-for-better-or-for-worse/

Vayena, E., & Gasser, U. (2016). Between Openness and Privacy in Genomics. *PLOS Medicine*, *13*(1), e1001937. https://doi.org/10.1371/journal.pmed.1001937

Wellcome. (2017). Policy on data, software and materials management and sharing. Retrieved 26 January 2019, from https://wellcome.ac.uk/funding/guidance/policy-data-software-materials-management-and-sharing

Wellcome. (n.d). Developing an outputs management plan. Retrieved 26 January 2019, from https://wellcome.ac.uk/funding/guidance/developing-outputs-management-plan

Weller, M. (2016, January 11). Advantages & disadvantages of openness. Retrieved 25 January 2019, from https://medium.com/open-knowledge-in-he/advantages-disadvantages-of-openness-bb9790c06c1b#.1nbb95i8e