



NAVIGATING RISK IN VENDOR DATA PRIVACY PRACTICES

An Analysis of Springer
Nature's SpringerLink

October 2024

© 2024 SPARC, subject to a Creative Commons Attribution 4.0 International License



NAVIGATING RISK IN VENDOR DATA PRIVACY PRACTICES

An Analysis of Springer Nature's SpringerLink

SPARC*

October 2024

© 2024 SPARC, subject to a Creative Commons Attribution 4.0 International License



Project Statement

Becky Yoose of LDH Consulting Services prepared this report in collaboration with Nick Shockey of SPARC as part of the Navigating Risk in Vendor Data Privacy Practices Project at <https://sparcopen.org/our-work/privacy-and-surveillance-community-of-practice>.

The analysis of this report is based off of the methodology used in the 2023 SPARC report *Navigating Risk in Vendor Data Privacy Practices: An Analysis of Elsevier's ScienceDirect*, located at <https://zenodo.org/records/10078610> (subsequently referred to in this report as the ScienceDirect Report). The ScienceDirect Report is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Parts of the ScienceDirect Report text have been adapted for inclusion in this report.

The following analysis reflects publicly available information at the time of review (Fall/Winter 2023) and focuses on the North American context. The SpringerLink Privacy Policy was updated in June 2024, but there are no changes that impact the findings from the analysis conducted in 2023. The applicability of our findings may be limited in other regions and may change over time as privacy policies and practices are revised. If you have additional information that could impact this analysis, please email Nick Shockey at nick@sparcopen.org.

This report is not an analysis of whether SpringerLink complies with applicable privacy laws nor should it be construed as legal advice. For additional information about the limitations of the report, as well as future research opportunities, please see the "Limitations and Further Investigation" section of the report.

This report is published by SPARC, a project of the New Venture Fund, and licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Acknowledgements

This report benefitted from feedback and guidance from Michele Gibney, Dorothea Salo, and John Mark Ockerbloom. The analysis builds on prior work by the Licensing Privacy Project, which provided the framework for the data privacy assessment. In doing this work, we are also grateful for the related work by many organizations, including the Library Freedom Project and Library Futures.

Contents

Project Statement.....	1
Acknowledgements	1
Executive Summary	3
Introduction.....	5
Analysis Methodology	8
Analysis: Experience and Assessment.....	13
Analyzing SpringerLink’s Contracts, Policies, and Publication Documentation	13
Investigating Data Collection and Tracking on the SpringerLink Website.....	28
Overall Data Privacy Assessment.....	35
Limitations and Further Investigation	40
Key Findings.....	41
Suggested Actions	43
Appendix A – Documents Used for Rubric Analysis.....	48
Appendix B – Glossary	50
Appendix C – Web Tracker List and Counts	54
Appendix D – Advertiser List	58

Executive Summary

Serious threats to privacy have followed libraries' transition from buying materials to licensing content. User tracking that would be unthinkable in a physical library setting now happens routinely through licensed platforms. This change has shifted control over library user data (and whether it is collected or kept at all), including personal data about what people search for and what they read, from the local library to third-party vendors.

This report focuses specifically on SpringerLink, the platform providing most users access to the articles of the world's second-largest academic publisher, Springer Nature. SpringerLink provides a case study in the encroachment of the broader surveillance-based data brokering economy into academic systems. Combined with our 2023 report on Elsevier's ScienceDirect platform, this analysis illustrates the wide range of privacy risks inherent in the business models and practices in the academic scholarship marketplace. The risks for each vendor are different, however, and academic institutions should carefully consider the distinct concerns posed by each.

Unlike Elsevier, Springer Nature does not currently operate a full breadth of products across the research lifecycle, nor are they part of a larger corporate structure that could enable the direct use of user information in non-academic or data brokering products. However, SpringerLink is a conduit for data collection by an extensive number of third parties which risks the monetization of patron data by these external firms in ways that directly conflict with library privacy standards and could negatively impact users themselves. These firms include companies with troubling data use practices, including one that has been sued by the FTC for allegedly selling precise geolocation data that could be used to track individuals' presence at sensitive locations.¹

The online ad exchanges and data brokering economies that many of these third-party firms participate in are technically complex, often opaque, and rapidly evolving. While the concept of "personalized advertising" may seem innocuous, it obscures the extent of information collected about individuals and the potential for data disclosed in one context, such as to sell advertisements via an online ad exchange, to be collected by others, reassociated with individuals after deidentification/anonymization ("reidentification"), and used for entirely different purposes.²

Based on our analysis, SpringerLink's data privacy practices undermine basic library privacy standards and practices. The following are examples of current SpringerLink practices found in our analysis that conflict

1 Federal Trade Commission. "FTC v Kochava, Inc.," August 29, 2022. <https://www.ftc.gov/legal-library/browse/cases-proceedings/ftc-v-kochava-inc>.

2 Tau, Byron. *Means of Control: How the Hidden Alliance of Tech and Government Is Creating a New American Surveillance State*. First edition. New York: Crown, 2024. See introduction (p. xv-xxvii) and chapters 15-16 (p. 171-195).

with the ALA Code of Ethics,³ the Library Bill of Rights,⁴ and the IFLA Statement on Privacy in the Library Environment:⁵

- SpringerLink uses web beacons, cookies, and other invasive web surveillance methods (“online trackers”) to track user behavior outside and beyond the SpringerLink website.
- 200 third parties—including data brokers, advertisers, and marketing companies—listed on SpringerLink’s website are allowed to collect user information that “may be used by those companies to build a profile,” along with what appear to be additional unlisted companies found only in our public website analysis.⁶
- SpringerLink’s Privacy Policy explicitly states that it does not cover data collected by these third parties, which is subject to their own privacy policies.
- Users’ ability to manage their choices or provide meaningful consent is deeply undermined through broken links to cookie management screens, broken links to third-party privacy policies, and the sheer volume of reading required to review all third-party privacy policies (likely stretching to thousands of pages).

Significant expertise and capacity are required to understand Springer Nature’s privacy practices and the potential downstream uses of data that may be collected through their products. Extremely few, if any, libraries have such resources available to closely review even a single vendor, further exacerbating the power asymmetry between vendors and libraries.

Libraries should question why these third parties are permitted access to patron data and press for limitations on data collection and use that are durable, verifiable, and not limited to a particular jurisdiction. This report closes with options that institutions may consider to mitigate risks to their faculty and students over the short and longer term. Collaborative efforts, such as SPARC’s Privacy & Surveillance Community of Practice, can play a key role in supporting future action to address the real privacy risks posed by vendors’ increasing connection to the wider data brokering economy.

3 ALA. “ALA Code of Ethics.” <https://www.ala.org/tools/ethics>.

4 ALA. “Library Bill of Rights.” <https://www.ala.org/advocacy/intfreedom/librarybill>.

5 IFLA. “IFLA Statement on Privacy in the Library Environment.” <https://www.ifla.org/publications/node/10056>.

6 SpringerLink “Manage your cookies preferences” banner, “Cookies that help show personalised advertising.” See the section below on “Investigating Data Collection and Tracking on the SpringerLink Website” for documentation regarding the unlisted companies.

Introduction

The following report provides a detailed analysis of the privacy practices of Springer Nature's SpringerLink product. This analysis continues SPARC's work to assist libraries in understanding and taking action to address the complex, rapidly evolving landscape of potential surveillance risks across the policies, notices, and contract language that comprise key vendors' privacy practices. Combined with our 2023 privacy analysis of Elsevier's ScienceDirect platform, this report's findings further underline the urgency for libraries to better understand these risks and, where necessary, to act to mitigate them. While this analysis and recommended actions are grounded in the library context, these findings will raise pressing issues for faculty, administrators, and policymakers to consider as well.

This report focuses specifically on SpringerLink, the platform through which most users will access the articles of the world's second-largest academic publisher, Springer Nature. Compared to our earlier analysis of Elsevier's ScienceDirect, SpringerLink is notable for allowing 200 third-parties to collect information from users of the website and for disclaiming responsibility for data collection and use by these third parties. Our analysis of SpringerLink is a case study in the encroachment of the broader data brokering economy into academic systems, standing in stark contrast to librarians' professional commitment to resisting user surveillance.

The 200 third parties that may collect information on SpringerLink users are described as "advertising partners" in the cookie management preferences menu, which contains the only user-accessible list of said "partners." These companies include advertisers, data brokers, data intelligence companies, and retailers. The data markets these kinds of companies participate in are known to expose data for further uses by fourth parties—including other data brokers and government and law-enforcement agencies.⁷

The online ad exchanges that drive the multi-billion dollar personalized advertising⁸ market offer snippets of information—often described as anonymized or pseudonymized—to serve customized ads against.⁹ This information might include what page someone is visiting, their exact geographic location, or unique advertising IDs associated with their device or user profile, and while the data does not directly contain a user's

7 Tau, Byron. *Means of Control: How the Hidden Alliance of Tech and Government Is Creating a New American Surveillance State*. First edition. New York: Crown, 2024. See chapters 15-16 (p. 171-195).

8 Personalized advertising uses personal data collected through user surveillance to offer ads specific to that user's interests and behaviors.

9 Tau, Byron. *Means of Control: How the Hidden Alliance of Tech and Government Is Creating a New American Surveillance State*. First edition. New York: Crown, 2024. See introduction (p. xv-xxvii) and chapter 16 (p. 184-195).

name or government identifier, it risks being readily reassociated with that user when combined with large data sets maintained by a growing number of data brokers.¹⁰

As a result, some advertised safeguards for user privacy (e.g. “[advertising partners] do not store directly personal information”) ring hollow as the information collected on library users through vendors’ platforms flows through wider data ecosystems. Information that may be transient for one party as an advertising market maker may be collected and retained by other entities participating in the ad exchange. Such tactics are known to be employed by data brokers, including those catering to law-enforcement and intelligence agencies.¹¹

The re-identification of supposedly “anonymized” user data collected by institutionally-licensed resources would fundamentally break library privacy commitments. While the opacity of data flows prevents our analysis from documenting the full extent of possible surveillance risks for SpringerLink (or other vendor systems), the seriousness of the threat it poses should lead libraries and institutional counsel to insist on durable and verifiable protections from these potential data uses. Usage data, search-related data, behavior data, login data, logs, and other identified or re-identifiable user information should be used solely for the narrow business purposes required to fulfill the terms of vendors’ contracts with libraries. Data from library users should not be collected by hundreds of third parties whose interests likely run counter to those of libraries and their users.

Like our 2023 ScienceDirect report, this analysis highlights the extent to which evaluating a vendor's data privacy practices has become increasingly complex for libraries due to contracts and documentation that are vague and/or difficult to understand as well as the opacity of data flows within and beyond a product or company. In the case of SpringerLink, the complexity is further compounded by numerous broken links and copy editing errors making policies more difficult to follow and creating ambiguity around key policy elements. Additionally, because Springer Nature disclaims responsibility for third-party data collection and use, users (as well as library negotiators and institutional counsel) must read 200 additional privacy policies—likely stretching into the thousands of pages—to understand how user data may be used. And as we will describe later in the report, the broken links to the cookie management screen can impede users’ ability to manage cookie preferences that affect the extent and type of tracking and data collection present when

10 Tau, Byron. *Means of Control: How the Hidden Alliance of Tech and Government Is Creating a New American Surveillance State*. First edition. New York: Crown, 2024. See introduction (p. xv-xxvii) and chapters 15-16 (p. 171-195).

11 Ibid.

using SpringerLink. Given broken links to a significant portion of these policies, meaningful user consent¹² is arguably impossible.¹³

Library workers, procurement professionals, institutional counsel, and other institutional workers tasked with evaluating a vendor's privacy practices typically do not have access to the dedicated resources needed to evaluate the full array of any given vendor's data privacy practices (including those of third-party data collectors) to understand the risks they may present and develop an effective institutional response.

In conjunction with SPARC's other vendor privacy reports, the following analysis aims to assist both users and institutions by identifying specific problematic data privacy practices through the review of SpringerLink's contracts, public documentation, and website. This analysis also suggests actions institutions can take to address these issues for SpringerLink and other vendor products that present similar concerns.

12 While we use the concept of consent in the report, we do not use it in the context of Springer Nature's compliance to any data protection regulation that may have a legal definition of user consent. Please see the Glossary in Appendix B for the report's definition of consent.

13 Please refer to the "Analyzing SpringerLink's Contracts, Policies, and Public Documentation" section for additional information about the forementioned issues.

Analysis Methodology¹⁴

The SpringerLink analysis closely mirrors the analysis in the 2023 ScienceDirect Report.¹⁵ Except for one major change, explained in the front-end website data collection section below, the methodology remained largely unchanged between the two analyses. We will still explain the methodology, for transparency and for those who have not yet read the 2023 ScienceDirect Report.

The structure of the SpringerLink analysis borrows from the data privacy criteria laid out in the Vendor Contract and Policy Rubric from the Licensing Privacy Project.¹⁶ Created to evaluate the data privacy risks in content platform contracts and policies, the rubric measures risk in eight privacy domains:

PRIVACY DOMAIN	DEFINITION
<i>Data collection</i>	What user data the vendor collects, how they collect it, where they collect user data, and their stated rationale for collecting it
<i>User data rights</i>	What controls users have over the vendor's ability to collect, retain, use, and share user data
<i>Data disclosure</i>	What user data the vendor shares, which parties the vendor shares data with, the reasons the vendor shares data, and how data sharing is controlled/determined
<i>Data processing</i>	What user data the vendor uses, for what purpose, and how data use is controlled/determined
<i>Privacy policy</i>	Public privacy statements on the vendor's service or website, as well as any internal vendor privacy policies the vendor provides the library
<i>Data ownership</i>	Who owns the data in the vendor service or product, and what rights come with data ownership in specific business scenarios
<i>User surveillance</i>	What tracking or logging mechanisms the vendor uses to collect user data, and the level of control users have over vendor tracking/logging behavior while using the service
<i>Data security and accountability</i>	How the vendor protects user data in transit and storage from unauthorized access or use, how the vendor works to prevent and respond to data breaches or leaks, and what checks are in place to ensure compliance with vendor security and privacy policies and industry standards

Privacy practices in each domain are measured against three levels of data privacy based on a consensus of library data privacy standards, guidelines, and practices from the profession and major professional organizations such as ALA and IFLA (referred to as Minimum Viable Privacy or MVP). Each level measures risk to

14 Parts of this section were adapted from SPARC's 2023 report, Yoose, Becky, and Nick Shockey. "Navigating Risk in Vendor Data Privacy Practices: An Analysis of Elsevier's ScienceDirect." SPARC, November 7, 2023. <https://doi.org/10.5281/zenodo.10078610>.

15 Yoose, Becky, and Nick Shockey. "Navigating Risk in Vendor Data Privacy Practices: An Analysis of Elsevier's ScienceDirect." SPARC, November 7, 2023. <https://doi.org/10.5281/zenodo.10078610>.

16 Licensing Privacy. "Assessing Contracts." <https://publish.illinois.edu/licensingprivacy/contracts/>.

patron privacy, with the caveat that a vendor meeting MVP does not automatically mean that the vendor is adequately protecting patron privacy as a whole.¹⁷

The SpringerLink analysis assessed readily available materials: contracts, public documentation, and the front-end (patron-facing) website. No materials analyzed were in areas that require special permission to access (e.g., a support site that requires users to log in before accessing help documentation).

CONTRACTS, POLICIES, AND PUBLIC DOCUMENTATION

While the Vendor Contract and Privacy Rubric is built to assess the vendor contract and privacy policy, our analysis goes far beyond these specific documents to provide a broader picture of the data privacy practices of SpringerLink. The documents used in the analysis include the following:

- Signed Springer Nature Master License Agreement contracts from eight academic libraries and two library consortiums in the US (two of the contracts are available in the SPARC Contract Library: <https://sparcopen.org/our-work/big-deal-knowledge-base/contracts-library/>)
- California Privacy Rights Act Policy
- Cookies Consent/Cookie Policy
- SpringerLink Privacy Policy
- SpringerLink Terms and Conditions
- Springer Nature Privacy Policy
- Springer Nature Profile Privacy Policy
- Springer Nature Terms and Conditions
- Springer Support Center documentation
 - » Account deletion
 - » Data security
 - » Manage cookie settings
 - » Personalization features
 - » About Springer Nature Profile

Links to the archived versions of the above documents can be found in Appendix A.

17 LDH Consulting Services. "Developing the Vendor Contract and Policy Rubric," January 2022. <https://publish.illinois.edu/licensingprivacy/files/2022/03/Licensing-Privacy-Vendor-Rubric-White-Paper.pdf>.

FRONT-END WEBSITE DATA COLLECTION

The second half of the analysis focused on investigating data collection and tracking behavior occurring on the public-facing SpringerLink website. The investigation examined four specific web pages at different levels of the website to capture a broader picture of the possible collection and tracking throughout the entire site:

- The SpringerLink home page (<https://link.springer.com/>)
- Journals A-Z (<https://link.springer.com/journals/a/1>)
- The search results page for a keyword search (<https://link.springer.com/search?query=cat+behavior>)
- A page for an Open Access journal article (<https://link.springer.com/article/10.1186/s12915-022-01369-1>)

The investigation used several mature and widely used tools available to the general public to inspect website traffic:

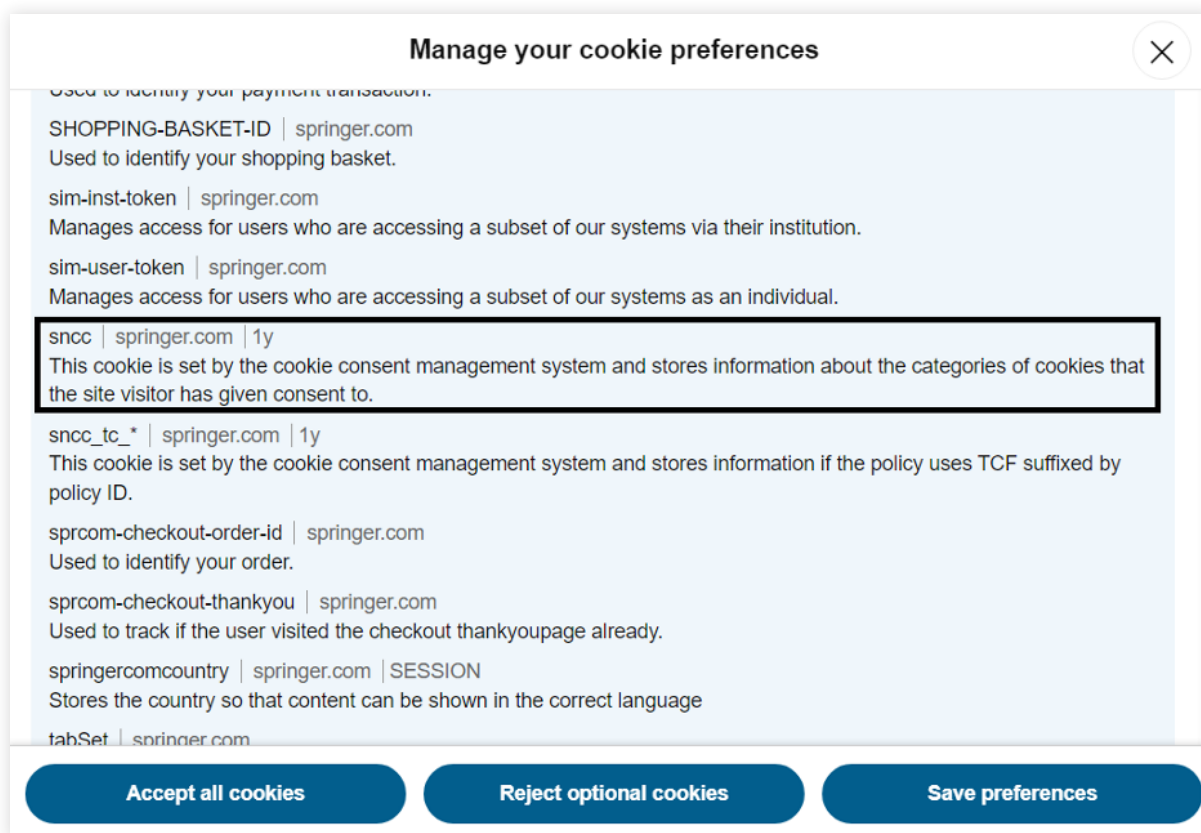
- Website Evidence Collector (WEC; <https://github.com/EU-EDPS/website-evidence-collector>), a tool developed by the European Data Protection Supervisor to document what information is collected and transmitted by a specific website, including web cookies, local storage, and data requests and traffic
- NoScript (<https://noscript.net/>), a browser add-on that detects and blocks JavaScript and other executable content on websites
- uBlock Origin (<https://ublockorigin.com/>), a browser add-on that detects and blocks ads, trackers, and other privacy-invasive content
- Web Developer Tools in Firefox and Chrome

Each URL was tested with a local installation of WEC. The results from each test were spot-checked by the browser plugins installed in Firefox (NoScript) and Chrome (uBlock Origin) and with the network monitor tool in the Web Developer Tools in both Firefox and Chrome.

Unfortunately, we were unable to use Blacklight for the SpringerLink analysis due to the tool's limitations. Blacklight (<https://github.com/the-markup/blacklight-collector>) is a tool developed by the nonprofit news organization The Markup (<https://themarkup.org/>) to detect several standard surveillance methods, including ad trackers, canvas fingerprinting, key logging, and tracking by Google and Facebook. Blacklight was used for the previous ScienceDirect analysis. When attempting to test the SpringerLink website with Blacklight we found that the tool was unable to be configured to navigate SpringerLink's cookie banner which requires an acceptance or management of cookies before one can use the website. This limitation somewhat impacted

the ability to test the SpringerLink website for canvas fingerprinting¹⁸ and recording of users' mouse activity and keystrokes.

FIGURE 1



A partial list of cookies used by SpringerLink from the cookie management window.

The SpringerLink cookie banner did not impact the WEC tests thanks to the tool's ability to set a cookie value in test parameters. WEC's `--set-cookie` parameter allows a user to set a value for a specific cookie. In the case of SpringerLink and other websites that use cookie banners, the choices made by the user is stored in a cookie; therefore, one can run multiple tests on the same web page with different types of choices

18 Canvas fingerprinting is a specific browser fingerprinting method that uses the HTML5 canvas element to track unique users on a website. More details about canvas fingerprinting can be found in the glossary in Appendix B.

(i.e., rejecting all non-essential cookies, accepting all cookies, accepting some non-essential cookies). The cookie containing the user's consent choices can be found through the network tab in the web developer tools available in many browsers. SpringerLink's website uses the **sncc** cookie to store the choices a user makes in the cookie banner.

For example, a **sncc** value of "**P=17:V=38.0.0&C=C01&D=true**" indicates that the user rejected all non-essential cookies and a value of "**P=17:V=38.0.0&C=C01,C02,C03,C04&D=true**" indicates that the user accepted all cookies.¹⁹ Therefore, we were able to run two different tests for every URL: one test that accepted all cookies and one test that rejected all non-essential cookies. The ability to set a cookie value in WEC became invaluable in our SpringerLink analysis when the disparity of tracking results between the two tests for each URL became very apparent, about which we will go into detail in the Analysis section below.

¹⁹ Readers wishing to replicate the analysis should be aware that the values may have changed by the time of publication, particularly the values for V and P. Readers should confirm the values for the sncc cookie before replicating the WEC tests.

Analysis: Experience and Assessment

ANALYZING SPRINGERLINK'S CONTRACTS, POLICIES, AND PUBLIC DOCUMENTATION

The ScienceDirect report discussed the web of documentation users must navigate to understand the data privacy practices and policies of a particular vendor. Awareness of problematic data privacy practices is difficult to achieve in such a web, particularly when key pieces of information are only found in a single document. Like ScienceDirect, SpringerLink has a similar web of documentation that users, library workers, and institutional counsel must go through to learn about data privacy practices; however, SpringerLink presents an additional set of barriers not previously discussed in the ScienceDirect report, such as broken links to important privacy controls and confusing copyediting errors in policy statements. These additional barriers are also not unique to SpringerLink. Like the concerns and issues brought forth in the ScienceDirect report, they are representative of practices and barriers found in other vendors.

CONTRACT ANALYSIS

One common trait found in content platform contracts is the absence of robust data privacy language surrounding personal data. Out of the contracts analyzed for the SpringerLink report, only two contracts—those from the University of California and Iowa State University—directly addressed personally identifiable information. The following language comes from Section 8 of the California Digital Library contract with Springer Nature:

The Licensor agrees that no personally identifiable information, including but not limited to log-ins recorded in system logs IP addresses of patrons accessing the system, saved searches, usernames and passwords, will be shared with third parties, except in response to a subpoena, court order, or other legal requirement. If Licensor is compelled by law or court order to disclose personally identifiable information of Authorized Users of patterns of use, Licensor shall provide the Licensee with adequate prior written notice as soon as is practicable, so that Licensee or Authorized Users may seek protective orders or other remedies. Licensor will notify Licensee and Authorized Users as soon as is practicable if the Licensor's systems are breached and the confidentiality of personally identifiable information is compromised.²⁰

20 CDL contract available at https://cdlib.org/services-groups/collections/licensed_resources/redacted_licenses/Springer_Nature_Transformative_Agreement_signed_20210114_Redacted.pdf; Similar language can be found in Section 4.3 of the Iowa State contract available at <https://sparcopen.org/wp-content/uploads/2021/08/Springer-Nature-2021-2023-Iowa-State.pdf>.

In addition, only one contract contained specific information regarding personal data in watermarking²¹:

[from Section 4.15] These watermarks shall not contain user-related information, including but not limited to an account number, IP address, and usernames. If digital watermarking technology is implemented, Licensor will notify Licensee at least thirty (30) days in advance of implementation, and Licensor will provide the technical specifications for the technology used.²²

Other contracts lacked this specific language around personally identifiable information. Instead, the majority of contracts contained the following language around “usage data.” The following is a representative example of the language around “usage data” found in these contracts:

4.3 Where feasible, Licensor shall collect data on usage of the Content and process these according to the COUNTER Code of Practice and according to applicable privacy and data protection laws (the “Usage Data”). The Usage Data will be made available for download by Licensee or third-party harvesting service contracted by Licensee, with single password log-in through a secure website, provided that these statistics are strictly for the Licensee’s own internal use and Licensor shall not be required to disclose any information to the Licensee which it is prohibited from disclosing to the Licensee due to any legal or regulatory constraint imposed upon it, including without limitation any applicable privacy or data protection legislation or regulations or contractual obligations.²³

The organization referenced in the example language, COUNTER, is a non-profit organization that provides a reporting framework to enable consistent usage statistics across different vendor platforms.²⁴ The Code of Practice details the collection, processing, and reporting of data about the usage of electronic resources.²⁵ COUNTER reports range from the database level to the item level. Some COUNTER reports, such as Item Reports, rely on some personal data to calculate unique requests via user session:

A user session is defined any of the following ways: by a logged session ID + transaction date, by a logged user ID (if users log in with personal accounts) + transaction date + hour

21 Watermarks in this context refer to the data encoded in a digital file that contains information about the license, copyright, and the licensed user(s) who downloaded, printed, and/or accessed the file.

22 CDL contract available at https://cdlib.org/services-groups/collections/licensed_resources/redacted_licenses/Springer_Nature_Transformative_Agreement_signed_20210114_Redacted.pdf.

23 Oregon State University Contract, <https://sparcopen.org/our-work/big-deal-knowledge-base/contracts-library/>.

24 COUNTER Metrics. “About COUNTER.” <https://www.countermetrics.org/about/>.

25 “COUNTER Code of Practice Release 5.0.2 — COUNTER Code of Practice Release 5.0.2 Documentation.” <https://cop5.projectcounter.org/en/5.0.2/>.

of day (day is divided into 24 one-hour slices), by a logged user cookie + transaction date + hour of day, or by a combination of IP address + user agent + transaction date + hour of day.²⁶

In the case of COUNTER, a user session can be defined by a user logging into their account at a particular time. If a user is not logged in, then a user session can be calculated through an IP address, the user agent (i.e., browser information), and time of user activity. In both cases, the personal information required to calculate a unique user session according to COUNTER is only a small subcollection of the personal information that is collected and tracked when using SpringerLink. The “Investigating Data Collection and Tracking” section provides further detail as to the extent of personal information collection and tracking. The contract language surrounding “usage data” is vague enough that contract evaluators must not assume that all personal data collected, processed, and disclosed by SpringerLink is covered under the term “usage data.”

POLICIES AND PUBLIC DOCUMENTATION

As we found in the ScienceDirect Report, the contracts included in this analysis have very little information regarding SpringerLink’s data privacy practices or responsibilities. And, as with the ScienceDirect Report, this report relies heavily on documents outside signed legal contracts to learn more about SpringerLink’s data privacy practices. SpringerLink presents the SpringerLink Privacy Policy when we click on the “Privacy Policy” link in the footer of the SpringerLink home page. However, if users are on another Springer Nature platform, such as BioMed Central, and click on the “privacy statement” link in the footer, users will be presented with a privacy policy that is specific to BioMed Central.²⁷ Springer Nature platforms have privacy statements that are specific to their platforms; therefore, someone who might be familiar with BioMed Central’s privacy policy will need to familiarize themselves with the SpringerLink Privacy Policy and not assume that both platforms have the same content in their privacy policies.

An additional layer of privacy policies involves Springer Nature’s privacy policy, accessible through the “privacy” link in the footer of Springer Nature’s website. The Springer Nature privacy policy is separate from the platform privacy policies. On top of all these policies is yet another privacy policy in the form of the Springer Nature Profile privacy policy for users who create accounts with Springer Nature. In the case of SpringerLink, the three privacy policies—SpringerLink, Springer Nature, and Springer Nature Profile—provide similar but not exactly the same information regarding data privacy practices.

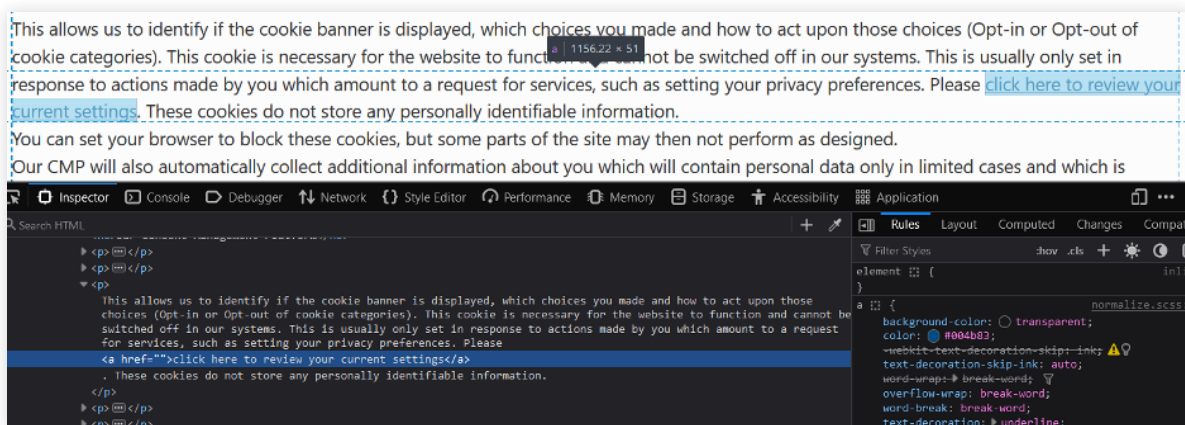
26 Project COUNTER. “7.3 Counting Unique Items,” <https://cop5.projectcounter.org/en/5.0.2/07-processing/03-counting-unique-items.html>.

27 BioMed Central. “Privacy Statement,” May 2, 2024. <https://www.biomedcentral.com/privacy-statement>.

PRIVACY POLICY AND PRIVACY MANAGEMENT – BROKEN FUNCTIONALITY AND ERRORS

At time of publication, the SpringerLink Privacy Policy suffered from numerous broken links and copyediting errors that further reduce the policy's clarity. While these basic site functionality issues may be unintentional, they nevertheless present users with a frustrating user experience when trying to manage their privacy choices. For example, clicking on the "click here to review your current settings" under "Section VIII. Analytics" (where the policy discusses cookie preferences) does not open the cookie management window, but instead leads the reader back to the top of the privacy policy page due to the link target being left empty.

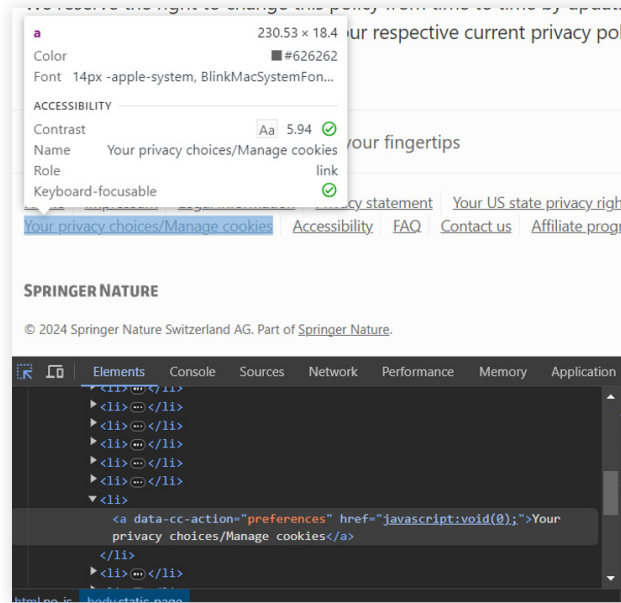
FIGURE 2



The Inspector tab in Web Developer tools showing the code from part of the SpringerLink Privacy Policy web page, with the "a href" HTML tag for the "click here to review your current settings" link left empty.

Another example of a broken link is the "Your privacy choices/Manage cookies" link in the footer of the SpringerLink Privacy Policy page. Clicking the link does not open the cookie management window, nor does it do anything else as the link is coded as `javascript:void(0)`, allowing the text to appear as a link but does nothing when clicked.

FIGURE 3



The Inspector tab in Web Developer tools showing the code from part of the SpringerLink Privacy Policy web page, with the "a href" HTML tag for the "click here to review your current settings" link left empty.

In fact, clicking on "manage cookies" links throughout SpringerLink (and the Springer Nature policy pages included in this analysis) may or may not actually bring up the cookie management window, which is problematic for reasons we will explain in the next section.

In addition to the broken links, readers must navigate through copyediting errors, such as misspellings (e.g., "AdSens" for Google AdSense in Section VII), and repeated text (e.g., repeated sentences in Section VI).

FIGURE 4

VI. Automated decision making

We do not use your personal data for automated decision making which produces legal effects concerning you or similarly significantly affects you; however we do use your personal data to offer you content and services which we believe may be of interest.

We do not use your personal data for automated decision making which produces legal effects concerning you or similarly significantly affects you, however we do use your personal data to offer you content and services which we believe may be of interest.

Section VI of the SpringerLink Privacy policy in its entirety, which repeats the same sentence twice.

PRIVACY POLICY – USER CONSENT CONCERNS

Beyond the broken functional level, the SpringerLink Privacy Policy presents a wall of dense text that is likely to confuse readers who are not accustomed to navigating legalistic language in privacy policies. The November 2023 version²⁸ of the Privacy Policy is over 730 sentences long, with approximately 9000 words in total.^{29,30} Assuming an average reading speed of 238 words per minute,³¹ it would take the average user almost 40 minutes to read through the Policy. While the Privacy Policy has some detail about how SpringerLink collects, processes, and discloses personal data, the specifics are lost within the wall of text, particularly around the collection, use, and disclosure of data for marketing and advertising purposes.

Additionally, there is repetition of a key passage in SpringerLink's Privacy Policy that states they are not responsible for either the privacy of personal data collected by third parties through plug-ins or for the purposes for which the data is processed (more on why this is a critical point in the next section). The exact same sentence is used twice: once in Section IX.1. about social media plug-ins (which is to be expected) and again in Section VII, where there is no mention of plug-ins.³² Section VII mentions cookies and web beacons, though, so it would be reasonable for users to interpret "plug-in" to refer to "cookies and web beacons." Nevertheless, it should not be left to library workers, users, and other members of the general public to guess at such an important component of the policy.

US CONSUMER-SPECIFIC DATA COLLECTION AND DISCLOSURE CONCERNS

The ScienceDirect report found that Elsevier's policy specific for US consumers (i.e., U.S. Consumer Privacy Notice) had more detailed and concise information about data collection, processing, and disclosure of personal information, but readers of the ScienceDirect Privacy Policy could easily miss the link to the supplemental policy in the policy text. The equivalent policy for Springer Nature, confusingly titled "California Privacy Rights Act" (CPRA), even though it also covers other states, receives no mention in the main text of the Privacy Policy. The only indication that such a policy exists is a link in the footer of the web page of the policy labeled "Your US state privacy rights," which directs the user to the Springer Nature website (and therefore may put any collection of personal data from visiting the page under the Springer Nature Privacy Policy and not the SpringerLink

28 While the readability and length analysis were performed on the November 2023 version of SpringerLink's Privacy Policy, our findings here (and across other sections of this report) are similarly applicable to the most recent version at the time of publication (posted in June 2024).

29 The word count estimate is from a Microsoft Word document containing the SpringerLink Privacy Policy text.

30 For comparison, SpringerLink's Privacy Policy is more than double the word count for the 2023 version of Elsevier's Privacy Policy used in the ScienceDirect Report, which was approximately 3750 words.

31 Brysbaert, Marc. "How Many Words Do We Read per Minute? A Review and Meta-Analysis of Reading Rate." *Journal of Memory and Language* 109 (December 1, 2019): 104047. <https://doi.org/10.1016/j.jml.2019.104047>.

32 "Please note that we neither have the control of the extent of personal data that is collected by the respective plug-in provider nor do we know the processing's purpose or the period your personal data will be retained."

policy³³). Despite additional broken links in the footer to manage cookies and the initial impression that the policy is only for California residents (the end of the policy has a supplemental notice for other states with data protection regulations), the policy provides a clearer picture of the overwhelming amount of data that Springer Nature may collect. Springer Nature CPRA Policy states the categories user data collected, including:

- **Identifiers:** identifiers such as a real name, alias, postal address, unique personal identifier (such as a device identifier; cookies, beacons, pixel tags, mobile ad identifiers and similar technology; customer number, unique pseudonym, or user alias; and other forms of persistent or probabilistic identifiers), online identifier, internet protocol address, email address, telephone number, account name, classifications under California or federal law (e.g. race, color, national origin, religion, age, sex, gender, gender identity) and other similar identifiers.
- **Online Activity:** Internet and other electronic network activity information, including, but not limited to, browsing history, search history, and information regarding your interaction with websites, applications or advertisements.
- **Internet or other similar network activity:** Browsing history, search history, information on a consumer's interaction with a website, application, or advertisement.
- **Protected classification characteristics under California or federal law:** Age (40 years or older), race, color, ancestry, national origin, citizenship, religion or creed, marital status, medical condition, physical or mental disability, sex (including gender, gender identity, gender expression, pregnancy or childbirth and related medical conditions), sexual orientation, veteran or military status, genetic information (including familial genetic information).
- **Biometric information:** Genetic, physiological, behavioral, and biological characteristics, or activity patterns used to extract a template or other identifier or identifying information, such as, fingerprints, faceprints, and voiceprints, iris or retina scans, keystroke, gait, or other physical patterns, and sleep, health, or exercise data.
- **Geolocation Data:** Physical location or movements.
- **Inferences:** inferences drawn from any of the information identified above to create a profile about you reflecting your preferences, characteristics, psychological trends, predispositions, behavior, attitudes, intelligence, abilities, and aptitudes.³⁴

The Springer Nature CPRA Policy lists the types of uses for the personal information citing “business purposes” as specified under the California law, including marketing and advertising purposes. In Section 2 of the policy, we learn that Springer Nature collects personal information from a wide range of sources, including from data brokers. These data sources also include “...joint marketing partners, ...publishing partners, data

33 At the time of the analysis, the SpringerLink and Springer Nature Privacy Policies are similar in nature but both contain unique information, nor is it guaranteed that this similarity will continue depending on future policy updates.

34 Springer Nature. “California Privacy Rights Act.” <https://www.springernature.com/gp/legal/ccpa>.

analytics providers, government entities, ...[and] social networks.”³⁵ Collecting personal information from a wide range of third party vendors including data brokers raises the question of how this personal data is used. The SpringerLink Privacy Policy only mentions user profiling in relation to “(behavioural) Advertising and profiling” under processing personal data and in the use of Hotjar, a behavior and analytics tool.³⁶ Beyond the profile created through inference data, there is little transparency as to what data is collected from which source and for what specific purpose.

The following sections 3 and 4 of the CPRA Policy provide an example for library workers when researching whether a vendor “sells” personal data. Springer Nature states in Section 3 that they do not sell personal information, but then makes the clarification that they do not sell personal information “in exchange for monetary consideration.”³⁷ The company goes on to explain that personal data such as identifiers, online activity, and inferences are provided to online advertisers and social networks, which **may** be considered a “sale” under the California Consumer Privacy Act (CCPA) and CPRA. Section 4, conversely, lists all the categories of personal information that is shared with (rather than sold to) third parties, including identifiers, online activity, and inferences which are shared with affiliates, marketing partners, data analytics providers, and government entities.³⁸ While the statement “we do not sell your personal information” may seem to look out for the best interest of the end-user—particularly when this statement is made in communications to the public or media—the reality for many vendors is that personal information is collected by or disclosed to third parties—no monetary exchange required.³⁹

A BRIEF NOTE ABOUT THE LEGAL BASIS FOR PROCESSING PERSONAL DATA – ARTICLE 6 OF GDPR AND “BUSINESS PURPOSES” UNDER CCPA

The SpringerLink Privacy Policy is littered with the phrase “Article 6 sec. 1 sent. 1 lit. [a-f] GDPR” as an explanation for the data collection, tracking, disclosure, and processing activities listed in the Policy. Similarly, the Springer Nature CPRA Policy cites the CCPA’s “business purposes” as justification for the processing of personal information listed in the policy. While this analysis is not a legal review of Springer Nature’s data practices, nor is it an analysis to determine if a vendor like Springer Nature is complying with data protection regulations such as GDPR or CCPA, it is still worthwhile to briefly discuss what library workers and others should consider when evaluating such statements in vendor policies, contracts, and documentation.

35 Ibid.

36 SpringerLink. “Privacy Statement.” <https://link.springer.com/privacystatement>.

37 Springer Nature. “California Privacy Rights Act.” <https://www.springernature.com/gp/legal/ccpa>.

38 Ibid.

39 Related to the ScienceDirect report and the policies reviewed within, the Elsevier U.S. Consumer Privacy Notice (<https://www.elsevier.com/legal/us-consumer-privacy-notice>) states that Elsevier does not sell personal information in Section 3. Nevertheless, Elsevier still shares personal information with third parties for behavioral and targeted advertising purposes per Section 4. Therefore, a public statement saying that a company does not sell personal information may be technically true in regards to a “sale” defined by U.S. data protection laws, but does not account for the fact that personal information is still disclosed to third parties for purposes that consumers may find objectionable.

Article 6 of GDPR⁴⁰ and the definition of “business purpose” in CCPA⁴¹ detail the types of processing activities allowed under their respective data protection regulations. However, despite this attempt to narrow the specific types of data processing through regulation, these lists of lawful or allowed data processing contain loopholes that can be, and in some cases are, exploited:

- In the case of GDPR:
 - » Article 6 (1)(a) allows for “lawful” processing of personal data if the person (“data subject”) gives consent for “one or more specific purposes.”⁴² GDPR provides the conditions needed to obtain consent⁴³ but at the same time consent can also be gamed, such as claiming that using a service with a Terms of Service that was not read by a website user is considered consent.⁴⁴
 - » Article 6 (1)(f) allows for “lawful” processing of personal data “for the purposes of the legitimate interests pursued by the controller or by a third party”⁴⁵ but “legitimate interests” includes actions such as direct marketing⁴⁶ and the broadness of the definition can lead to deceptive design practices.⁴⁷
- In the case of CCPA, “business purposes” includes several advertising, analytics, and marketing activities.⁴⁸

Put another way, complying with GDPR or CCPA requirements does not automatically equate to meeting library privacy standards, guidelines, and best practices. It is prudent for library workers to be aware of loopholes in data protection regulations whenever they encounter statements regarding “lawful processing,” “legitimate interests,” and “business purposes” in vendor data privacy policies, contracts, and other documents.

40 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance) (2016). <http://data.europa.eu/eli/reg/2016/679/2016-05-04/eng>.

41 “California Consumer Privacy Act of 2018,” July 15, 2024. https://coppa.ca.gov/regulations/pdf/coppa_act.pdf, pg. 21.

42 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance) (2016). <http://data.europa.eu/eli/reg/2016/679/2016-05-04/eng>.

43 Ibid.

44 Martin, Baily. “Are Terms of Service a Loophole in GDPR Consent?” *Georgetown Law Technology Review*, January 23, 2021. <https://georgetownlawtechreview.org/are-terms-of-service-a-loophole-in-gdpr-consent/GLTR-01-2021/>.

45 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance) (2016). <http://data.europa.eu/eli/reg/2016/679/2016-05-04/eng>.

46 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance) (2016). <http://data.europa.eu/eli/reg/2016/679/2016-05-04/eng>, as cited in “GDPR and the “Legitimate Interests” Loophole.” <https://www.andrew.legal/blog/2018/3/27/gdpr-and-the-legitimate-interests-loophole>.

47 Kyi, Lin, Sushil Ammanaghatta Shivakumar, Cristiana Teixeira Santos, Franziska Roesner, Frederike Zufall, and Asia J. Biega. “Investigating deceptive design in GDPR’s legitimate interest.” In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, pp. 1-16. 2023, <https://dl.acm.org/doi/full/10.1145/3544548.3580637>.

48 “California Consumer Privacy Act of 2018,” July 15, 2024. https://coppa.ca.gov/regulations/pdf/coppa_act.pdf, pg 21.

Library workers should also be aware that these regulations may not offer the level of privacy protections necessary to meet the standards and best practices of the library profession, as well as certain US-state laws about library privacy. Finally, library workers should understand that by the time data protection bills become laws, the rapid changes in technology and data-centered business models blunt the impact of many of the data protections offered by these regulations.

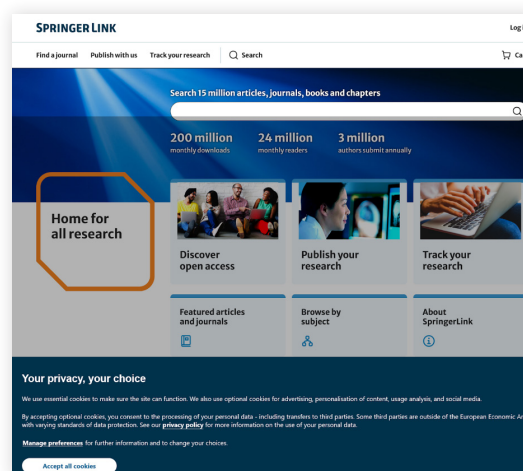
SPRINGERLINK COOKIE MANAGEMENT – ISSUES AND CONCERNS^{49,50}

Functionality and Design

SpringerLink's Cookie Policy, like ScienceDirect's Cookie Notice, does not mention the other technologies that collect user data and track users across SpringerLink. Some additional information about third party trackers and cookies can be found in the SpringerLink Privacy Policy, albeit buried in the middle of a lengthy dense document. Information is sparse on third party cookies and other tracking technologies in SpringerLink's Cookie Policy; however, the policy references a "Privacy Preference Centre" where users can block or "deactivate" cookies on the site.⁵¹ Unfortunately, at the time of publication, the link the policy redirects users to in the footer of the SpringerLink Cookie Policy page is broken like the footer link in the SpringerLink Privacy Policy.

Regardless of link functionality, all end users located in the US should encounter the cookie banner shown in Figure 5 when visiting SpringerLink for the first time.

FIGURE 5



The SpringerLink home page, with a translucent black screen overlay save the cookie banner at the bottom of the page.

49 It is important to note that the cookie banner for SpringerLink differs by locality. This report's analysis was conducted in the United States; therefore, we are analyzing the design, content, and functionality of the US version of the cookie banner. A brief test of the SpringerLink homepage using a VPN pointing to a server in Germany showed a different banner design and consent choices for users based in Germany.

50 Readers may notice the differences between several of the cookie banner screenshots used in this report. Some elements of the US version of SpringerLink's cookie banner changed during the analysis, primarily the toggles for turning on and off specific cookie categories (i.e., radio buttons or sliders) and the three buttons at the bottom of the cookie management window (i.e., button order and color). The cookie banner first presented to users based in the US did not have noticeable changes throughout the analysis time period.

51 "Cookies Consent." <https://link.springer.com/cookiepolicy#preference-centre>.

The cookie banner acts as a wall that prohibits the user from searching or navigating the website, forcing the user to go through the cookie wall before using the site.⁵² The most prominent design feature of the cookie banner is the “Accept all cookies” button, highlighted in white against a dark background. End users who don’t instinctively (or unconsciously) click on the “Accept” button as they enter the site might notice the “Manage preferences” in smaller font above the button. The smaller link does not stand out like the big white button on the dark background, which could leave users the first impression that they have to accept all cookies to use the site.

This design choice is an example of a “deceptive pattern,” a design term⁵³ describing user interface design choices that influence users into making choices or doing something that they did not fully understand or consent to or, conversely, discourage user choices that would negatively impact a business or organization. In the case of Springer Nature’s cookie banner, the design choice to only offer one big button that is sharply contrasted from the rest of the text, including the link for users to reject optional cookies, is an example of several deceptive patterns such as visual interference (hiding or otherwise obscuring information through low contrast and small text)⁵⁴ and misdirection (designing to focus the user’s attention on one choice or option, distracting users from paying attention to another choice or option).⁵⁵ Researchers in multiple studies have also found that deceptive designs in cookie consent banners increases acceptance of all cookies and decreases the chances of users rejecting optional cookies.⁵⁶ In addition, the requirement for users to navigate to an additional screen with multiple clicks in order to reject some or all optional cookies—as well as access information about the cookies being used to collect personal data—falls into the deceptive patterns of “sneaking” (withholding or otherwise obscuring relevant information from users)⁵⁷ and “obstructing” via “longer than necessary” (requiring additional steps in order for a user to access or use settings around data

52 At the time of the analysis, clicking on the Privacy Policy link in the cookie banner message redirects users to the SpringerLink Privacy Policy page sans cookie banner, where users will encounter the broken links for managing cookies in the main text and footer of the page. However, clicking on the Privacy Policy link in the cookie banner on the Springer Nature main site will take users to the Springer Nature Privacy Policy, which is obscured behind the cookie banner.

53 Like the term consent, we use the phrase “deceptive pattern” in a technical and user experience context, where it is a term of art, and not in the context of Springer Nature’s compliance to any state or federal regulation in the United States.

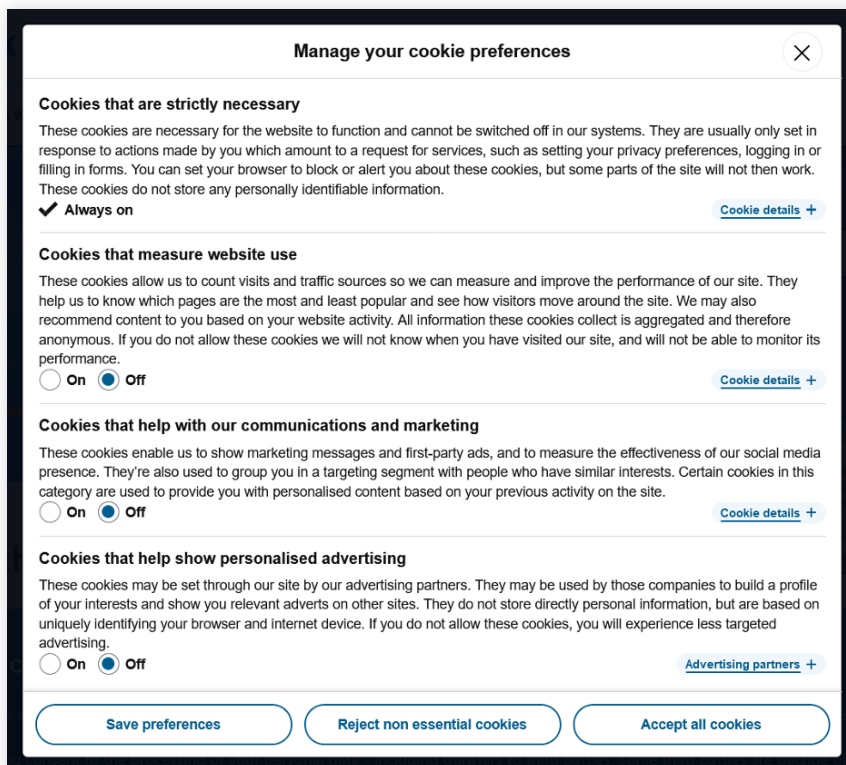
54 Deceptive Patterns. “Visual Interference.” <https://www.deceptive.design/types/visual-interference>.

55 Mejtoft, Thomas, Erik Frångsmyr, Ulrik Söderström, and Ole Norberg. “Deceptive Design: Cookie Consent and Manipulative Patterns.” BLED 2021 Proceedings, January 1, 2021. <https://aisel.aisnet.org/bled2021/36/>.

56 Nouwens, Midas, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. “Dark Patterns after the GDPR: Scraping Consent Pop-Ups and Demonstrating Their Influence.” In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–13, 2020. <https://doi.org/10.1145/3313831.3376321>; Graßl, Paul, Hanna Schraffenberger, Frederik Zuiderveen Borgesius, and Moniek Buijzen. “Dark and Bright Patterns in Cookie Consent Requests.” *Journal of Digital Social Research* 3, no. 1 (February 8, 2021): 1–38. <https://doi.org/10.33621/jdsr.v3i1.54>; Bielova, Nataliia, Laura Litvine, Anyisia Nguyen, Mariam Chammat, Vincent Toubiana, and Estelle Hary. “The Effect of Design Patterns on (Present and Future) Cookie Consent Decisions,” 2813–30, 2024. <https://www.usenix.org/conference/usenixsecurity24/presentation/bielova>.

57 Deceptive Patterns. “Sneaking.” <https://www.deceptive.design/types/sneaking>.

FIGURE 6



The SpringerLink cookie preference screen in late 2023.

privacy).⁵⁸ All of these deceptive patterns impact an end-user's choice to accept all cookies or not—an impact that may have a much larger consequence than the user realizes as we will discuss in the next section.

For those users who click on the “Manage preferences” link in the cookie banner, they are then presented with the “Manage your cookie preferences” pop up window shown in Figure 6.

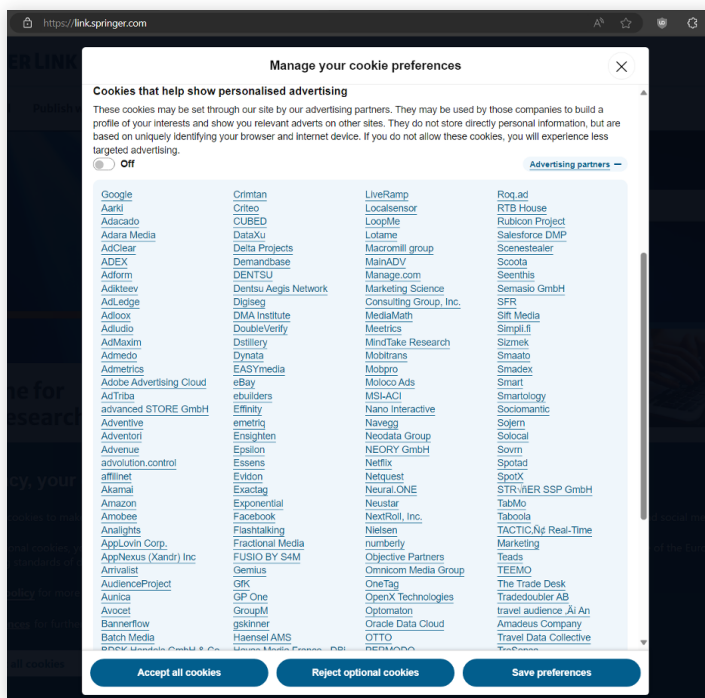
Users are given three options: accept all cookies, reject all optional cookies, or customize which cookies are allowed or not and save their preferences. The first three categories of cookies—strictly necessary, measuring website use, and communications and marketing—provide a high-level summary about how a person's data is used by the cookies under a particular category. These categories provide additional information under “Cookie details” with specific cookies, and the type of data being collected by each cookie.

58 European Data Protection Board. “Guidelines 03/2022 on Deceptive Design Patterns in Social Media Platform Interfaces: How to Recognise and Avoid Them.” European Data Protection Board, February 14, 2023. https://www.edpb.europa.eu/system/files/2023-02/edpb_03-2022_guidelines_on_deceptive_design_patterns_in_social_media_platform_interfaces_v2_en_0.pdf.

200 Advertising Partners – The Limits of User Notice and Consent

Springer Nature breaks the notice and consent model⁵⁹ of tracking permission by making it impossible for users to fully understand the terms they are agreeing to. The true extent of the consequences of clicking on “Accept all cookies” is only revealed when a user clicks on the “Advertising partners” link under the personalized advertising cookie category, where the user is presented with a list of 200 companies shown in Figure 7.

FIGURE 7



A partial list of the advertising partners included in the personalized advertising section of the SpringerLink cookie preference screen (US version).

These partners include advertisers, data brokers, data intelligence companies, and retailers. The companies listed in this category range from the well-known—Google, Amazon, Netflix, and Facebook—to companies that an average user or library worker has likely never heard of. Clicking on the individual companies’ names sometimes brings users to that company’s privacy policy but also sometimes results in a “not found” 404 page or otherwise unresponsive site. A full list of advertising partners is in Appendix D with notes regarding broken links found at the time of analysis.

59 The notice and consent model is largely formed from the “Notice/Awareness” and “Choice/Consent” principles from the FTC Fair Information Practice Principles (<https://web.archive.org/web/20090331134113/http://www.ftc.gov/reports/privacy3/fairinfo.shtm>).

There are several possible reasons why Springer Nature allows these partners to set cookies on the SpringerLink website. Web analytics applications such as Google Analytics and Hotjar use trackers such as cookies (as noted in the other tracker categories in the cookie preference screen) to monitor website traffic, which often includes collecting user data. Social media trackers included in plugins embedded into the website allows for easier sharing of articles on social media platforms but also enables these companies to track users outside of those platforms. Nevertheless, the large number of advertisers, marketers, and data brokers in Springer Nature's partner list indicates likely financial incentives for Springer Nature to include trackers from these partners in SpringerLink. Social media platforms use data collected from trackers to, among other purposes, provide targeted ads which contribute to a company's revenue stream. Trackers from advertisers, marketers, and data brokers, on the other hand, can also provide income or, in the case of Springer Nature, exchange personal information for non-monetary consideration as noted in the CPRA Policy to the site hosting said trackers. In addition to advertisements, the information collected from these trackers—used to create profiles from information ranging from unique digital fingerprints to browsing behavior which may reveal sensitive personal information—can be fed into the data economy, including being sold by data brokers to other brokers or interested buyers.

While many of the larger, more well-known companies in the list have received public attention from various publications and legal actions regarding their problematic privacy practices,⁶⁰ users and library workers alike may miss lesser-known companies with similar problematic practices that may result in real world harm to users. For example, the list includes Kochava, a data broker that the Federal Trade Commission is (at the time of publication) suing for allegedly selling geolocation data collected from millions of mobile devices, which could be used to track the movements of unique individuals to and from sensitive locations such as domestic abuse and homeless shelters and reproductive health clinics.⁶¹ It is not reasonable to expect an average user, library worker, or even institutional counsel to know all two hundred companies in this list, nor is it reasonable to expect the same people to be aware of all the problematic data privacy practices of over two hundred companies.

Springer Nature states that the cookies *may* be set by their advertising partners, which means that users and library workers are left guessing exactly which third parties are receiving their personal information. Even though the summary for this cookie category states that the cookies in this category do not “directly [store] personal information,” the rest of the sentence lists data points that, when combined with other data

60 These companies include Amazon (e.g. <https://www.wired.com/story/amazon-failed-to-protect-your-data-investigation/> and <https://www.politico.eu/article/data-at-risk-amazon-security-threat/>), Google (e.g. <https://www.reuters.com/legal/google-settles-5-billion-consumer-privacy-lawsuit-2023-12-28/> and <https://www.ftc.gov/legal-library/browse/cases-proceedings/102-3136-google-inc-matter>), and Meta (e.g. <https://www.nbcnews.com/tech/social-media/timeline-facebook-s-privacy-issues-its-responses-n859651> and <https://www.npr.org/2022/12/23/1145303268/facebook-meta-cambridge-analytica-privacy-settlement>).

61 Federal Trade Commission. “FTC v Kochava, Inc.,” August 29, 2022. <https://www.ftc.gov/legal-library/browse/cases-proceedings/ftc-v-kochava-inc>.

points collected about the same user, creates a digital fingerprint that can identify an individual.⁶² Since the data collected by third parties is also processed and stored by those third parties and is not governed by the SpringerLink Privacy Policy or contract language, users should be concerned about how these third parties may use personal data collected through the SpringerLink website—including the possibility of reidentifying the users from deidentified data.

Users wishing to change their cookie preferences on SpringerLink during the same browsing session may be met with broken links depending on which page they may be on the website, which may lead to confusion and frustration on the part of the user who may not realize that the link works on some pages but not others. SpringerLink is not the only vendor that uses a cookie banner for notice and consent purposes, but they are an example of how vendors use tools that, in theory, are supposed to provide users control over their privacy choices but instead “nudge” users into making choices that primarily benefit the company and its third party partners.

Combined with the deceptive designs present in the cookie banner, the extensive list of advertising partners shows the limits of the notice-and-consent model for managing user privacy. Users who need to access an article for research may click on the “Accept all cookies” button without realizing the amount of tracking they just “agreed” to so they can access the article. Library workers who are doing research on SpringerLink’s data privacy practices will most likely not have the resources and time to check the privacy policies of 200 advertising partners in the cookie management window.

In both scenarios, library workers, users, and institutional counsel interested in reading every privacy policy for all advertising partners would face a daunting task. If, for example, each policy were roughly the same length as SpringerLink’s Privacy Policy (approximately 9,000 words)⁶³, the total amount of reading would amount to 1.8 million words, just a little more than the combined word count of the first five volumes of George R.R. Martin’s *A Song of Ice and Fire* series.⁶⁴ At an average reading speed,⁶⁵ this task would require more than three full weeks solely dedicated to this task—not factoring in the time required to re-read policies anytime they are revised.

62 Re-identification of aggregated or de-identified data can include linking multiple data sets (e.g., Netflix [<https://arxiv.org/abs/cs/0610105>], NYC Taxis [<https://perma.cc/5LZG-YZM8>]), and combining quasi-identifiers in the same data set (e.g., AOL [<https://www.nytimes.com/2006/08/09/technology/09aol.html>]).

63 The word count estimate is from a Microsoft Word document containing the SpringerLink Privacy Policy text.

64 Davis, Sarah S. “The Word Count of 175 Favorite Novels.” Broke by Books (blog), May 12, 2019. <https://brokebybooks.com/the-word-count-of-175-favorite-novels/>; A Song of Ice and Fire Wiki. “A Song of Ice and Fire/Statistical Analysis.” https://iceandfire.fandom.com/wiki/A_Song_of_Ice_and_Fire/Statistical_Analysis.

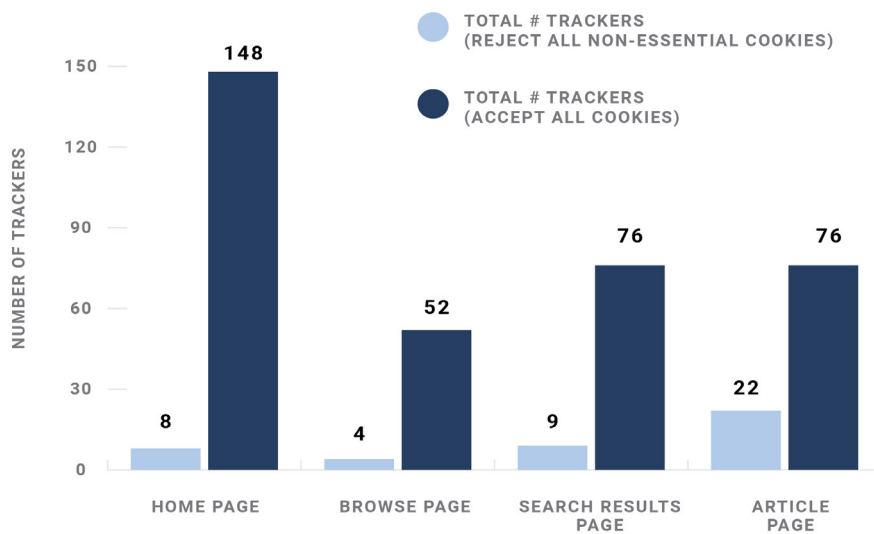
65 The average reading speed for English is 238 words per minute. See Brysbaert, Marc. “How Many Words Do We Read per Minute? A Review and Meta-Analysis of Reading Rate.” *Journal of Memory and Language* 109 (December 1, 2019): 104047. <https://doi.org/10.1016/j.jml.2019.104047> for a detailed explanation.

INVESTIGATING DATA COLLECTION AND TRACKING ON THE SPRINGERLINK WEBSITE

The cookie banner had a significant impact on how we tested the SpringerLink website, as covered in the methodology section. This impact, nevertheless, enabled us to test the website in such a way that we can demonstrate the stark contrast in data collection and tracking between accepting all cookies and rejecting all optional cookies, including the cookies that may send user data to the 200 advertising partners listed in the cookie management window in addition to other companies not listed in the window but found in our front-end website analysis.

As in our 2023 ScienceDirect analysis, our SpringerLink analysis takes a multifaceted approach to test front-end website data collection and tracking. Popular browser plugins such as uBlock Origin and NoScript grant quick access to the list of cookies and executable content (e.g., JavaScript) found on a web page. Built-in browser web developer tools provide a more detailed view of real-time network traffic for a specific page, including what data trackers and scripts are requesting and collecting. More advanced tools such as Website Evidence Collector (WEC) offer a better overview of the types of tracking present on multiple pages across the SpringerLink website and the information being collected, stored, and transmitted. WEC output several text files after running each tool. These files contain lists of specific trackers on the site and overall inspection report files. WEC also produces a summary report of the findings in a more friendly, human-readable HTML file.

FIGURE 8



The comparison of total number of trackers by web page as reported by WEC.

The WEC reports for the two sets of tests—accepting all cookies and rejecting all optional cookies—showed a substantial difference in the number of trackers and cookies on all four URLs tested for the analysis, with the highest disparity being on the SpringerLink home page.

While there are still trackers from the “necessary” cookie category in the “reject all optional cookies” WEC tests, the tracking came nowhere near the extent of the “accept all cookies” tests. Several domains attached to trackers and cookies listed in the “accept all cookies” WEC reports are from companies that seem not to be listed in the 200 businesses in the cookie management window nor in any privacy-related documentation reviewed in the analysis, including:

- 360yield.com (Improve Digital)
- ivitrack.com
- mediavine.com
- mediawallahscript.com (Mediawalla)
- postrelease.com (Nativo)
- revcontent.com
- sharethrough.com
- tapad.com⁶⁶

A list of all cookie and web beacon hosts found in the WEC reports can be found in Appendix C.

A particular concern found in the WEC reports for the “accept all cookies” tests is the use of local storage for user tracking and data collection. Local storage stores data in the client web browser, but unlike cookies, that data can persist after user sessions and browser shutdowns. Data in local storage also do not automatically expire after a certain timespan, as cookies do. This persistent storage and lack of automatic deletion enables websites to track users for potentially long periods of time, creating a profile with a detailed history of website activity and user behavior. As mentioned in the ScienceDirect report, local storage (along with web beacons) can bypass common mechanisms and user practices that mitigate surveillance. For example, data in local storage can persist until a user adjusts additional settings in their browser (e.g., in Firefox, setting the browser to delete cookies, site data, and cached web content at browser close). However, this example assumes that most users change default settings, which research shows is not the case.⁶⁷

66 It should be noted that testing the SpringerLink cookie banner using a VPN pointing to a server in Germany showed some, but not all, of these specific companies included in the cookie management window's list of business partners.

67 See also, Groeger, Lena V. “Set It and Forget It: How Default Settings Rule the World.” ProPublica, July 27, 2016. <https://www.propublica.org/article/set-it-and-forget-it-how-default-settings-rule-the-world>.

The WEC reports found multiple local storage objects when a user accepts all cookies on SpringerLink websites. These local storage objects include key entries for Permutive, a data analytics platform aimed at publishers and advertisers. The data being collected in the Permutive keys can uniquely identify an individual and build to bypass common protections to prevent user tracking, such as browsers that can block third party cookies. For example, the `permutive-id` key is used to identify users as long as the local storage on the user's device is not cleared.⁶⁸ Therefore, a SpringerLink user can be identified across multiple sessions, and their multiple-session behavior trail analyzed, using the `permutive-id` unless they manually clear their browser's local storage.

Permutive keys are also being used to collect digital fingerprints of users (i.e., unique browser and system information), and there are additional Permutive keys that collect personal information that identifies not only the Internet Service Provider (ISP) the person is currently using to access SpringerLink, but the zip code where the person is located during the time of the session:

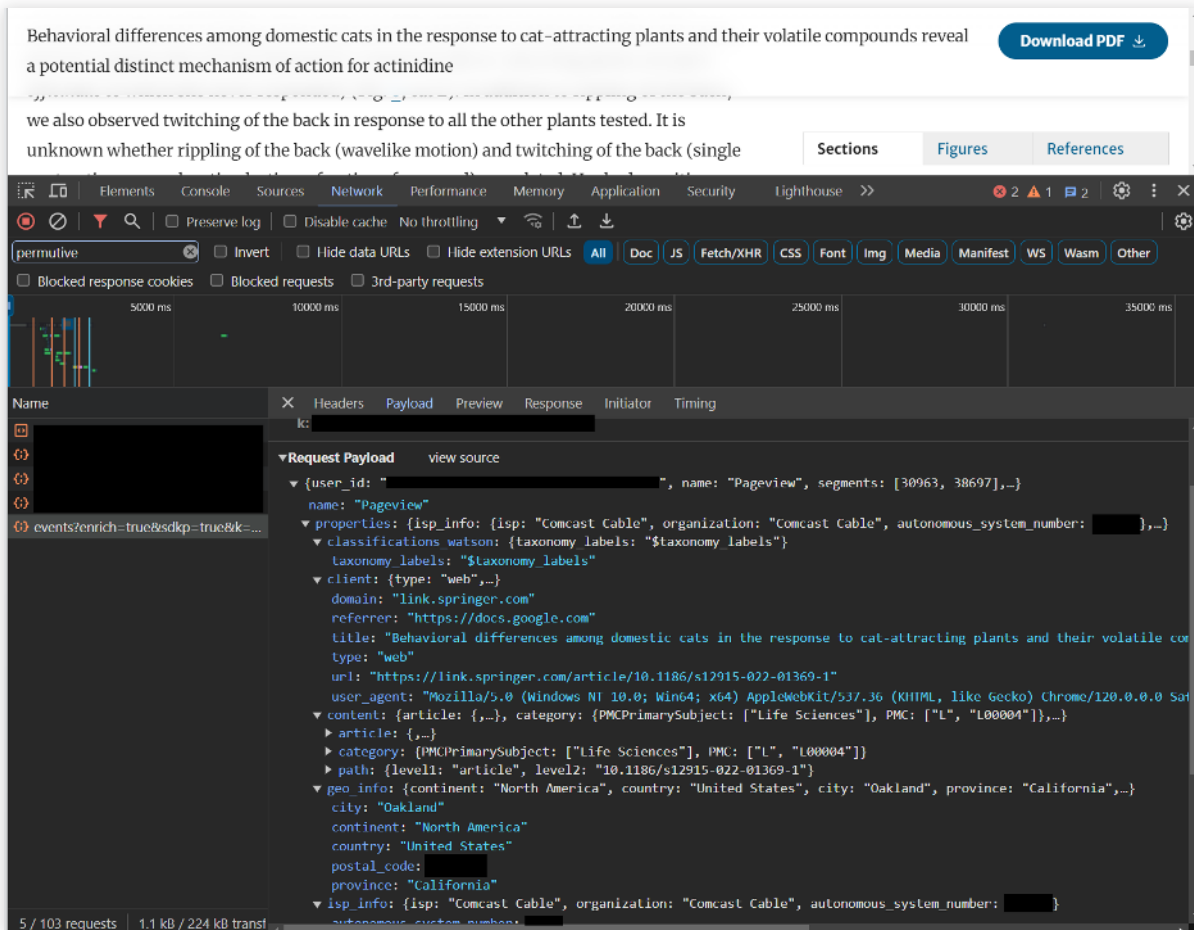
FIGURE 9

10	link.springer.com	permutive-data-enrichers	<pre> { "enricher:geopip": { "session_id": [REDACTED] "data": { "geo_info": { "continent": "North America", "country": "United States", "city": "Oakland", "province": "California", "postal_code": [REDACTED] }, "isp_info": { "isp": "Comcast Cable", "organization": "Comcast Cable", "autonomous_system_number": [REDACTED] "autonomous_system_organization": " } } } } </pre>
----	-------------------	--------------------------	--

The semi-redacted entry for the "permutive-data-enrichers" key from the WEC report on the SpringerLink article page.

68 Permutive. "Q. What Is the Difference between the Permutive ID vs PXID?," October 13, 2022. <https://support.permutive.com/hc/en-us/articles/4650502329500-Q-What-is-the-difference-between-the-Permutive-ID-vs-PXID>.

FIGURE 10



The Network tab in Web Developer Tools showing a partial (and semi-redacted) payload for Permutive trackers on the SpringerLink article page tested for the analysis.

According to Permutive's documentation, the ISP information captured by the Permutive local storage object can be used to group users from a similar ISP together in order to create a "cohort" based on ISP,⁶⁹ which could lead to profiling entire campuses that serve as an ISP for their campus community. Permutive "enriches" this ISP information through data from Maxmind, a "leading provider of GeoIP intelligence."⁷⁰ Permutive

69 Permutive. "Targeting Users by Their Internet Service Provider (ISP)," July 13, 2021. <https://support.permutive.com/hc/en-us/articles/360010209739-Targeting-users-by-their-Internet-Service-Provider-ISP>.

70 Permutive. "Location and ISP Enrichment with Maxmind," September 7, 2023. <https://support.permutive.com/hc/en-us/articles/360010261819-Location-and-ISP-Enrichment-with-Maxmind>.

can also integrate with over 200 “data partners” that can combine data to create profiles of users based on demographic, lifestyles, interests, and behaviors.⁷¹ It is unclear whether Springer Nature’s use of Permutive takes advantage of these integrations—but, even without these data partners, SpringerLink’s use of trackers and receipt of data from Google Analytics, Hotjar, Amplitude, Criteo, and other advertising and data analytics platforms can provide Springer Nature a comprehensive picture of individual users.

An additional concern is the use of the `permutive-data-models` key which appears to be referencing Permutive’s modeling capabilities, where users are assigned to one or more “cohorts” that may be based on behaviors or interests.⁷²

FIGURE 11

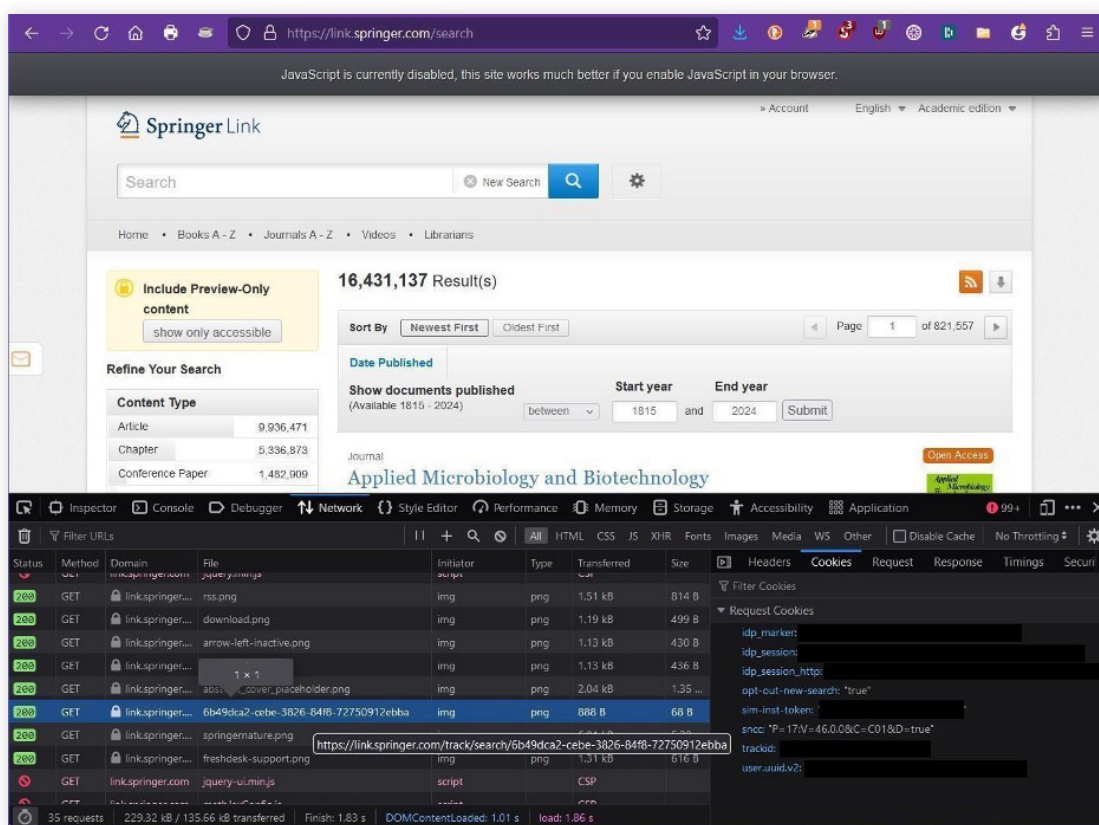
2	link.springer.c...	<code>permutive-data-models</code>	<pre> { "models_data": { [REDACTED] "ip": { [REDACTED] "const": "0" }, "precision": 16, "min": 0.05405351519584656, "max": 0.05936770141124725 } }, "models_last_updated": 17035435827 } </pre>
---	------------------------------------	------------------------------------	---

The “`permutive-data-models`” key and semi-redacted contents from the WEC report on the SpringerLink article page.

- 71 Permutive. “Q. Which Third-Party Data Providers Do You Integrate With?,” July 13, 2021. <https://support.permutive.com/hc/en-us/articles/360010262699-Q-Which-third-party-data-providers-do-you-integrate-with>; see also <https://docs.google.com/spreadsheets/d/1SFJDDb70HVKOXje7kyC6nvVRTax2cX4WYCDnEFJCQY/edit#gid=0> which has 50 data partners listed.
- 72 Permutive. “Introduction to Cohorts,” October 4, 2023. <https://support.permutive.com/hc/en-us/articles/360010174519-Introduction-to-Cohorts>.

It is unclear from the values in the permutive-data-models if Springer Nature is referring to a Classification Model⁷³ or a Lookalike Model⁷⁴ but these models rely on machine learning to analyze behavioral data to identify users and assign them to one or more “cohorts.” Even if the models can only be applied to audiences created with “first party data”—meaning that third party data cannot be used to create models in Permutive—these models are still based off of user data, possibly collected from multiple sessions attached to a permutive-id in local storage that will not automatically delete.

FIGURE 12



The Network tab in Web Developer Tools showing a 1x1 clear gif web beacon on the SpringerLink search results page that contains the “sncc” cookie set by the cookie management system. JavaScript was turned off in the browser at the time of the screenshot.

73 Permutive. “Classification Models Overview,” October 21, 2022. <https://support.permutive.com/hc/en-us/articles/5489960808988-Classification-Models-Overview>.

74 Permutive. “Modeling Explained,” October 24, 2023. <https://support.permutive.com/hc/en-us/articles/360013366399-Modeling-Explained>.

While Permutive keys dominated the local storage lists for the “accept all cookies” WEC tests, the cookies and web beacon lists show a plethora of third party trackers collecting personal information such as digital fingerprints. Some of these trackers—such as trackers from Yahoo Analytics, Bluekai, Microsoft Ads, and Criteo—contain pseudonymous user or session ids. Other trackers such as Google Analytics and R.O.EYE (acquired by Acceleration Partners)⁷⁵ contain digital fingerprint information. This information can be used to construct detailed user profiles, particularly when these trackers remain persistent across multiple user sessions.

For the average user, the extent and scale of user tracking recorded in the front-end website tests most likely does not line up with their expectations when they click on the “Accept all cookies” button on the cookie banner when they first visit SpringerLink. For example, users who are nudged to click the “Accept all cookies” button (see the earlier section about deceptive patterns) and not the small link to manage preferences may miss that their data may be shared with 200 advertising partners—all because they needed to use this particular academic content platform for publications that they may not have access to elsewhere. Users may also be unable to readily manage cookie preferences after their initial selection from the SpringerLink Privacy Policy or other privacy-adjacent policy pages due to broken or redirecting links to manage cookies.⁷⁶ Users can bypass the problematic cookie banner and mitigate the amount of tracking by SpringerLink through turning off JavaScript (JS) while using the site—the banner does not appear when JS is turned off, and searches and full article access are not affected by turning off JS. Testing this option using NoScript and Firefox’s Web Developer Tools shows that turning off JS has the same effect as rejecting all non-essential cookies (see Figure 12 - the **sncc** is set for rejection of all non-essential cookies with JS turned off). However, this option may not be viable on computers or devices that do not allow for user adjustments to browser settings, such as public access computers. Turning off JS also does not eliminate all tracking by SpringerLink, with two pseudonymous tracking ids being present in a cookie in a HTML file.

75 <https://roeye.com/>.

76 At the time of publication the broken links identified earlier in the report are still broken or redirecting to the Privacy Policy.

OVERALL DATA PRIVACY ASSESSMENT

Our analysis indicates that there are no privacy domains where most of SpringerLink's data practices meet or exceed Minimum Viable Privacy (MVP).⁷⁷ The following table breaks down each privacy domain and the significant findings for each privacy level.

PRIVACY DOMAIN:

Data Collection

EXCEEDS MVP

—

MEETS MVP

- Users can disable JavaScript without impacting core site functionality, such as searching and downloading the full text of Open Access articles, blocking most tracking scripts and cookies in the process.
- There is no indication that SpringerLink collects physical or behavioral biometric information, though their documentation states that they may collect this type of personal information.
- The Privacy Policy and Supplemental CPRA Policy list the general types of data collected and the general business reasons for why the data is collected.

DOES NOT MEET MVP

- Multiple types of personal information—including biometric data, direct and indirect identifiers, and behavioral data—may be collected through use of SpringerLink.
- Springer Nature collects personal information about users from third parties, such as marketing partners, data analytics partners, social networks, and data brokers.
- The Privacy Policy and privacy-adjacent documentation is vague on specific types of data collected, such as not indicating the collection of ISP and zip code data.
- Users are presented with a cookie banner that uses deceptive design patterns to nudge users to accept all cookies, enabling data collection by SpringerLink and 200 advertising partners, including at least one company currently being investigated for problematic data privacy practices.
- Depending on where the user is on SpringerLink, they may be hampered by broken links when they wish to change their cookie preferences.
- As noted in the Privacy Policy and privacy-adjacent documentation, the categories of data collected (e.g., usage and location data) and the sources where data is harvested (third parties including data brokers) are extensive. In addition, there is the creation of inferred data that can contain potentially sensitive data about an individual's characteristics based on collected personal data. These practices do not abide by data minimization standards.

77 MVP is defined in the Analysis Methodology section.

PRIVACY DOMAIN:

*User Data Rights***EXCEEDS MVP**

—

MEETS MVP

- Users can opt-out of non-essential data collection via trackers and cookies in the cookies management window.
- Depending on applicable regulations, SpringerLink users have rights to access, correct, restrict processing, delete, and export personal information.

DOES NOT MEET MVP

- SpringerLink users outside of jurisdictions that have data protection and privacy regulations (e.g., GDPR, CCPA/CPRA) most likely are not afforded the same data rights as those within these jurisdictions.
- Depending on where the user is on SpringerLink, they may be hampered by broken links when they wish to change their cookie preferences to opt out of non-essential data collection at any time.
- The user data rights specified in the CPRA Policy only apply to California Consumers and are not extended to other states beyond residents in Virginia, Colorado, and Connecticut.

PRIVACY DOMAIN:

*Data Disclosure***EXCEEDS MVP**

—

MEETS MVP

- The Privacy Policy and CPRA Policy provides information about what data is disclosed to specific (but not all) third parties, along with the business case for such disclosure.

DOES NOT MEET MVP

- Personal information is disclosed to third parties for behavioral advertising and targeted marketing purposes.
- The use of social media plugins such as Twitter and Facebook allow for disclosure of personal information to these social media sites. In addition, the use of Facebook Custom Audience ties personal information to a user's account.
- While the Privacy Policy and the CPRA Policy state GDPR and CCPA regulations as rationale for disclosure, additional research may likely be needed to assess the privacy risks with each type of disclosure.
- There is no mention in the Privacy Policy or other privacy-adjacent documentation that all personal information disclosed to third parties is de-identified or aggregated before disclosure.
- While there is a statement in the Privacy Policy about the disclosure of personal information by court order, the statement also states that SpringerLink may disclose personal data if they are "legally entitled" to do so. Outside of two contracts analyzed in the report, there is no definitive statement about the circumstances when SpringerLink will disclose personal information in the case of a law enforcement or governmental request.
- SpringerLink provides personal information (i.e. identifiers, online activity, and inferences) to online advertising services and social networks.

PRIVACY DOMAIN:

*Data Processing***EXCEEDS MVP**

—

MEETS MVP

- The Privacy Policy and CPRA Policy provide information about what data is processed along with the business case for such processing.
- The Privacy Policy states that personal information will be retained until the primary purpose has been fulfilled (or as long as required by law).

DOES NOT MEET MVP

- Personal information, including behavioral information, is processed for targeted, personalized advertising and marketing.
- While the Privacy Policy makes multiple statements about the aggregation and “anonymization” of personal information, the rigor and extent of de-identification of personal data are unknown in several instances throughout the policy.
- The Privacy Policy explicitly states that SpringerLink does not have control over the processing of personal data by third parties collecting said data via trackers placed on their website. The Cookie Policy explicitly states that advertisers and service providers may use the information obtained from cookies to (1) track your browser across multiple websites; (2) build a profile of web visitors based on their activity; and (3) target advertisements to web visitors.

PRIVACY DOMAIN:

*Privacy Policy***EXCEEDS MVP**

—

MEETS MVP

- The publicly available SpringerLink Privacy Policy generally explains the collection, use, and disclosure of personal data provided by the user and third parties.

DOES NOT MEET MVP

- The language in the Privacy Policy or its supplement policies are not included in the signed contracts, leaving little recourse for libraries or users with personal accounts to object before the policy changes take place.
- The Privacy Policy is not written in a clear, concise language for a general audience.
- The Privacy Policy, Cookie Policy, and CPRA Policy contains specific data collection, use, and disclosure practices that contradict library or organizational privacy policies that adhere to data privacy standards and best practices in the library profession.

PRIVACY DOMAIN:

*Data Ownership***EXCEEDS MVP**

—

MEETS MVP

—

DOES NOT MEET MVP

- It is unclear if personal data is deleted at the end of the business relationship, including aggregated and de-identified personal data.
- There seems to be no explicit personal data ownership statements in the analyzed contracts or the public privacy documentation.
- There is no explicit statement about subscribers or users exercising data ownership rights in the case of a merger or acquisition (M&A). The Privacy Policy states that SpringerLink will disclose personal data as part of an M&A, which implies personal data is a company-owned asset.

PRIVACY DOMAIN:

*User Surveillance***EXCEEDS MVP**

—

MEETS MVP

- Users are not required to create a separate personal account to access core site functionality (i.e. search, full access to Open Access articles).
- Users can opt-out of non-essential data collection via certain trackers and cookies in the cookies management window (dependent on functional links on a specific page to manage cookies).

DOES NOT MEET MVP

- SpringerLink uses third party beacons, cookies, and other web tracking methods that could be used to track user behavior outside of the SpringerLink website.
- SpringerLink uses third party services such as Google AdWords Remarketing, Facebook Custom Audience, Hotjar, and Amplitude cookies and web beacons to collect personal data, including user behavior on SpringerLink.
- Users are presented with a cookie banner that uses deceptive design patterns to nudge users to accept all cookies, enabling user tracking by at least one company currently being investigated for potentially problematic data privacy practices (Kochava).
- Personal information may be collected by and disclosed to 200 advertisers, marketers, retailers, data intelligence companies, and data brokers through the use of online trackers (i.e., cookies, web beacons, and local storage objects). Depending on the privacy policy of the third parties, personal information may also be disclosed to fourth parties well outside the constraints of the SpringerLink Privacy Policy and contracts.
- Depending on where the user is on SpringerLink, they may be hampered by broken links when the user wants to control the level of surveillance conducted by different categories of web trackers.
- Springer Nature reserves the right to monitor all data, including log files, to detect "misuse" of SpringerLink content, in which Springer Nature also reserves the right to deny user access to content. This monitoring most likely includes personal information such as behavioral data.
- Users who reject all non-essential cookies when presented the choice in the cookie banner are still tracked on SpringerLink via multiple tracking IDs in SpringerLink web beacons and cookies.

PRIVACY DOMAIN:*Data Security and Accountability***EXCEEDS MVP** —**MEETS MVP** —**DOES NOT MEET MVP**

- There are no descriptions of SpringerLink's data security practices outside of the vague "reasonable state of the art security measures" phrase in the Privacy Policy or in the contracts reviewed in our analysis.
- While the policy makes a general statement regarding limited access to personal data, there is no mention of encryption of personal information outside the use of HTTPS for the SpringerLink website.
- The Privacy Policy states "periodic" reviews of security and privacy policies, but does not give specific timetables (i.e., annually, bi-annually, etc.).
- Outside of "periodic" policy reviews, there seems to be no formal data security or privacy audit process in place, including using independent third parties for the audit and for licensees to review the audit report.
- Only two contracts contain language around user notification after a data breach involving personal information. Other information around incident response is lacking from all contracts and public documentation.

*Limitations and Further Investigation*⁷⁸

Our analysis encompassed a variety of public documentation, contracts, and the front-end (user-facing) website, providing a broad overview of SpringerLink's data privacy practices. Nevertheless, there are limitations to our analysis based on the scope and nature of the resources used. Reviewing compliance with specific data protection and privacy regulations was beyond the scope of our analysis, partly because a legal review heavily depends on which libraries and users fall under which jurisdictions or scopes of specific regulations. Sections of the rubric overlap with rights and requirements governed by regulations, such as data user rights. Still, our analysis is not a legal review of whether SpringerLink complies with specific laws.

The analysis provides a reasonably high-level summary of what data is collected, processed, disclosed, and retained based on publicly available information. Contract language variability prevented a complete picture of data flows and practices, particularly with the inconsistent inclusion of data protection schedules and signed DPAs among the contracts. The lack of sign-in credentials also limited the ability of the analysis to determine differences between an authenticated session and a guest/public session (e.g., any additional data collection that might happen when a user is logged into the site). Some of these limitations can be addressed in future analyses, but lack of access to the back-end systems will most likely leave gaps in any external analysis. The documentation for the Permutive API is behind a password protected site; therefore, we were unable to investigate additional Permutive keys such as `permutive-data-queries` and the meanings behind the data values stored within.

78 This section was adapted from Yoose, Becky, and Nick Shockey. "Navigating Risk in Vendor Data Privacy Practices: An Analysis of Elsevier's ScienceDirect." SPARC, November 7, 2023. <https://doi.org/10.5281/zenodo.10078610>.

Key Findings

Based on our analysis, SpringerLink’s data privacy practices undermine basic library privacy standards.

SpringerLink’s use of and Springer Nature’s partnership with 200 data brokers, advertisers, and other commercial companies—along with what appear to be additional unlisted companies found only in our public website analysis—run counter to libraries’ responsibility to protect the misuse of behavioral data, as stated in the 2021 ALA resolution regarding behavioral data surveillance:

Resolved, that the American Library Association, on behalf of its members

1. stands firmly against behavioral data surveillance of library use and users...
3. calls on libraries and vendors to apply the strictest privacy settings by default, without any manual input from the end-user;
4. urges libraries, vendors, and institutions to not implement behavioral data surveillance or use that data to deny services;⁷⁹

Many third parties collecting “deidentified” user data through SpringerLink likely have the means to reidentify this data and associate it with individuals. The Website Evidence Collector tests revealed third party trackers collecting digital fingerprints, zip codes, and ISP information, among other points of personal information. Even if advertisers and marketing companies don’t directly collect identifiable data points such as names and email addresses, they do routinely collect data around a person’s online behavior: what websites they visit, search terms used, and digital fingerprints. These companies along with data brokers can then create a profile of an individual and associate it with that individual (“reidentification”) using collected data and inferred data (i.e., data about a person inferred by analysis of other collected data points).

In addition, **the SpringerLink Privacy Policy explicitly states that SpringerLink is not responsible for the data collected by third parties through trackers placed on its own website.** This policy decision leaves library workers and users without meaningful recourse on how these third parties handle personal data collected from their use of SpringerLink. It also means that users have to read the individual privacy policies of every third party in order to understand what data uses they are consenting to—an undertaking that would likely require 120 hours to complete.⁸⁰

79 ALA. “Resolution on the Misuse of Behavioral Data Surveillance in Libraries.” <https://www.ala.org/advocacy/intfreedom/datasurveillanceresolution>.

80 The “Analyzing SpringerLink’s Contracts, Policies, and Public Documentation” section contains a breakdown of the calculation of total reading time.

Poor quality control over policy copyediting and link functionality on Springer Nature's website limits users' ability to fully understand and consent to privacy terms. The barriers of poor copyediting, broken or misdirecting links, and confusing navigation between privacy-adjacent policies only serve to obscure, though perhaps unintentionally, problematic practices. Combined with the design of the US version of the cookie consent and management banners, it is unclear whether the average user can provide informed consent. Furthermore, achieving meaningful informed consent is unlikely if users are nudged to accept all cookies at the first banner screen or presented with hundreds of links to third party privacy policies (including some that are broken) with the expectation that users would either be able to read through all of them before using SpringerLink or implicitly consent to these third party privacy policies if they click on the "Accept All Cookies" button on the first banner screen.

Institutions should not solely rely on a company's privacy policy to protect user data.⁸¹ As described in our 2023 ScienceDirect report, US courts have not consistently recognized privacy policies as contracts.⁸² Specific regulations, such as the California Online Privacy Protection Act (CalOPPA) and GDPR, require companies to publish privacy policies publicly, but there is no legal guarantee that these policies would be considered enforceable contracts.⁸³ In addition, any changes in the Privacy Policy do not automatically trigger a renegotiation of the signed contract, leaving libraries with limited options when a change in policy or documented practices further conflicts with patron privacy. Public privacy policies have many functions, but they primarily serve to minimize compliance liability and are not designed for, or to be readable by, the average user.⁸⁴ Relying on the company's word—via the privacy policy—without having stringent data privacy protections explicitly spelled out in the signed contract is not a viable strategy for protecting library user privacy.

81 This subsection was adapted from the 2023 ScienceDirect Report (CC BY 4.0).

82 Citron, Danielle Keats, and Daniel J. Solove. "Privacy Harms." SSRN Scholarly Paper. Rochester, NY, February 9, 2021. <https://doi.org/10.2139/ssrn.3782222>.

83 Tarr, Madelyn. "Accountability Is the Best (Privacy) Policy: Improving Remedies for Data Breach Victims Through Recognition of Privacy Policies as Enforceable Agreements." *Georgetown Law Technology Review* 3, no. 1 (2018): 40. <https://georgetownlawtechreview.org/wp-content/uploads/2019/01/3.1-Tarr-pp-162-201.pdf>.

84 Cranor, Lorrie Faith. "Necessary but Not Sufficient: Standardized Mechanisms For Privacy Notice and Choice." *J. on Telecomm. & High Tech.*, 10 (2012): 36. https://www.law.nyu.edu/sites/default/files/upload_documents/Cranor%20-%20Necessary%20but%20Not%20Sufficient.pdf; Waldman, Ari Ezra. "Privacy, Notice, and Design." SSRN Scholarly Paper. Rochester, NY, March 16, 2016. <https://doi.org/10.2139/ssrn.2780305>.

Suggested Actions⁸⁵

Users and institutions must consider the privacy risks that come with vendors such as Springer Nature whose relationship with data brokers (e.g. Oracle, LiveRamp), advertisers, and data intelligence companies (e.g., Kantar, Virtual Minds) show the increasing connection between the publishing and the data brokering economies. Data collection and use that may be less concerning when limited to a business solely involved in publishing (e.g. search history, logs of content accessed, sensitive personal information, geolocation data) are significantly more concerning when the same data may be disclosed to others or integrated into data brokering products (e.g. risk management, insurance, health agencies, law enforcement). These possibilities underline the urgency of a proactive, multi-pronged approach to protecting the privacy and best interests of users, both collectively and by individual institutions.

Given the 200 “advertising partners” that SpringerLink allows to track users, it would be reasonable to assume that SpringerLink users will be tracked in much the same way as on e-commerce, news, and social media sites. The disavowment of Springer Nature’s responsibility for protecting user data collected by their third party “advertising partners” combined with the data privacy practices described in the analysis compounds the privacy risks to users and institutions on top of the risks found in weak or missing contract language. The widespread data collection, user tracking and surveillance, and disclosure of user data inherent to these business models run counter to the library’s commitment to patron privacy as described in the ALA Library Bill of Rights (LBoR).⁸⁶ The library’s responsibility to protect the patron’s right to privacy is expanded upon in the interpretation of the LBoR:

The right to privacy includes the right to open inquiry without having the subject of one’s interest examined or scrutinized by others, in person or online... Lack of privacy and confidentiality has a chilling effect on users’ selection, access to, and use of library resources. All users have a right to be free from any unreasonable intrusion into or surveillance of their lawful library use.⁸⁷

While libraries continue to describe themselves as protectors of patron privacy, the marketplace has shifted toward business models based on commercial surveillance.⁸⁸ Libraries must be proactive in changing their

85 This section was adapted from Yoose, Becky, and Nick Shockey. “Navigating Risk in Vendor Data Privacy Practices: An Analysis of Elsevier’s ScienceDirect.” SPARC, November 7, 2023. <https://doi.org/10.5281/zenodo.10078610>.

86 ALA. “Library Bill of Rights.” <https://www.ala.org/advocacy/intfreedom/librarybill>.

87 ALA. “Privacy: An Interpretation of the Library Bill of Rights.” <https://www.ala.org/advocacy/intfreedom/librarybill/interpretations/privacy>.

88 This shift has drawn concern from the U.S. Federal Trade Commission, which has begun a rulemaking process related to “commercial surveillance and data security.” See <https://www.ftc.gov/legal-library/browse/federal-register-notices/commercial-surveillance-data-security-rulemaking>. As part of this process, SPARC submitted comments that outline concerns related to commercial surveillance in academic markets. See https://downloads.regulations.gov/FTC-2022-0053-1082/attachment_1.pdf.

dealings with content platform providers to reflect this reality and better protect users. Libraries can consider several actions in both the shorter and longer term. Immediate actions libraries might consider include:

- **Reviewing existing contracts for inclusion of explicit and robust language about protecting personally identifiable information**
 - » The base contract language found in many SpringerLink contracts for the analysis only provides limited protection for “Usage Data” as defined in COUNTER.
 - » Negotiate adding an addendum to the contract if there is none with the current contract. Check if your institution has a Data Processing Addendum (DPA) that explicitly details how to process personally identifiable information. Consider using the SPARC Data Privacy Addendum once it becomes available if there is no existing DPA.⁸⁹
- **Urging Springer Nature to address critical issues around privacy management on the SpringerLink website**
 - » Document and communicate issues with Springer Nature representatives about privacy management functionality on the website, such as broken links for managing cookie preferences.
 - » Push for non-deceptive designs for the cookie banner, including providing users a clear, prominent option to reject all non-essential cookies when first presented with the cookie banner.
- **Reviewing and revising policies and settings around public computing security and privacy**
 - » Install privacy-protecting browsers (e.g., Firefox, Tor) and browser add-ons (e.g., uBlock Origin, Privacy Badger) on public computers to limit tracking through cookies or beacons.
 - » Configure public computers to automatically reset the system after each user session (i.e., reimaging). This will remove data stored on the computer from the user session, including persistent data in local storage.

Library workers can use the following resources to assist them in reviewing and securing public computing:

- » Electronic Frontier Foundation's (EFF) Cover Your Tracks (<https://www.eff.org/pages/cover-your-tracks>)
- » ALA's Library Privacy Guidelines and Checklists (<https://www.ala.org/advocacy/privacy/guidelineschecklists>), specifically *Assistive Technologies* and *Public Access Computers and Networks*

89 The SPARC DPA is currently in a pilot phase. Readers can receive updates about the DPA or express interest in piloting it by signing up for the Privacy and Surveillance Community of Interest — https://docs.google.com/forms/d/e/1FAIpQLSdxJc_Ev6hp50KbLYXRUdgKpilL0PB9VluT7hE2nXgv3whprg/viewform.

- » Choose Privacy Every Day's "Guidelines for Private Online Searching & Browsing" (<https://web.archive.org/web/20230928235615/https://chooseprivacyeveryday.org/guidelines-for-private-online-searching-browsing/>)
 - » Security-in-a-Box (<https://securityinabox.org/en/>)
 - » Digital Privacy & Technology Guide's Privacy Resources for Library Tech Management (<https://guides.masslibsystem.org/digital-privacy-and-technology/resources-for-lib-tech-management>)
- **Educating campus users about Springer Nature's data privacy practices**
 - » Use the forthcoming talking points that are currently in production by the Resource Library Working Group of SPARC's Privacy & Surveillance Community of Practice for library workers to use in instructional or reference sessions about how personal data is collected, used, and shared by Springer Nature and third parties based on the Privacy Policy and other public documentation.⁹⁰
 - » Provide information and training around privacy-protecting tools for campus users who want to protect their privacy while using SpringerLink.
 - » Update library privacy notices or other public policies about vendor privacy practices to include the level of tracking and disclosure of user data by SpringerLink and their "advertising partners."
 - » Work with campus privacy advocates, concerned campus community members, IT, instructional technologists, and people involved in vendor selection and negotiation in addressing the campus' continued relationship with Springer Nature (and other vendors with similar data privacy practices) and how the library will protect user privacy on the users' behalf.

Library workers can use the following resources to assist them in their patron communication and education work:

- » ALA's Privacy Field Guides (<https://libraryprivacyguides.org/>), specifically *How to Talk About Privacy*, *Privacy Policies*, and *Vendors and Privacy*
- » Digital Privacy & Technology Guide's Privacy Resources for the Public (<https://guides.masslib-system.org/digital-privacy-and-technology/resources-for-the-public>)
- » Library Freedom Project Resources (<https://libraryfreedom.org/resources/>)
- » Digital Library Federation (DLF) Privacy and Ethics in Technology Working
 - Group Digital Privacy Instruction Curriculum (<https://osf.io/sebhf/>)
 - Advocacy Action Plan (<https://osf.io/2smrf/>)
- » Digital Shred's Privacy Literacy Toolkit (<https://sites.psu.edu/digitalshred/>)

90 SPARC. "Privacy and Surveillance Community of Practice." <https://sparcopen.org/our-work/privacy-and-surveillance-community-of-practice/>.

It is critical to note that the above steps are only short-term workarounds that limit some of the more well-documented surveillance harms of profitable data-intensive business portfolios. Though entrenched vendor data practices do not change overnight, longer-term strategies with widespread community support could create productive pressure for changes in Springer Nature's data privacy practices. Toward this end, there are additional steps that individual libraries can consider taking over the longer term. These include:

- **Negotiating stronger privacy terms in vendor contracts.** In the case of vendors like Springer Nature whose sites allow extensive third-party tracking, libraries should ask pointed questions about their inclusion and consider pushing back on their presence moving forward. The Navigating Risk in Vendor Data Privacy Practices Project will provide tools and resources for libraries to negotiate, such as model contract language for contract negotiations. SPARC's Privacy & Surveillance Community of Practice⁹¹ can also support libraries in sharing information and experiences in trying to secure stronger privacy contract terms.
- **Better understanding the risks posed by publishers' third party partners and the markets they participate in.** Data fed into online ad exchanges and other systems controlled by third parties fundamentally threatens libraries' commitment to privacy. Working knowledge of these systems can better equip libraries to educate users, push back in negotiations, and make decisions about which vendors to work with in the future. SPARC will provide relevant support through our Privacy & Surveillance Community of Practice and additional learning opportunities and resources.
- **Removing confidentiality clauses that prevent sharing of privacy terms.** Libraries must be free to discuss the privacy terms they successfully negotiate (and those they do not) in order to effectively learn from and leverage others' experiences. By refusing to sign NDAs and sharing terms, libraries can reduce information asymmetries with vendors when negotiating. SPARC's brief resource on pushing back against NDAs⁹² may be useful to libraries pursuing this strategy.
- **Recalibrating relationships with vendors whose privacy practices conflict with library expectations and pose risks to users.** If a vendor is unwilling to commit to sufficient contractual privacy guarantees, libraries may consider adjusting their spend with that provider over time. As more libraries unbundle from "Big Deal" subscription packages,⁹³ there is growing experience with how to successfully navigate the transition to a dramatically lower spend with a given vendor while maintaining access to content through alternative means.⁹⁴ Those institutions interested

91 SPARC. "Privacy and Surveillance Community of Practice." <https://sparcopen.org/our-work/privacy-and-surveillance-community-of-practice/>.

92 SPARC. "Pushing Back Against Confidentiality Clauses & Non-Disclosure Agreements." <https://sparcopen.org/our-work/big-deal-knowledge-base/confidentiality-clauses-and-ndas>.

93 SPARC. "Big Deal Cancellation Tracking." <https://sparcopen.org/our-work/big-deal-cancellation-tracking/>; Aiwuyor, Jessica. "More Research Libraries Decline 'Big Deal' Subscription Contracts with Publishers." Association of Research Libraries (blog), May 14, 2020. <https://www.arl.org/news/more-research-libraries-decline-big-deal-subscription-contracts-with-publishers>.

94 SPARC. "Recommendations for Providing Alternative Access After a Big Deal Cancellation." <https://sparcopen.org/our-work/negotiation-resources/alternative-access>.

in considering an unbundling project can find resources through SPARC, such as the Big Deal Cancellation Tracker,⁹⁵ Unbundling Profiles,⁹⁶ and Negotiations Community of Practice.⁹⁷

Through a combination of the above actions, individual academic libraries can better shore up privacy protections while raising awareness about vendor data privacy practices with patrons. This can reinforce and extend existing privacy advocacy work on campus. However, individual libraries can only change the landscape so much. Institutions and consortia can share their experience pushing back on vendor surveillance practices, leverage individual advances, and collaborate to build and support privacy-preserving scholarly infrastructure.

Libraries continue to have the power to shift the marketplace to once again reflect librarianship's commitment to patron privacy. Any large-scale approach to create alternatives will take time and resources. More importantly, large-scale approaches require a long-term commitment from a critical mass of individuals and organizations. Past actions to build protection mechanisms and meaningful alternatives were primarily successful thanks to the countless hours of privacy work, relationship-building, and advocacy by both individual library workers and organizations. Deliberate, dedicated, sustained action on a local level is a necessary foundation for cross-institutional collaborations that can shape markets and provide valuable resources, partnerships, and commitment needed for these actions to succeed.

In collaboration with the many other organizations and individuals that are actively leading library privacy work, SPARC is committed to supporting libraries in addressing vendor privacy concerns. This work is an essential component of SPARC's broader commitment to ensure that privacy is foundational to equitable systems for openly sharing knowledge.

95 SPARC. "Big Deal Cancellation Tracking." <https://sparcopen.org/our-work/big-deal-cancellation-tracking/>.

96 SPARC. "Unbundling Profiles." <https://sparcopen.org/our-work/big-deal-knowledge-base/unbundling-profiles>.

97 SPARC. "Negotiation Community of Practice." <https://sparcopen.org/our-work/negotiations-community-of-practice/>.

APPENDIX A

– Documents Used for Rubric Analysis

Most of the links to the documents point to archived versions of the page in the Wayback Machine.

CONTRACTS

- **Iowa State University, Springer Nature Contract (2021-2023)**
 - » <https://sparcopen.org/wp-content/uploads/2021/08/Springer-Nature-2021-2023-Iowa-State.pdf>
- **Oregon State University, Springer Nature Contract (2023-2024)**
 - » <https://sparcopen.org/wp-content/uploads/2019/05/Oregon-State-University-Springer-Nature-License-Agreement.pdf>
- **University of California, Springer Nature Contract (2020-2023)**
 - » https://cdlib.org/services-groups/collections/licensed_resources/redacted_licenses/Springer_Nature_Transformative_Agreement_signed_20210114_Redacted.pdf

7 additional contracts were directly shared with SPARC for the purpose of this analysis but are not publicly accessible online. These contracts were not subject to confidentiality clauses or NDAs.

WEBSITE DOCUMENTATION

- **Springer Nature California Privacy Rights Act Policy**
 - » <https://web.archive.org/web/20231202130117/https://www.springernature.com/gp/legal/ccpa>
- **Springer Nature Privacy Policy**
 - » <https://web.archive.org/web/20231209191432/https://www.springernature.com/gp/legal/privacy-statement>
- **Springer Nature Profile Privacy Policy**
 - » <https://web.archive.org/web/20231002114236/https://my-profile.springernature.com/privacy-policy>

- **Springer Nature Terms and Conditions**
 - » <https://web.archive.org/web/20231123105828/https://www.springernature.com/gp/legal/general-terms-of-use/11033520>
- **Springer Support Center Documentation**
 - » About Springer Nature Profile
 - <https://web.archive.org/web/20240227103331/https://support.springernature.com/en/support/solutions/articles/6000248072-about-springer-nature-profile>
 - » Data Security
 - <https://web.archive.org/web/20231004164325/https://support.springer.com/en/support/solutions/articles/6000210497-data-security>
 - » Manage Cookie Settings on SpringerNature Websites
 - <https://support.springer.com/en/support/solutions/articles/6000250352-manage-cookie-settings-on-springernature-websites>
 - » Springer Nature Profile Terms of Use and Privacy Policy
 - <https://web.archive.org/web/20230530133644/https://support.springer.com/en/support/solutions/articles/6000248233-terms-of-use-and-privacy-policy>
- **SpringerLink Cookie Policy**
 - » <https://web.archive.org/web/20231207104102/https://link.springer.com/cookiepolicy>
- **SpringerLink Privacy Policy**
 - » <https://web.archive.org/web/20231207115401/https://link.springer.com/privacystatement>
- **SpringerLink Terms and Conditions**
 - » <https://web.archive.org/web/20231207170643/https://link.springer.com/termsandconditions>

APPENDIX B⁹⁸

– Glossary

N.B. – Several term definitions come from the Licensing Privacy Glossary (<https://publish.illinois.edu/licensingprivacy/glossary/>), published under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Aggregation

A method of de-identification that reduces the granularity of personal data through grouping data into categories or ranges (such as using age ranges to report on user birthdate or age data). Aggregation carries some risk of re-identifying an individual if there are outliers in the data set, if the data set size is small overall, or if categories or ranges are granular enough to identify an individual in the data set.

Behavioral tracking

The practice of surveilling users' actions over a period of time.

Canvas fingerprinting

A specific browser fingerprinting method that uses the HTML5 canvas element to track unique users on a website. This method draws shapes and text on a page hidden from the end user. Differences or variations in the drawing indicate specifics about the user's browser and computer, including hardware and software. These fingerprints are typically used to protect against fraud/abuse but can also be used to track behavior for user profiling, personalization, and advertising purposes.

Consent

The act of an individual freely giving or denying permission for the use or disclosure of an individual's data. Explicit consent requires an affirmative action from the individual, while implicit consent implies consent, such as continued use of a service or website.

Cookies

A small data file sent from the website and stored on the user's computer or the user's browser. Web cookies can be used to manage user authorization and session management, as well as track users through web analytic software and other tracking software. Some cookies last until a web session ends ("session" cookies), while other cookies ("persistent" cookies) can remain on the user's computer after the end of a session. A website can have web cookies from the site itself ("first-party" cookies) as well as cookies from external sites ("third-party" cookies).

98 This appendix is adapted from the 2023 ScienceDirect Report (CC BY 4.0).

Data breach (or data leak)

The unauthorized access of personal data by an individual, organization, or system process. A data breach is an intentional act of gaining unauthorized access, such as an attacker gaining access to personal data for the purpose of identity theft, while a data leak is unintentional, such as an employee losing a laptop or mobile device containing personal data.

Data brokers

Entities that sell personal data collected from private and public data sets.

De-identification

The process of transforming personal data to remove identifiable aspects of the data. This includes a variety of methods, including aggregation, stripping or truncating personal data, or removal or pseudonymization of personal data. De-identified data carries a risk of re-identification, meaning that an individual person can be identified from the dataset by reattaching parts of the dataset back to the individual.

Digital fingerprinting

A set of data about a person's device, browser, and other hardware and software that, when combined, can identify an individual. Often referred to as browser fingerprinting.

Digital surveillance

The act of monitoring and capturing a person's activities through various technologies, including web analytics, cookies, trackers, and other data observation and capture techniques.

Domain name system (DNS)

The service that associates numerical Internet Protocol (IP) addresses with domain names, which are more friendly for everyday human use. A domain name can point to another domain name through a Canonical Name (CNAME) record. The domain name mapped to the other domain name is called an alias.

Incident response

The process of responding to and managing a data breach or leak. This can include:

- Identification and detection of a breach or leak
- Containing and eliminating the cause of the breach or leak
- Communications with affected parties

Local storage

HTML5 local storage allows for offline storing of data in the client web browser. Compared to cookies, local storage has a larger storage capacity to store potentially personally identifiable data, persists after the end of the user session and browser shutdowns, and does not have an automatic expiration date.

Opt-in/out

Opt-in is a choice made by a user involving active affirmation, such as checking a box or toggle to turn on specific data sharing or collection settings.

Opt-out is a choice made by a user through inaction unless the user makes an action to choose otherwise. An example is a product collecting user data until the user unchecks the box that controls the data collection setting.

Personal data

Data relating to an identifiable individual. This includes single points of data that can identify a person (direct identifier); data that, when combined, can identify a person (indirect identifiers); and data about a person's behaviors (behavioral data).

Direct identifiers can include:

- Name
- Email or physical address
- Government or organization-issued identification numbers
- Account username and password
- Biometric information
- IP address
- Device information (operating system, browser, device unique ID, etc.)

Indirect identifiers can include:

- Age or date of birth
- Race/ethnicity
- Gender identity
- Education level
- Major or minor field of study
- Disability status
- Veteran status
- Geographical information, such as regions or zip code

Behavioral data can include:

- Search history
- Electronic content access histories
- Circulation histories
- Website activity
- Geolocation history

Personalized advertising

The practice of using personal data collected through user surveillance (online and offline) to market ads specific to that user's interests and behaviors.

Web analytics

The collection and analysis of website data. Web analytic applications track and capture data from a variety of sources, including user data. Depending on the application, user data can range from search term results, page hits, and landing/exit pages to user demographic data, behavioral data, and even data of users visiting sites outside of the original website.

Web beacon

A web tracking method that is often undetectable by the end user. Web beacons commonly take the form of a small image file (often a single transparent pixel) loaded with a web page or email but can also be embedded in visible graphics (e.g., buttons, banners) and HTML elements. Beacons can collect personal data such as behavioral data, digital fingerprints, and IP addresses.

APPENDIX C

– Web Tracker List and Counts

COOKIE AND WEB BEACON HOST LIST

The following is a list of the unique domains serving cookies and web beacons found in the website test WEC reports at the time of the analysis (Fall 2023).

1. 237ac0daefa111ce1a8460ed3e06aede.safeframe.google syndication.com
2. 2e4b93d1-a8ae-4a89-8885-6109135ac0de.edge.permutive.app
3. 2e4b93d1-a8ae-4a89-8885-6109135ac0de.prmutv.co
4. a6b6296477ccd11a7abfd7763835ee5f.safeframe.google syndication.com
5. aa.agkn.com
6. ad.360yield.com
7. ads.stickyadstv.com
8. analytics.google.com
9. analytics.twitter.com
10. api.eu.amplitude.com
11. api.permutive.com
12. b932819bef02cce0a817cec565f70bc6.safeframe.google syndication.com
13. bat.bing.com
14. c.bing.com
15. cart.springer.com
16. cdn.amplitude.com
17. cdn.jsdelivr.net
18. cdn.pbgrd.com
19. cdn.permutive.com
20. cdn.polyfill.io
21. cdnjs.cloudflare.com
22. ck.solocpm.com
23. ck.tangooserver.com
24. cm.g.doubleclick.net
25. cmp.springer.com

26. cmp-static.springer.com
27. collect.springer.com
28. contextual.media.net
29. ualtr-sync.teads.tv
30. delivery.pbgrd.com
31. dis.criteo.com
32. dpm.demdex.net
33. eb2.3lift.com
34. exchange.mediavine.com
35. fonts.googleapis.com
36. gum.criteo.com
37. hb.yahoo.net
38. i.liadm.com
39. i6.liadm.com
40. ib.adnxs.com
41. idp.springer.com
42. jadservice.postrelease.com
43. lantern.roeye.com
44. lantern.roeyecdn.com
45. link.springer.com
46. match.adsrvr.org
47. match.sharethrough.com
48. media.springernature.com
49. mug.criteo.com
50. notifier-configs.airbrake.io
51. pagead2.google syndication.com
52. partner.mediawallahscript.com
53. pixel.rubiconproject.com
54. push-content.springernature.io
55. r.casalemedia.com
56. rtb-csync.smartadserver.com
57. s.ad.smaato.net
58. s.marvellousmachine.net
59. s3.amazonaws.com

60. script.hotjar.com
61. secure.adnxs.com
62. securepubads.g.doubleclick.net
63. simage2.pubmatic.com
64. siteintercept.qualtrics.com
65. springer.met.vgwort.de
66. springeronlineservice.freshdesk.com
67. sslwidget.criteo.com
68. static.ads-twitter.com
69. static.criteo.net
70. static.hotjar.com
71. stats.g.doubleclick.net
72. sync.crowdctrl.net
73. sync.outbrain.com
74. sync-t1.taboola.com
75. t.co
76. tags.bluekai.com
77. tpc.googlesyndication.com
78. trends.revcontent.com
79. ups.analytics.yahoo.com
80. visitor.omnitags.com
81. ws.rqtrk.eu
82. www.dwin1.com
83. www.google.com
84. www.google-analytics.com
85. www.googletagmanager.com
86. www.googletagservices.com
87. www.mainadv.com
88. x.bidswitch.net
89. zn0ccsrg9grqrxatc-springernature.siteintercept.qualtrics.com

NUMBER OF TRACKERS BY PAGE – WEC RESULTS

Number of Trackers by Page

REJECTED NON-ESSENTIAL COOKIES SESSION

PAGE	1ST AND 3RD PARTY COOKIES	LOCAL STORAGE	WEB BEACONS	TOTAL # TRACKERS
<i>Home page</i>	6	0	2	8
<i>Browse page</i>	2	1	1	4
<i>Search results page</i>	7	0	2	9
<i>Article page</i>	8	8	6	22

Number of Trackers by Page

ACCEPT ALL COOKIES SESSION

PAGE	1ST AND 3RD PARTY COOKIES	LOCAL STORAGE	WEB BEACONS	TOTAL # TRACKERS
<i>Home page</i>	95	19	34	148
<i>Browse page</i>	20	16	16	52
<i>Search results page</i>	37	18	21	76
<i>Article page</i>	36	18	22	76

APPENDIX D

– Advertiser List

The following is a list of companies and the URLs for each company listed under the “Cookies that help show personalized advertising” section in the SpringerLink cookie banner at the time of the analysis (Fall/Winter 2023).

Please note that the URLs provided below are presented “as-is” from the cookie banner list and may be broken (i.e., page no longer exists or has moved to another URL). URLs that are broken at the time of publication of this report will appear in **bold**.

1. Aarki (<http://corp.aarki.com/privacy>)
2. Adacado (<https://www.adacado.com/privacy-policy-april-25-2018/>)
3. Adara Media (<https://adara.com/2018/04/10/adara-gdpr-faq/>)
4. AdClear (<https://www.adclear.de/datenschutzerklaerung/>)
5. ADEX (<http://theadex.com>)
6. Adform (<https://site.adform.com/uncategorized/product-and-services-privacy-policy/>)
7. **Adikteev (<https://www.adikteev.com/eu/privacy/>)**
8. **AdLedge (<https://adledge.com/data-privacy/>)**
9. Adloox (<http://adloox.com/disclaimer>)
10. **Adludio (<https://www.adludio.com/terms-conditions/>)**
11. AdMaxim (<http://www.admaxim.com/admaxim-privacy-policy/>)
12. Admedo (<https://www.admedo.com/privacy-policy>)
13. Admetrics (https://admetrics.io/en/privacy_policy/)
14. Adobe Advertising Cloud (<https://www.adobe.com/privacy/general-data-protection-regulation.html>)
15. AdTriba (<https://privacy.adtriba.com/>)
16. advanced STORE GmbH (<https://www.ad4mat.com/en/privacy/>)
17. Adventive (<https://www.adventive.com/legal/privacy-policy>)
18. Adventori (<https://www.adventori.com/fr/avec-nous/mentions-legales/>)
19. Advenue (<https://www.advenuemediaco.uk/privacy-policy/>)
20. advolution.control (<http://advolution.de/privacy.php>)
21. affilinet (<https://www.affili.net/uk/footeritem/privacy-policy>)
22. Akamai (<http://www.akamai.com/compliance/privacy>)
23. Amazon (<https://www.amazon.co.uk/gp/help/customer/display.html?nodeId=201909010>)

24. Amobee (<https://www.amobee.com/trust/privacy-guidelines>)
25. Analights (<https://analights.com/docs/analights-consent-to-the-processing-of-personal-data-en.pdf>)
26. AppLovin Corp. (<https://www.applovin.com/privacy>)
27. AppNexus (Xandr) Inc (<https://www.xandr.com/privacy/>)
28. Arrivalist (<https://www.arrivalist.com/privacy#gdpr>)
29. AudienceProject (<https://privacy.audienceproject.com/>)
30. Aunica (<https://aunica.com/privacy-policy/>)
31. Avocet (<http://avocet.io/privacy-portal>)
32. Bannerflow (<https://www.bannerflow.com/privacy>)
33. Batch Media (<https://theadex.com/privacy-opt-out/>)
34. BDSK Handels GmbH & Co. KG (<https://www.xxxlutz.de/c/privacy>)
35. Beeswax (<https://www.beeswax.com/privacy.html>)
36. Betgenius (<https://ssl.connextra.com/resources/Connextra/privacy-policy/index-v2.html>)
37. Blismedia (<https://blis.com/privacy/>)
38. Bombora (<http://bombora.com/privacy>)
39. Booking.com (<https://www.booking.com/content/privacy.en-gb.html>)
40. C3 Metrics (<https://c3metrics.com/privacy>)
41. Cablato (<https://cablato.com/privacy-policy/>)
42. Celtra (<https://www.celtra.com/privacy-policy/>)
43. Centro (<http://privacy.centro.net/>)
44. Cint (<https://www.cint.com/participant-privacy-notice/>)
45. Clinch (<https://clinch.co/pages/privacy.html>)
- 46. Cloud Technologies (<http://green.erne.co/assets/PolicyCT.pdf>)**
47. Cloudflare (<https://www.cloudflare.com/security-policy/>)
48. Commanders Act (<https://www.commandersact.com/en/privacy/>)
49. comScore (<https://www.comscore.com/About-comScore/Privacy-Policy>)
50. Crimtan (<https://crimtan.com/privacy-ctl/>)
51. Criteo (<https://www.criteo.com/privacy>)
52. CUBED (<http://cubed.ai/privacy-policy/gdpr/>)
53. DataXu (<https://docs.roku.com/published/userprivacypolicy/en/gb>)
54. Delta Projects (<http://www.deltaprojects.com/data-collection-policy/>)
55. Demandbase (<https://www.demandbase.com/privacy-policy/>)
56. DENTSU (http://www.dentsu.co.jp/terms/data_policy.html)
57. Dentsu Aegis Network (http://www.dentsu.com/termsfuse/data_policy.html)

58. Digiseg (<http://www.digiseg.io/GDPR/>)
- 59. DMA Institute (<https://www.dma-institute.com/privacy-compliance/>)**
60. DoubleVerify (<https://www.doubleverify.com/privacy/>)
61. Distillery (<https://distillery.com/privacy-policy/>)
62. Dynata (<https://www.opinionoutpost.co.uk/en-gb/policies/privacy>)
63. EASYmedia (<https://www.rvty.net/gdpr>)
64. eBay (<https://www.ebay.com/help/policies/member-behaviour-policies/user-privacy-notice-privacy-policy?id=4260#section12>)
65. ebuilders (<https://www.mediamonks.com/privacy-notice>)
66. Effinity (<https://www.affiliation.com/politique-confidentialite.html>)
67. emetriq (<https://www.emetriq.com/datenschutz/>)
68. Enshighten (<https://www.ensighten.com/privacy-policy/>)
69. Epsilon (<https://www.conversantmedia.eu/legal/privacy-policy>)
70. Essens (<https://essens.no/privacy-policy/>)
71. Evidon (<https://www.crownpeak.com/privacy>)
72. Exactag (<https://www.exactag.com/en/data-privacy>)
73. Exponential (<http://exponential.com/privacy/>)
74. Facebook (<https://www.facebook.com/about/privacy/update>)
75. Flashtalking (<http://www.flashtalking.com/first-party-ad-serving/>)
- 76. Fractional Media (<https://www.fractionalmedia.com/privacy-policy>)**
77. FUSIO BY S4M (<http://www.s4m.io/privacy-policy/>)
78. Gemius (<https://www.gemius.com/cookie-policy.html>)
79. GfK (<https://sentic.net/>)
80. Google (<https://www.google.com/policies/technologies/partner-sites/>)
81. GP One (http://www.gsi-one.org/templates/gsi/en/files/privacy_policy.pdf)
82. GroupM (<https://www.groupm.com/privacy-policy>)
83. gskinner (<https://createjs.com/legal/privacy.html>)
84. Haensel AMS (<https://haensel-ams.com/data-privacy/>)
85. Havas Media France - DBi (<https://www.havasgroup.com/data-protection-policy/>)
86. hurra.com (<http://www.hurra.com/impressum>)
87. IBM (<https://www.ibm.com/customer-engagement/digital-marketing/gdpr>)
- 88. Ignition One (<https://www.ignitionone.com/privacy-policy/gdpr-subject-access-requests/>)**
89. Impact (<https://impact.com/privacy-policy/>)
90. Index Exchange (<http://www.indexexchange.com/privacy/>)
91. Infectious Media (<https://impressiondesk.com/>)

92. Innovid (<http://www.innovid.com/privacy-policy>)
93. Integral Ad Science (<http://www.integralads.com/privacy-policy>)
94. intelliAd (<https://www.intelliad.de/datenschutz>)
95. Interpublic Group (<https://www.interpublic.com/privacy-notice>)
96. IPONWEB (<http://www.bidswitch.com/privacy-policy/>)
97. Jivox (<http://www.jivox.com/privacy/>)
98. Kantar (<https://www.kantarmedia.com/global/privacy-statement>)
99. Kochava (<https://www.kochava.com/support-privacy/>)
100. LifeStreet (<http://www.lifestreet.com/privacy>)
101. Liftoff (<https://liftoff.io/privacy-policy/>)
102. LiveRamp (<https://liveramp.com/service-privacy-policy/>)
103. Localsensor (<https://www.localsensor.com/privacy.html>)
104. LoopMe (<https://loopme.com/privacy/>)
105. Lotame (<https://www.lotame.com/about-lotame/privacy/>)
106. Macromill group (<https://www.metrixlab.com/privacy-statement/>)
107. MainADV (<http://www.mainad.com/privacy-policy>)
108. Manage.com (<https://www.manage.com/privacy-policy/>)
109. Marketing Science Consulting Group, Inc. (http://mktsci.com/privacy_policy.htm)
110. MediaMath (<http://www.mediamath.com/privacy-policy/>)
111. Meetrics (<https://www.meetrics.com/en/data-privacy/>)
112. MindTake Research (<https://www.mindtake.com/en/reppublika-privacy-policy>)
113. Mobitrans (<http://www.mobitrans.net/privacy-policy/>)
114. Mobpro (<http://mobpro.com/privacy.html>)
115. Moloco Ads (<http://www.molocoads.com/private-policy.html>)
- 116. MSI-ACI (<http://site.msi-aci.com/Home/FlexTrackPrivacy>)**
117. Nano Interactive (<http://www.nanointeractive.com/privacy>)
118. Navegg (<https://www.navegg.com/en/privacy-policy/>)
119. Neodata Group (<https://www.neodatagroup.com/en/security-policy>)
120. NEORY GmbH (<https://www.neory.com/privacy.html>)
121. Netflix (<http://www.netflix.com/google-3PAS-info>)
122. Netquest (<https://www.nicequest.com/us/privacy>)
123. Neural.ONE (<https://web.neural.one/privacy-policy/>)
124. Neustar (<https://www.home.neustar/privacy>)
125. NextRoll, Inc. (<https://www.nextroll.com/privacy>)

126. Nielsen (<https://www.nielsen.com/us/en/legal/privacy-statement/digital-measurement/>)
127. numberly (<https://numberly.com/en/privacy/>)
128. Objective Partners (<https://www.objectivepartners.com/cookie-policy-and-privacy-statement/>)
129. Omnicom Media Group (<https://www.omnicommediagroup.com/disclaimer.htm>)
130. OneTag (<https://www.onetag.net/privacy/>)
131. OpenX Technologies (<https://www.openx.com/legal/privacy-policy/>)
132. Optomaton (<http://optomaton.com/privacy.html>)
133. Oracle Data Cloud (<https://www.oracle.com/legal/privacy/marketing-cloud-data-cloud-privacy-policy.html>)
134. OTTO (<https://www.otto.de/shoppages/service/datenschutz>)
135. PERMODO (<http://permodo.com/de/privacy.html>)
136. Picalate (<http://www.picalate.com/privacypolicy/>)
137. Platform161 (<https://platform161.com/cookie-and-privacy-policy/>)
138. PMG (<https://www.pmg.com/privacy-policy/>)
139. Polar (<https://privacy.polar.me/>)
140. Protected Media (<http://www.protected.media/privacy-policy/>)
141. Publicis Media (<https://www.publicismedia.de/datenschutz/>)
142. PubMatic (<https://pubmatic.com/legal/privacy-policy/>)
143. PulsePoint (<https://www.pulsepoint.com/privacy-policy>)
144. Quantcast (<https://www.quantcast.com/privacy/>)
145. Rackspace (<http://www.rackspace.com/gdpr>)
146. Rakuten Marketing (<https://rakutenadvertising.com/legal-notice/services-privacy-policy/>)
147. Relay42 (<https://relay42.com/privacy>)
148. Remerge (<http://remerge.io/privacy-policy.html>)
149. Resolution Media (<https://www.nonstoppartner.net>)
150. Resonate (<https://www.resonate.com/privacy-policy/>)
151. RevJet (<https://www.revjet.com/privacy>)
152. Roq.ad (<https://www.roq.ad/privacy-policy>)
153. RTB House (<https://www.rtbhouse.com/privacy-center/services-privacy-policy/>)
- 154. Rubicon Project (<https://rubiconproject.com/rubicon-project-advertising-technology-privacy-policy/>)**
155. Salesforce DMP (<https://www.salesforce.com/company/privacy/>)
- 156. Scenestealer (<https://scenestealer.co.uk/privacy-policy/>)**
157. Scoota (<https://www.scoota.com/privacy>)
158. Seenthis (<https://seenthis.co/privacy-notice-2018-04-18.pdf>)
159. Semasio GmbH (<https://www.semasio.com/privacy>)

160. SFR (<http://www.sfr.fr/securite-confidentialite.html>)
161. Sift Media (<https://www.sift.co/privacy>)
162. Simpli.fi (<https://simpli.fi/simpli-fi-services-privacy-policy/>)
163. Sizmek (<https://www.sizmek.com/privacy-policy/>)
164. Smaato (<https://www.smaato.com/privacy/>)
165. Smadex (<http://smadex.com/end-user-privacy-policy/>)
166. Smart (<http://smartadserver.com/company/privacy-policy/>)
167. Smartology (<https://www.smartology.net/privacy-policy/>)
- 168. Sociomantic (<https://www.sociomantic.com/privacy/en/>)**
169. Sojern (<https://www.sojern.com/privacy/product-privacy-policy/>)
170. Solocal (<https://client.adhslx.com/privacy.html>)
171. Sovrn (<https://www.sovrn.com/privacy-policy/>)
172. Spotad (<http://www.spotad.co/privacy-policy/>)
173. SpotX (<https://www.spotx.tv/privacy-policy/>)
- 174. STRÖER SSP GmbH (https://www.stroer.de/fileadmin/de/Konvergenz_und_Konzepte/Daten_und_Technologien/Stroer_SSP/Downloads/Datenschutz_Stroer_SSP.pdf)**
175. TabMo (<http://static.tabmo.io.s3.amazonaws.com/privacy-policy/index.html>)
176. Taboola (<https://www.taboola.com/privacy-policy>)
177. TACTIC™ Real-Time Marketing (<http://tacticrealt.com/privacy/>)
178. Teads (<https://teads.tv/privacy-policy/>)
179. TEEMO (<https://teemo.co/fr/confidentialite/>)
180. The Trade Desk (<https://www.thetradedesk.com/general/privacy-policy>)
181. Tradedoubler AB (<http://www.tradedoubler.com/en/privacy-policy/>)
182. travel audience, An Amadeus Company (<https://travelaudience.com/product-privacy-policy/>)
183. Travel Data Collective (<https://www.yieldr.com/privacy/>)
184. TreSensa (<http://tresensa.com/eu-privacy/index.html>)
185. TripleLift (<https://triplelift.com/privacy/>)
- 186. TruEffect (<https://trueffect.com/privacy-policy/>)**
187. TrustArc (<https://www.trustarc.com/privacy-policy/>)
188. UnrulyX (<https://unruly.co/privacy>)
189. usemax (Emego GmbH) (<http://www.usemax.de/?l=privacy>)
190. Verizon Media (<https://www.verizonmedia.com/policies/ie/en/verizonmedia/privacy/index.html>)
- 191. Videology (<https://videologygroup.com/en/privacy-policy/>)**
192. Vimeo (https://vimeo.com/cookie_policy)
193. Virtual Minds (<https://virtualminds.de/datenschutz/>)

194. Vodafone GmbH (<https://www.vodafone.de/unternehmen/datenschutz-privatsphaere.html>)
195. Waystack (https://www.wayfair.com/customerservice/general_info.php#privacy)
196. Weborama (<https://weborama.com/en/weborama-privacy-commitment/>)
197. White Ops (<https://www.whiteops.com/privacy>)
198. Widespace (<https://www.widespace.com/legal/privacy-policy-notice/>)
199. Wizaly (<https://www.wizaly.com/terms-of-use#privacy-policy>)
200. ZMS (<https://zms.zalando.com/#>)

October 2024

© 2024 SPARC, subject to a Creative Commons Attribution 4.0 International License

