

Data privacy in our libraries

HSLG Journal Club 1st Dec 2022

Caitríona Lee

Information Specialist, Health Research Board



The paper: Data tracking in research

Data tracking in research: aggregation and use or sale of usage data by academic publishers.

A briefing paper of the Committee on Scientific Library Services and Information Systems of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation).

DFG-Committee on Scientific Library Services and Information Systems. Zenodo, 2021. doi:10.5281/ZENODO.5937995

Available from: <https://zenodo.org/record/5937995>

Note

Disclaimer:

Any opinions stated in this Journal Club are just opinions, not advice or instructions

I am not a lawyer

I am not an IT professional, a behavioural scientist or a data privacy expert

I cannot advise on security or legal issues

For any concerns, talk with your IT professionals, vendors or a legal expert (and write up the results for HINT please, we'd love to know too)

The plan

1. Quick overview of the journal club paper
2. Discussion of the topics arising in the paper
3. Discussion of our own knowledge and awareness of the topics

Data tracking in research: Contents

Contents

1. Description of the current situation
2. The transformation of the major publishers and their relationship with the academic community
 - 2.1 Consequences of the transformation of publishers into data analytics businesses
3. Types of Data mining
 - 3.1 Third Party Data through Microtargeting
 - 3.2 Bidstream Data and port scanning
 - 3.3 “Spyware”

Data tracking in research: Section 1

1. Description of the current situation

Expansion of academic publishers' business model:

Traditional content provision - publishing journals and search tools



Information analytics, wrap-around services and data brokering

Data tracking in research: Section 2

2. The transformation of the major publishers and their relationship with the academic community

Provide technical proprietary services for the entire research process and the analysis of research-related data

e.g. Elsevier's Professional Services contact with the Netherlands, Germany's DEAL agreements with Springer Nature and Wiley, Elsevier's subcontracting with the European Commission on Open Science Monitor

Potential consequences: industrialisation of knowledge through tracking, potential risks to researchers, sale of data for potentially unethical work

Data tracking in research: Section 3

Types of Data mining

- Third Party Data through Microtargeting
- Bidstream Data and port scanning
- “Spyware”

Examples of technical methods used to track users and collect data for aggregation and/or sale

Questions: Relevance

**Having read this briefing paper on data privacy in libraries by a prominent German research organisation,
is this issue relevant for our libraries, specifically health libraries?**

Questions: Risks

Risks

What risks could there be to health library users from **usage data collection, data aggregation, profile building, or data resale** by organisations that wrap around the entire scholarly and research complex from journal publishing, primary study research tools, preprint servers and repositories, search platforms, annotation tools, metrics collection, research network tools, research evaluation tools, plagiarism tools, online training tools, and online organisational platforms?

Questions: Professional choices

Professional choices

What choices can health librarians employ on this issue?

- We often have little control over IT-related questions
- We have to supply e-journals and e-books to our readership. Alternatives are often not possible in healthcare - specific titles are required
- We have to track some usage to prove value
- We do not often cannot choose against the lowest-cost product/a product by a company with which we already have a contract

Questions: Health libraries

Health libraries and librarians

Lots of coverage from public and academic libraries but not from health:

- Are we too small – solo and small libraries just can't do and be everything?
- Are the aspects relating to health already covered by research libraries?
- Are we protective of our user privacy by not discussing the products we use and the issues they may have with privacy?

Questions: Geography and legislation

Geography and legislation

This paper is written by German librarians and researchers. Are we less at risk in Europe with our **data protection laws** (GDPR, the Digital Services Package, the upcoming Cyber Resilience Act), or do we assume we are?

With the hazards of **data transfer** and the invalidity of the EU-US Privacy Shield, how safe is data transfer from the EU to third countries?

Questions: Privacy is over

Privacy is over

Is privacy in practice **a thing of the past** – a person must accept being tracked by the library, the vendor and third parties who purchase the data?

Is the **philosophy of privacy** as a right not longer in use – as long as purchasers of user data do not leak the data or harm the user directly, they are entitled to track users around the net?

Future harms – the lack of privacy may not harm us so much now in this time and country, but it harms others now and our conditions can change.

Questions: User behaviour

User behaviour

- We typically give library users information about data protection and use of cookies on our library websites.
- Do we think the users read the notices, understand what is implied and perhaps change their cookie settings, or do we all just click OK to all and go straight to what we want?
- Could we assume the permissions don't limit data collection much, as so much data can be collected under 'legitimate reasons'?

Questions: Situational awareness

Situational Awareness

- What level of awareness do we individually and collectively as a profession have, or need to have, of detailed issues of data privacy?
- How much IT knowledge do we need to understand what is in question?
- We probably trust our vendors with our data, but do we know what they do with it?
- How much awareness can we have of our vendors' work as information analytics companies, data brokers, data prediction companies?

Thanks!

Thanks for participating in this session of the HSLG Virtual Journal Club!

Some follow-up reading: Books

- Clark ACR, Val SED. Radical resistance: Libraries, defiance, and data surveillance. In: *Digital totalitarianism*. Routledge,, 2022 <https://www.routledge.com/Digital-Totalitarianism-Algorithms-and-Society/Filimowicz/p/book/9781032002415>
- Lamdan S. *Data cartels: The companies that control and monopolize our information*. Stanford: Stanford University Press 2022 <https://www.sup.org/books/title/?id=33205>
- Pedley P. *A practical guide to privacy in libraries*. Facet Publishing, 2020. <https://www.facetpublishing.co.uk/page/detail/a-practical-guide-to-privacy-in-libraries/?k=9781783304684>
- Véliz C. *Privacy is power*. Penguin, 2021. <https://www.penguin.co.uk/books/442343/privacy-is-power-by-carissa-veliz/9780552177719>
- Zuboff S. *The age of surveillance capitalism*. London, UK: Profile Books, 2019. <https://www.publicaffairsbooks.com/titles/shoshana-zuboff/the-age-of-surveillance-capitalism/9781610395694/>

Some follow-up reading: Articles

- Gardner GJ. Aiding and Abetting: Third-Party Tracking and (In)secure Connections in Public Libraries, 2021.
<http://eprints.rclis.org/42329/>
- Custers B, Malgieri G. Priceless data:: why the EU fundamental right to data protection is at odds with trade in personal data. Computer Law & Security Review 2022;45:105683.
<https://doi.org/10.1016/j.clsr.2022.105683>
- Lamdan S. Social Media Privacy: A Rallying Cry to Librarians. Publications and Research, 2015.
https://academicworks.cuny.edu/cl_pubs/52
- Pooley J. Surveillance Publishing. The Journal of Electronic Publishing, 2022.
<https://doi.org/10.3998/jep.18744>
- Sutlieff L, Chelin J. 'An absolute prerequisite': The importance of user privacy and trust in maintaining academic freedom at the library. Journal of Librarianship and Information Science 2010.
<https://doi.org/10.1177/09610006103689162>

Some follow-up reading: Websites

- Data Doubles. Data Doubles. Data Doubles. 2022. <https://datadoubles.org/>
- Electronic Frontier Foundation. Cover Your Tracks. 2022. <https://coveryourtracks.eff.org/>
- EDUCAUSE. Higher Education Community Vendor Assessment Toolkit. 2021. <https://library.educause.edu/resources/2020/4/higher-education-community-vendor-assessment-toolkit#tools>
- Have I Been Pwned. Have I Been Pwned: Check if your email has been compromised in a data breach. 2022. <https://haveibeenpwned.com/>
- Library Freedom. Library Freedom: Resources. 2021. <https://libraryfreedom.org/resources/>
- Stop Tracking Science. Stop Tracking Science. Open Science Future. 2022. <https://open-science-future.zbw.eu/en/stop-tracking-science-2/>

Some follow-up reading: Blogs

- Cukier E. What the vendor saw: Digital surveillance in libraries. ALA Intellectual Freedom Blog. 2022. <https://www.oif.ala.org/oif/what-the-vendor-saw-digital-surveillance-in-libraries/>
- DiVittorio K, Gianelli L. Ethical financial stewardship: one library's examination of vendors' business practices. In the Library with the Lead Pipe. 2021. <https://www.inthelibrarywiththeleadpipe.org/2021/ethical-financial-stewardship/>
- Lamdan S. The quiet invasion of 'Big Information'. Wired 2022. <https://www.wired.com/story/big-information-relx-privacy-surveillance-data/>
- Keller D. User privacy vs. platform transparency: The conflicts are real and we need to talk about them. 2022. <https://cyberlaw.stanford.edu/blog/2022/04/user-privacy-vs-platform-transparency-conflicts-are-real-and-we-need-talk-about-them-0>
- Hanson C. User tracking on academic publisher platforms. 2019. <https://www.codyh.com/writing/tracking.html>

Some follow-up reading: Reports

- Chen G, Posada A, Chan L. Vertical Integration in Academic Publishing : Implications for Knowledge Inequality. In: Mounier P, ed. Connecting the Knowledge Commons — From Projects to Sustainable Infrastructure : The 22nd International Conference on Electronic Publishing – Revised Selected Papers. Marseille: : OpenEdition Press 2019. <http://books.openedition.org/oep/9068>
- Christl W. Corporate Surveillance in Everyday Life. Cracked Labs, 2017. https://crackedlabs.org/dl/CrackedLabs_Christl_CorporateSurveillance.pdf
- IFLA. IFLA Statement on Privacy in the Library Environment. 2015. <https://www.ifla.org/publications/ifla-statement-on-privacy-in-the-library-environment/>
- SPARC. Addressing the alarming systems of surveillance built by library vendors. SPARC. 2021. <https://sparcopen.org/news/2021/addressing-the-alarming-systems-of-surveillance-built-by-library-vendors/>
- Vaughan J. Library Privacy Policies. Library Technology Reports vol. 56, no. 6, 2020 <https://journals.ala.org/index.php/ltr/issue/view/771>