Schwerpunktinitiative „Digitale Information“

Empfehlungen zu Methoden zur Kontrolle des Zugriffs auf wissenschaftliche Informationsressourcen

Ein gemeinsames Papier der Schwerpunktinitiative Digitale Information der
Allianz der deutschen Wissenschaftsorganisationen und
Deutscher Bibliotheksverband e.V. (dbv), 29.11.2019

Präambel

Der Zugang zu digitalen Inhalten für Wissenschaftlerinnen und Wissenschaftler ist im Wandel begriffen. In Deutschland verfolgen alle Wissenschaftsorganisationen und ihre Infrastruktureinrichtungen das gemeinsame Ziel, Open Access als Standardpublikationsmodell für wissenschaftliche Informationen zu etablieren. Dies geht konform mit der Open Access-Strategie des BMBF, die als Leitprinzip das Etablieren von Open Access als Standard des wissenschaftlichen Publizierens benennt. Der freie Zugang ohne technische, rechtliche oder finanzielle Barrieren garantiert die optimale Nutzbarkeit der Informationen und sichert die Rechte der Wissenschaftlerinnen und Wissenschaftler sowohl in ihrer Rolle als Autorinnen und Autoren als auch bei der Nachnutzung der publizierten Informationen.

Subskriptionsbasierte wissenschaftliche Informationen unterliegen allerdings noch Zugangsbeschränkungen, die über Authentifizierungs- und Autorisierungsverfahren kontrolliert werden. Traditionell wird die Zugangsberechtigung vor allem über IP-basierte Verfahren geprüft. Daneben existieren bereits seit langem Single-Sign-On-Verfahren (z. B. Shibboleth), die sich im Verlagsumfeld jedoch bisher nicht auf breiter Basis durchgesetzt haben. In den letzten Jahren sind nun Weiterentwicklungen der Verfahren zur Zugangskontrolle durch kommerzielle Anbieter zu beobachten, die u.a. von dem Wunsch nach Vereinfachung getrieben sind, aber auch zum Ziel haben, personalisierte Angebote zu ermöglichen. Dabei besteht grundsätzlich die Gefahr, dass personenbezogene Daten verstärkt gesammelt, weitergegeben und zur Analyse des Nutzungsverhaltens der Wissenschaftlerinnen und Wissenschaftler verwendet werden.

Der freie Zugang zu Information und der Schutz der Privatsphäre sind essentielle Konzepte für das Bibliotheks- und Informationswesen und grundlegende Voraussetzungen für die Ausübung der Menschenrechte auf Meinungsfreiheit und geistige Freiheit. Der IFLA-Verhaltenskodex fordert daher alle Informationsdienstleister auf, „Maßnahmen [zu] ergreifen, um die Sammlung personenbezogener Informationen über Nutzerinnen und Nutzer und ihre genutzten Dienste einzuschränken“[1].

Diesem Grundsatz sind auch die Informationsinfrastruktureinrichtungen in Deutschland verpflichtet. Sie setzen sich daher für den Schutz der personenbezogenen Daten ihrer Nutzerinnen und Nutzer

[1] International Federation of Library Associations-Erklärung zum Thema Bibliotheken und geistige Freiheit, https://www.ifla.org/DE/publications/node/8767.

ein. Die Verwendung und Speicherung dieser Daten muss transparent und unter Zustimmung der Nutzerinnen und Nutzer erfolgen. Die folgenden Empfehlungen geben in diesem Sinne eine Orientierung und setzen den Rahmen für die Entwicklung und Einführung neue Technologien für den Zugriff auf elektronische Informationsressourcen in Deutschland.

Fallbeispiel: Das Projekt RA21 – Resource Access in the 21st Century[2]

"Resource Access for the 21st Century" (RA21) war ein gemeinsames Projekt der International Association of Scientific, Technical and Medical Publishers und der National Information Standards Organization (NISO), das in seinem Kern darauf abzielt, SAML-basierte Verfahren (z.B. Single-Sign-On) als Standardtechnologie für den Zugriff auf elektronische wissenschaftliche Informationsressourcen durchzusetzen und so perspektivisch die derzeit weit verbreitete Zugriffskontrolle über IP-Adressbereiche abzulösen.

Ziel von RA21 war die Entwicklung eines vereinheitlichten und vereinfachten Single-Sign-On-Verfahrens (SSO) zum Zugriff auf elektronische Ressourcen. Bei diesem Verfahren wird die Berechtigung der anfragenden Nutzerinnen und Nutzer zwischen dem Anbieter (Service Provider) und der Heimateinrichtung (Identity Provider) ausgehandelt. Daher muss zunächst die Heimateinrichtung der oder des Anfragenden über einen Dienst (Discovery Service[3]) ermittelt werden. Hierfür ist im Rahmen von RA21 eine Best Practice Referenz-Implementierung eines verbesserten Discovery Services entwickelt worden, der von allen teilnehmenden Anbietern gemeinsam genutzt werden kann. Wurde die jeweilige Heimateinrichtung einmal ausgewählt, wird die Information im Browser gespeichert, so dass nun ein SeamlessAccess.org-kompatibler Discovery Service diese automatisiert auslesen kann. Für die Nutzerinnen und Nutzer entfällt auf diese Weise die erneute Suche nach der Heimateinrichtung, die bei den aktuell genutzten SSO-Verfahren häufig als unkomfortabel empfunden wird.

Im nächsten Schritt werden die anfragenden Nutzerinnen und Nutzer zur Authentifizierung an ihren Identity Provider (IdP) weitergeleitet. Der Identity Provider übermittelt an den Service Provider (SP) die Daten, auf deren Grundlage der Zugriff gewährt oder verweigert wird (Autorisierung). Welche Daten an den Service Provider übertragen werden, entscheidet auf technischer Ebene der Identity Provider. An der Verarbeitung dieser Daten ist die von RA21 vorgesehene Infrastruktur nicht beteiligt.

Bewertung und Empfehlungen

Webbasierte Single-Sign-On-Lösungen wie die von RA21 entwickelte und durch SeamlessAccess.org umgesetzte sind geeignet, den Zugang zu lizenzierten, zugangsbeschränkten Inhalten und Diensten zu steuern. Sie erlauben eine anonyme oder pseudonyme Nutzung solcher Ressourcen ebenso wie eine personalisierte Einzelnutzerregistrierung. Gegenüber der IP-basierten Zugangssteuerung bieten sie den Nutzerinnen und Nutzer vor allem Vereinfachungen bei der ortsunabhängigen Nutzung von Diensten außerhalb ihrer Einrichtung. Ferner gibt es administrative Vorteile; einzelne Sessions können technisch besser identifiziert und Nutzeraktivitäten auch bei anonymem Zugang granular nachvollzogen werden, wodurch z. B. eine missbräuchliche Nutzung verfolgt werden kann. Personalisierte Dienste können auf diesen technischen Möglichkeiten aufsetzen.

[2] Das RA21-Projekt endete offiziell am 30. Juni 2019. Zu diesem Zeitpunkt wurden die Empfehlungen aus dem Projekt in einen Betaphasendienst unter der Leitung der Coalition for Seamless Access (www.seamlessaccess.org) überführt.

[3] „Discovery Service" bezeichnet in diesem Kontext einen technischen Dienst zur Recherche und Auswahl der Heimateinrichtung und meint nicht das Recherche System, für das die gleiche Bezeichnung gebräuchlich ist.

Diesen Vorteilen steht potentiell die Gefahr gegenüber, dass Anbieter im Rahmen der Weiterentwicklung von SSO die Weitergabe von personenbezogenen Informationen durch den institutionellen Identity Provider über das erforderliche Maß hinaus als Vorbedingung für die Nutzung festlegen. Auch um Einrichtungen nicht zu benachteiligen, die keinen eignen Identity Provider betreiben, ist eine ausschließliche Festlegung auf SSO-basierte Verfahren aus Sicht der Informationsinfrastruktureinrichtungen nicht wünschenswert. Grundsätzlich sollte immer eine IP-basierte Zugangskontrolle als Alternative angeboten werden. Bei der Umsetzung webbasierter SSO-Verfahren muss sichergestellt werden, dass der Datenschutz auch im Sinne des Prinzips der Datensparsamkeit in vollem Maße umgesetzt wird (privacy by design).

Aus dieser generellen Bewertung ergeben sich die folgenden konkreten Empfehlungen:

1. Der Zugang zu Informationsressourcen muss grundsätzlich ohne Weitergabe personenbezogener Daten möglich sein. Die Einrichtungen sind aufgerufen, Nutzerinnen und Nutzer zu informieren und ihre Verantwortung für die technischen Rahmenbedingungen kritisch wahrzunehmen. Bibliotheken sollten sich hierzu mit der für sie zuständigen IT-Infrastruktureinrichtung (z. B. den Rechenzentren) abstimmen.
2. Generell sollten nur die Daten vom Identity Provider an den Service Provider übermittelt werden, die für den jeweils genutzten Dienst oder Inhalt notwendig sind. Die bloße Autorisierung des Zugangs zu Informationsressourcen erfordert im Rahmen von SSO-Verfahren keine personenbezogenen Attribute. Ausreichend sind die bereits seit längerem eingesetzten Attribute (eduPersonEntitlement und/oder eduPerson-ScopedAffiliation).
3. Für Dienste, die eine Personalisierung erfordern, sollte eine nationale und internationale Verständigung zur datenschutzkonformen Attributfreigabe angestrebt werden. Eine Regelung der Weitergabe von Attributen sollte transparent und servicegruppen-spezifisch erfolgen. Zur Personalisierung sollten pseudonyme, SP-spezifische Identifier bzw. Attribute zum Einsatz kommen, so dass ein User-Tracking über mehrere Service Provider hinweg nicht möglich ist.
4. Zur inhaltlichen Ausgestaltung bei der Übergabe von Attributen für Identity Provider sollten technische Empfehlungen erarbeitet werden. Bestehende Handlungsanleitungen wie der GEANT Data Protection Code of Conduct oder die Empfehlungen der Forschungsstelle Recht im DFN sollten dabei Berücksichtigung finden.
5. Da auch der vermittelnde Discovery Service zur Auswahl der Heimateinrichtung personenbezogene Daten sammeln und verarbeiten kann, sollte dieser Dienst von vertrauenswürdigen und neutralen Einrichtungen betrieben werden.
6. In der lizenzrechtlichen Umsetzung sollte die ausschließliche Festlegung auf Single-Sign-On-Technologien für die Zugangssteuerung zu Inhalten und Diensten nicht akzeptiert werden. Die Möglichkeit zur Nutzung über etablierte alternative Verfahren – etwa über IP-Steuerung – muss immer als zusätzliche Option gegeben sein.
7. Single-Sign-On-Technologie kann den Zugriff auf lizenzierte Inhalte vereinfachen. Grundsätzlich sollte jedoch die Open-Access-Transformation des wissenschaftlichen Publikationsmarktes vorrangig verfolgt werden, um den Zugang und die wissenschaftsadäquate Nachnutzung von Informationsressourcen unter Ausnutzung der digitalen Möglichkeiten sicherzustellen. Der offene Zugang zu wissenschaftlichen Inhalten ist am besten geeignet, Zugriffsprobleme zu lösen.