Handbuch IT in Bibliotheken

2023-12-14

Dieses Handbuch gibt einen knappen, umfassenden Überblick über die wesentlichen IT-bezogenen Themen in Bibliotheken. Beschrieben werden zum einen allgemeine Grundlagen wie technische Infrastruktur, Metadaten und IT-Management und zum anderen bibliotheksspezifische Dienste vom Bibliotheksmanagementsystem und Discovery-Systemen bis zu Repositorien und Werkzeugen für Kommunikation und Wissensmanagement.

Inhaltsverzeichnis

# Einleitung

## Motivation

Because the library has become software, it is no longer viable for our services to exist separately from our software. [...] Most importantly, all library staff must understand that our software is our library, and is everyone’s responsibility.

— (Hanson, Cody 2015)

Mit der wachsenden Bedeutung der **Informationstechnologie (IT) im Allgemeinen und für Bibliotheken im Besonderen** bleibt kaum ein Aspekt bibliothekarischer Aufgaben, der nicht durch IT unterstützt wird. Deutlich wird dies z.B. durch das stetig zunehmende Angebot an elektronischen Informationsmitteln, die Digitalisierung historischer Bestände, interoperable Metadaten oder auch die digitale Langzeitarchivierung. Die alltägliche Handhabung von IT (Smartphones, Automatisierung, Vernetzung, …) wird häufig einfacher, die zugrunde liegenden Systeme werden jedoch immer komplexer und erfordern entsprechend mehr Wissen zu ihrem Aufbau und Betrieb. Während sich einige Teile der IT in Bibliotheken nicht wesentlich von IT in anderen Bereichen unterscheiden, gibt es doch zahlreiche Aspekte von Bibliotheks-IT, die nicht oder nicht speziell genug an anderer Stelle behandelt werden. Das vorliegende Handbuch möchte diese Lücke schließen.

## Zielgruppe

Als Einführung und Nachschlagewerk wendet sich dieses Handbuch vor allem an Personen, die sich einen ersten **Überblick über die verschiedenen IT-Dienste an Bibliotheken** verschaffen wollen. Dies können z. B. Personen sein, die sich im Rahmen ihrer bibliothekarischen Ausbildung mit IT-Diensten in Bibliotheken beschäftigen oder die sich im Rahmen der Einarbeitung in eine neue Position mit IT-Diensten in Bibliotheken beschäftigen. Zur Veranschaulichung der Zielgruppen dienen mehrere sogenannte [Personas](#zielgruppe).

## Inhalt

Dieses Handbuch soll einen knappen und gleichzeitig umfassenden Überblick über die wichtigsten **IT-bezogenen Themen in Bibliotheken** geben. Die einzelnen Themenkapitel sind weitgehend unabhängig voneinander lesbar und mit Querverweisen verbunden. Die Kapitel bilden grob zwei Blöcke:

### Allgemeine technische Grundlagen

* [Technische Infrastruktur](#technische-infrastruktur) beschreibt grundlegende technischen Einrichtungen einer Bibliothek für den Betrieb von Prozessen und Dienstleistungen
* [Management von IT-Systemen](#management) beinhaltet die Einführung und den Betrieb von IT-Systemen allgemein
* [Anforderungsanalyse](#anforderungen) umfasst die Ermittlung und Erfüllung von Bedarfen und Anforderungen an IT-Systeme
* [Sicherheit & Datenschutz](#sicherheit-datenschutz) beschreibt Vorgaben und Maßnahmen zur Förderung der IT-Sicherheit und des Datenschutz
* [Daten & Metadaten](#daten-metadaten) stellt wichtige Begriffe, Standards und Prozesse der Datenverarbeitung in Bibliotheken vor

### Bibliothekspezifische Dienste

* [Bibliotheksmanagementsysteme](#bibliotheksmanagementsysteme) sind spezialisierte Informationssysteme für Arbeitsprozesse rund um Erwerbung, Erschließung, Ausleihe, Zugriff und Auffindbarmachung von Bibliotheksbeständen
* [Discovery & Retrieval](#discovery-und-retrieval) stellt Arten, Bestandteile und Funktionen von Rechercheplattformen vor und liefert Hinweise zu Auswahl und Betrieb von Discovery-Systemen
* [Digitalisierung](#digitalisierung) umfasst Prozesse und Werkzeuge zur Digitalisierung, Erschließung und Präsentation von Kulturgütern
* [Forschungsnahe Dienste](#forschungsnahe-dienste) beschreibt Dienste wie Repositorien und Forschungsdatenmanagement zur Unterstützung von Forschungsprozessen
* [Kommunikation](#kommunikation) beinhaltet Werkzeuge und Methoden interner und externer Kommunikation von Wissensmanagement bis Öffentlichkeitsarbeit

|  |
| --- |
| Gut zu wissen |
| Das Handbuch ist ein „lebendiges Buch“, das stetig ergänzt und aktualisiert werden kann und soll. **Tipps und Korrekturen** sind daher sehr willkommen! [Hinweise zur Mitarbeit](#mitarbeit) und Details zur Umsetzung des Handbuchs finden sich im Anhang. |

## Entstehungsgeschichte

Zur Erstellung des Handbuchs wurden zwischen April 2022 und Oktober 2023 [drei Book Sprints](https://www.th-wildau.de/book-sprint/) an der Bibliothek der Technischen Hochschule Wildau durchgeführt. Dabei trafen sich IT-affine [Expert\*innen aus dem Bibliotheksbereich](#autorinnen), um innerhalb von wenigen Tagen eine umfassende Übersicht zu den [**wichtigsten Themen**](#inhalt) **rund im IT in Bibliotheken** zu verfassen. Die Veranstaltung wurde mit Mitteln des [Publikationsfonds für Open-Access-Monografien des Landes Brandenburg](https://open-access-brandenburg.de/fonds/) gefördert. Kontaktinformationen, Neuigkeiten und Hintergrund zum Projekt finden sich auf der Seite <https://www.th-wildau.de/book-sprint/> und in den Artikeln von Bach (2022) und Christensen und Seeliger (2022).

## Beteiligte Autor\*innen

* Nicolas Bach, Student an der HdM Stuttgart  
  [nicolas.bach@posteo.de](mailto:nicolas.bach@posteo.de)
* Pascal-Nicolas Becker, Gründer und Geschäftsführer der The Library Code GmbH  
  <https://orcid.org/0000-0003-2169-1261> [pascal@the-library-code.de](mailto:pascal@the-library-code.de)
* Janna Brechmacher, Stabstelle IT in der Benutzungsabteilung der Staatsbibliothek zu Berlin  
  <https://orcid.org/0000-0001-7233-7153> [janna.brechmacher@sbb.spk-berlin.de](mailto:janna.brechmacher@sbb.spk-berlin.de)
* Jens Bemme, SLUB Dresden, Referat Saxonica und Kartensammlung (Citizen Science)  
  <https://orcid.org/0000-0001-6860-0924> [jens.bemme@slub-dresden.de](mailto:jens.bemme@slub-dresden.de)
* Sascha A. Carlin, Agile Coach für Führungskräfte in der Softwareentwicklung  
  [rqst@nvsbl.cm](mailto:rqst@nvsbl.cm)
* Anne Christensen, Bibliothekarin und Projektmanagerin bei effective WEBWORK sowie Lehrbeauftragte an verschiedenen Hochschulen  
  <https://orcid.org/0000-0001-7753-1078> [christensen@effective-webwork.de](mailto:christensen@effective-webwork.de)
* Jana Eger, Stadtbibliothek Chemnitz  
  [jana.eger@stadtbibliothek-chemnitz.de](mailto:jana.eger@stadtbibliothek-chemnitz.de)
* Matthias Finck, Professur für Informationstechnologie, Hochschule für Angewandte Wissenschaft Hamburg  
  <https://orcid.org/0000-0002-0848-2935> [matthias.finck@haw-hamburg.de](mailto:matthias.finck@haw-hamburg.de)
* Ulrike Golas, IT-Leitung der Universitätsbibliothek der TU Berlin  
  <https://orcid.org/0000-0002-6567-0000> [ulrike.golas@tu-berlin.de](mailto:ulrike.golas@tu-berlin.de)
* Tina Goldammer, stellv. Leiterin/Fortbildungen bei der SLUB/Sächsische Landesfachstelle für Bibliotheken  
  [tina.goldammer@slub-chemnitz.de](mailto:tina.goldammer@slub-chemnitz.de)
* Gerrit Gragert, Leitung IT-Services für die Digitale Bibliothek in der Staatsbibliothek Berlin  
  <https://orcid.org/0000-0002-0542-1555> [gerrit.gragert@sbb.spk-berlin.de](mailto:gerrit.gragert@sbb.spk-berlin.de)
* Lambert Heller, Leitung Open Science Lab an der TIB - Leibniz‐Informationszentrum Technik und Naturwissenschaften  
  <https://orcid.org/0000-0003-0232-7085> [lambert.heller@tib.eu](mailto:lambert.heller@tib.eu)
* Sina Hurnik, Studio für Kommunikationsdesign  
  <https://sinahurnik.com/>
* Clemens Kynast, Discoverysysteme & Bibliotheksautomatisierung an der ThULB Jena  
  [clemens.kynast@uni-jena.de](mailto:clemens.kynast@uni-jena.de)
* Lukas Lerche, UB Dortmund  
  <https://orcid.org/0000-0002-4027-6840> [lukas.lerche@tu-dortmund.de](mailto:lukas.lerche@tu-dortmund.de)
* Luis Moßburger  
  <https://orcid.org/0000-0002-5326-219X> [lmossburger@t-online.de](mailto:lmossburger@t-online.de)
* Stefanie Nagel, Abteilungsleiterin Open Science an der Universitätsbibliothek TU Bergakademie Freiberg  
  <https://orcid.org/0000-0001-8020-1440> [stefanie.nagel@ub.tu-freiberg.de](mailto:stefanie.nagel@ub.tu-freiberg.de)
* Silvia Polla, Forschungsdatenmanagement, Bibliothek  
  <https://orcid.org/0000-0002-2395-2448> [polla@wias-berlin.de](mailto:polla@wias-berlin.de)
* Michael Schaarwächter, Bibliotheks-IT an der UB Dortmund  
  <https://orcid.org/0000-0002-0180-5930> [michael.schaarwaechter@tu-dortmund.de](mailto:michael.schaarwaechter@tu-dortmund.de)
* Zoe Schubert, Wissenschaftliche Mitarbeiterin & Entwicklerin bei der TIB Hannover und der Staatsbibliothek zu Berlin  
  <https://orcid.org/0000-0001-9043-3632> [zoe.schubert@tib.eu](mailto:zoe.schubert@tib.eu)
* Frank Seeliger, Bibliotheksleiter TH Wildau  
  <https://orcid.org/0000-0003-0602-8082> [fseeliger@th-wildau.de](mailto:fseeliger@th-wildau.de)
* Britta Steinke, Referentin für Forschungsdatenmanagement  
  <https://orcid.org/0000-0001-6816-5168> [b.steinke@tu-berlin.de](mailto:b.steinke@tu-berlin.de)
* Katja Sternitzke, Wissenschaftliche Mitarbeiterin an der Staatsbibliothek zu Berlin  
  <https://orcid.org/0000-0002-9815-0490> [katja.sternitzke@sbb.spk-berlin.de](mailto:katja.sternitzke@sbb.spk-berlin.de)
* Ralf Stockmann, Referatsleiter Innovationsmanagement und Onlinedienste - Staatsbibliothek zu Berlin  
  <https://orcid.org/0000-0002-0977-5908> [ralf.stockmann@sbb.spk-berlin.de](mailto:ralf.stockmann@sbb.spk-berlin.de)
* Florian Strauß, Abteilungsleiter forschungsunterstützende Dienste Universitätsbibliothek Clausthal  
  <https://orcid.org/0000-0003-0168-0450> [florian.strauss@tu-clausthal.de](mailto:florian.strauss@tu-clausthal.de)
* Alexandra Streck, Online-Redakteurin, Marketing ZB MED - Informationszentrum Lebenswissenschaften  
  [streck@zbmed.de](mailto:streck@zbmed.de)
* Robert Strötgen, Leiter der Universitätsbibliothek der TU Braunschweig  
  <https://orcid.org/0000-0003-3320-5187> [r.stroetgen@tu-braunschweig.de](mailto:r.stroetgen@tu-braunschweig.de)
* Jakob Voß, Forschung und Entwicklung an der VZG Göttingen  
  <https://orcid.org/0000-0002-7613-4123> [jakob.voss@gbv.de](mailto:jakob.voss@gbv.de)
* Michael Voss, Consultant; IT-Expert-Voss  
  <https://orcid.org/0000-0002-7402-1598> [info@it-expert-voss.de](mailto:info@it-expert-voss.de)
* Kerstin Wendt, Stabsstelle Digitalisierung, Staats- und Universitätsbibliothek Hamburg  
  [kerstin.wendt@sub.uni-hamburg.de](mailto:kerstin.wendt@sub.uni-hamburg.de)
* Carolin Zapke, Fachreferat und ÖA an der UB Chemnitz  
  <https://orcid.org/0000-0002-2516-535X> [carolin.zapke@bibliothek.tu-chemnitz.de](mailto:carolin.zapke@bibliothek.tu-chemnitz.de)
* David Zellhöfer, Professor für Digitale Innovation in der öffentlichen Verwaltung an der HWR Berlin  
  <https://orcid.org/0000-0002-0403-457X> [david.zellhoefer@hwr-berlin.de](mailto:david.zellhoefer@hwr-berlin.de)

## Rechte an den Inhalten

Soweit nicht anders in Quellenangaben ausgewiesen, stehen alle Inhalte dieses Buches unter der Lizenz [Creative Commons Namensnennung 4.0 Deutschland](https://creativecommons.org/licenses/by/4.0/de/) (CC-BY 4.0 DE).

Das heißt: Sie dürfen das Material in jedwedem Format oder Medium vervielfältigen und weiterverbreiten (**Teilen**) und das Material remixen, verändern und darauf aufbauen (**Bearbeiten**) und zwar für beliebige Zwecke, inklusive kommerzielle Zwecke, unter der Bedingung, dass Sie angemessene Urheber\*innen- und Rechteangaben machen, einen Link zur Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden (**Namensnennung**).

# Technische Infrastruktur

|  |
| --- |
| Zusammenfassung |
| Dieses Kapitel beschreibt die grundlegenden technischen Einrichtungen einer Bibliothek für den Betrieb von Prozessen und Dienstleistungen wie [Verbuchung](#verbuchung), [Zugangskontrolle](#zugangskontrolle), [Internetzugang](#internetzugang) und [Arbeitsplätze](#gruppen--und-einzelarbeitsplätze). Neben [Diensten für Nutzer\*innen](#dienste-für-nutzerinnen) und der Basistechnologie [RFID](#rfdi) wird auch wesentliche [Dienste für Mitarbeiter\*innen](#dienste-für-mitarbeiterinnen) beschrieben. |

## Einleitung

Die technische Infrastruktur einer Bibliothek umfasst alle IT-Systeme, die die Prozesse und Dienstleistungen einer Bibliothek abbilden, unterstützen oder ergänzen. Neben dem zentralen [Bibliotheksmanagementsystem](#bibliotheksmanagementsysteme) zur Verwaltung und Bereitstellung von Bibliotheksbeständen und [Discovery-Systemen](#discovery-und-retrieval) zur Recherche gibt es zahlreiche weitere etablierte Anwendungen von IT in Bibliotheken. Die in diesem Kapitel vorgestellte Infrastruktur ist grob nach Hauptanwendungsfall gegliedert in

1. [Allgemeine Infrastruktur](#Xd6b1793515042b5f4de02aa4a13e076d01a9c07) für den grundlegenden Betrieb von Prozessen und Dienstleistungen in Bibliotheken wie Verbuchung und RFID
2. [Dienste primär für Nutzer\*innen](#dienste-für-nutzerinnen) wie Webseite, Internetzugang und Arbeitsplätze
3. [Dienste primär für Mitarbeiter\*innen](#dienste-für-mitarbeiterinnen) wie Intranet und mobiles Arbeiten

Weitgehend ausgeklammert, weil an anderer Stelle behandelt, bleiben [forschungsnahe Dienste](#forschungsnahe-dienste) wie Repositories und Open Data, Infrastruktur zur [Digitalisierung](#digitalisierung) sowie Anwendungen für [Kommunikation und Wissensmanagement](#kommunikation).

|  |
| --- |
| Info |
| Zu [Einführung und Betrieb](#management) und zur [Entwicklung samt Anforderungsanalyse](#anforderungen) von IT-Systemen gibt es eigene Kapitel. |

## Allgemeine Infrastruktur für Bibliotheksprozesse und -dienste

### Verbuchung

Zur Verbuchung zählen die Ausleihe und Rückgabe sowie die Verlängerung von Leihfristen. Für automatische Verbuchung müssen entsprechende Geräte vorhanden sein, um die für Medien eingesetzten Identifikationsmerkmale (Barcodes oder [RFID](#rfid)-Tags) zu lesen und ggf. auch zu schreiben. Außerdem muss eine Kommunikation mit dem BMS stattfinden, um darin den entsprechenden Ausleih-, Rückbuchungs-, oder Verlängerungsvorgang durchzuführen. Hierzu muss außerdem eine Identifikation der Nutzer\*innen erfolgen, also das Einlesen eines Ausweises oder das Anmelden mit gewissen Zugangsdaten und schließlich das Eintragen einer Ausleihe. Da es Nutzer\*innen möglich sein sollte, die Rückgabe von Medien nachweisen zu können, wird bei der Rückgabe mit Verbuchung immer eine Quittung erzeugt. Dies erfolgt sowohl an den Ausleihtheken als auch bei Rückgabeautomaten. Eine weitere Möglichkeit ist das [Bezahlen](#bezahlung) und Begleichen von Gebühren an gleicher Stelle.

### Automaten

Zur Gestaltung eines Angebots, das über die Servicezeiten hinausgeht, ist der Einsatz von Automaten für Standarddienstleistungen sinnvoll. Vorgänge wie Ausleihe, Verlängerung und Rückgabe von Medien, Bezahlung von Gebühren oder auch die Abholung von Bestellungen können mit Automaten weitgehend ohne Personal realisiert werden.

**Ausleihautomaten** bestehen aus Lesegeräten für Ausweise, Lesegeräten für Medien und Entsicherungsgeräten für Buchsicherungen. Der Verbuchungsprozess beinhaltet eine Anmeldung (Ausweisnummer und ggf. Passwort), den eigentlichen Verbuchungsprozess inklusive Entsicherung der Medien, Ausgabe einer Quittung auf Papier oder als E-Mail sowie eine manuelle oder automatisierte Abmeldung.

**Rückgabeautomaten** sind in zwei Gruppen zu unterteilen. Im einfacheren Fall ist der Ausleihautomat auch gleichzeitig ein Rückgabeautomat. Der Anmeldeprozess kann hier entfallen, denn es ist grundsätzlich nicht relevant, wer das Medium zurückgibt. Die zurückgebrachten Medien werden dann durch die Nutzer\*innen selbst auf einen Bücherwagen oder ein separates Regal gestellt und müssen anschließend durch das Bibliothekspersonal sortiert und eingestellt werden. Der zweite und komplexere Fall eines Rückgabeautomaten ist ein Eingabeschacht, der die Medien weiter transportiert, verbucht und sichert. Im einfachen Fall werden die Medien in einem Behälter gesammelt oder es ist ein automatisches [Sortiersystem](#sortiersysteme) angebunden. Medien, die nicht zum Bestand der Bibliothek gehören und andere Gegenstände müssen bei der Rückgabe erkannt und zurückgewiesen werden.

**Fernleihautomaten** dienen der personalfreien Abholung von bestellten Fernleihmedien oder anderen bestellten Dingen, die nicht durch die Ausleihautomaten verbucht werden können, etwa weil sie keinen RFID-Tag besitzen. Nach dem Einlegen des abzuholenden Mediums durch Mitarbeiter\*innen der Bibliothek werden Bestellende informiert (z.B. durch eine E-Mail), dass in einem Fach x etwas abzuholen sei. In diesem Vorgang kann man den Bestellenden auch ein PIN mitteilen, alternativ ist das Fach durch einen Bibliotheksausweis elektronisch zu öffnen. Wird das Fach geöffnet, wird zeitgleich das Medium im Nutzerkonto verbucht.

**Kassenautomaten** erlauben die [Bezahlung](#bezahlung) von offenen Gebühren oder auch den Kauf von Gutscheinen/Tickets für Dienstleistungen, die im Anschluss in Anspruch genommen werden.

### Mediensicherung

Mit dem Begriff wird die Methode beschrieben, am Ausgang der Bibliothek mittels einer Art von Schleusen Medien, ggf. auch Dinge des Interieurs, zu detektieren, die gesichert und insofern unverbucht sind. Eine absolute Absicherung gegen Verluste ist mit dieser Methode nicht zu erreichen. Der vornehmliche Einsatzzweck ist daher auch weniger die Verhinderung von vorsätzlichem Diebstahl, sondern vielmehr das Entdecken des versehentlichen Vergessens der Ausleihverbuchung.

Im Fall von **EM-Sicherung** oder **HF-RFID** (siehe [RFID](#rfid)) werden zur Erkennung ungesicherter Medien sogenannte Gates aufgestellt, durch die Menschen beim Verlassen der Bibliothek geleitet werden. In diesen Gates sind die entsprechende Detektionstechnologie sowie Alarmsysteme (Ton und/oder Licht) verbaut. Im Fall von **UHF-RFID** kann man auf Gates verzichten, hier genügen wegen der großen Reichweite auch Antennen, die an der Decke montiert sind. Für die Anzeige der detektierten Medien kann auch ein Monitor verwendet werden, auf dem dann etwa gleich das entsprechende Buchcover angezeigt wird.

Die Sicherungsanlagen können je nach Konfiguration auch als Besuchszähler genutzt werden. Die Aussagekraft der Zahlen ist zwar nicht exakt, aber hilfreich genug für eine Ermittlung der Auslastung.

### Bezahlung

In Bibliotheken fallen an verschiedenen Stellen Gebühren oder andere zu zahlende Beträge an. Diese werden einerseits elektronisch erzeugt und z.B. im Bibliothekssystem gespeichert (z.B. Überziehungsgebühren) oder fallen andererseits direkt an (z.B. Verkaufspreise für Dubletten, Tragetaschen, …). Benötigt werden dafür Geräte, die bei Bezahlung den Betrag direkt dem System verbuchen können, in dem die Beträge erfasst sind, z.B. Kassenautomaten mit Anbindung an das Bibliothekssystem. An diesen Automaten können sich Nutzer\*innen anmelden, erhalten eine Anzeige von offenen Posten und können sie direkt begleichen. Manche Bibliotheken erlauben lediglich die bargeldlose Zahlung, manche ermöglichen (auch) die Zahlung mit Bargeld. Unterschiedlich gehandhabt wird auch, ob die gesamte geschuldete Summe zu bezahlen ist oder ob einzelne Posten beglichen werden können. Mit dem Einsatz von Kassenautomaten können Bezahlvorgänge unabhängig von anwesendem Personal ermöglicht werden.

Auch die Verbindung eines Bezahlterminals mit einem Ausleihautomaten ist möglich. Hier sind dann alle Prozesse für die Nutzer\*innen an einem Gerät abwickelbar. Es sind Ausleihautomaten am Markt, die bargeldlose Bezahlung mit verschiedenen Bezahlssystmene (spezielle Debit-Karten, ec-Karten...) und/oder Bargeldzahlungen ermöglichen.

Die Systeme für Bezahlung und Gebühren sind meist an die jeweiligen Finanzstellen einer Universität oder einer Stadt(verwaltung) angebunden.

### Zugangskontrolle

Eine Zugangskontrolle zur Bibliothek ist relevant, wenn Nutzer\*innen die Bibliothek oder Teile davon auch ohne Anwesenheit von Bibliothekspersonal vor Ort nutzen können sollen ([Open Library](#open-library)). Im Regelfall sieht eine technische Umsetzung der Zugangskontrolle so aus, dass das Schließsystem des Gebäudes an das Identifikationssystem für die Nutzer\*innen angebunden ist. Dies lässt sich beispielsweise mit Kartenlesegeräten am Eingang oder mit einem RFID-Terminal lösen, welches die Bibliotheksausweise einlesen und dem Schließsystem nach festgelegten Regeln mitteilen kann, ob die betreffende Person zur Nutzung der Ressource berechtigt ist. Die Nutzung sollte außerdem kurzzeitig im Rahmen des Datenschutz protokolliert werden.

### Auslastungszählung

Die Ermittlung der Auslastung einer Einrichtung ist aus verschiedenen Gründen interessant: für statistische Zwecke, für eine Anzeige auf der Webseite als Service für Nutzer\*innen oder wenn für einen Notfall wichtig ist die ungefähre Anzahl der im Gebäude Anwesenden zu erfahren.

Bei der Auswahl einer geeigneten Lösung muss zuerst festgelegt werden, ob eine exakte Zählung nötig ist, oder eine Approximation der Anwesenheitszahlen ausreichend ist.

Folgende Umsetzungsmöglichkeiten sind für eine exakte Zählung in Erwägung zu ziehen: Auswertung der Zugangskontrollsysteme; Einsatz einer Drehschranke bzw. Vereinzelungsanlage, die nur Einzelpersonen durchlässt; Zählung durch Personal und Zählung durch CV (computer vision). Andere, relativ exakte, aber doch mit kleinen Ungenauigkeiten bei der Zählung behaftete Systeme sind die Zählung durch eine Lichtschranke, durch ein Radar etwa von einem Sicherungsgate (siehe [Zugangskontrolle](#zugangskontrolle)) und IR-Durchgangszähler mit Gruppenerkennung an der Decke montiert (bspw. Produkte der Firma [Irisys](https://www.irisys.net/)).

Bei allen exakten Zählsystemen ist insbesondere die Frage des Datenschutzes zu beachten, da aus der Zählung die Nachverfolgung der Nutzer\*innen nicht abgeleitet werden darf. Aus Datenschutzgründen ist auch der Einsatz von aufzeichnenden (Kamera-)Systemen im öffentlichen Raum immer sorgfältig abzuwägen.

Für die approximative Auslastungszählung eignen sich neben anonymisierten Varianten der o.g. Möglichkeiten insbesondere auch die Auslastungsmessung von anderen Infrastruktursystemen, etwa dem öffentlich WLAN. Wenn davon ausgegangen werden kann, dass ein\*e Nutzer\*in (im Durchschnitt) ein Gerät im öffentlichen WLAN anmeldet, dann kann dies als guter Indikator dienen. Sollen die Zahlen der im WLAN angemeldeten Geräte zur Messung der Auslastung eines Gebäudes dienen, sind allerdings umfangreiche Justierungen notwendig. So zeigte sich z.B. an der UB Dortmund, dass zu Prüfungszeiten die Studierenden in der UB pro Person jeweils mehrere Endgeräte im WLAN nutzten – vermutlich mehrheitlich Smartphone und Notebook. Eine andere Möglichkeit ist die Auswertung der Nutzung von Arbeitsplätzen, etwa der Anteil gerade aktiv genutzter öffentlicher PCs, die Messung von Temperaturabweichungen von Innenräumen, oder die Messung der Lautstärke. In Luzern wurde erfolgreich mit Sensoren in Sitzen bzw. unter Tischplatten experimentiert.

### Vor-Ort-Verlängerung

Die Vor-Ort-Verlängerung ist ein Angebot der Bibliothek für Mitarbeiter\*innen der zugehörigen Institution, also z.B. der Hochschule oder der Verwaltung. Hierbei wird der ausgeliehene Bestand nicht in der Bibliothek, sondern bei Nutzer\*innen vor Ort, also im Büro, im Labor etc. erfasst und die entsprechenden Leihfristen verlängert. Technische Hilfsmittel wie RFID und an das BMS angebundene Schnittstellen erleichtern diesen Prozess.

### Open Library

Unter Open Library werden zwei verschiedene Phänomene zusammengefasst: einerseits ist die [Open Library](https://openlibrary.org) ein offenes Projekt zum Aufbau einer digitalen Bibliothek von frei verfügbaren digitalen Büchern oder Digitalisaten. Andererseits bezeichnet der Begriff den Zugang zu einer Bibliothek durch technische Infrastruktur, ganz ohne oder mit wenigen personellen Ressourcen. Einige Bibliotheken bieten auf diese Weise einen durchgehenden Zugang, andere Bibliotheken nutzen die Möglichkeiten, um ihre Öffnungszeiten zu erweitern.

Zugang erhalten in diesen Fällen ausschließlich autorisierte Benutzer\*innen innerhalb und außerhalb der Öffnungszeiten der Bibliothek. In den Eingangsbereichen wird über Lesegeräte die Zugangsberechtigung geprüft. Hier kann es Altersbeschränkungen geben und auch ein gesperrter Bibliotheksausweis gewährt keinen Einlass.

Buchsicherungsanlagen dienen dem Diebstahlschutz von gesicherten Medien und speichern Vorfälle per Video oder verständigen einen Wachdienst. Vorgemerkte Medien liegen in Vormerkregalen, teilweise auch sogenannten intelligenten Abholregalen oder Automaten bereit. Da das Medium auf ein\*e Nutzer\*in vorgemerkt ist, kann es nur von diesen ausgeliehen werden.

Eine Ausleihe erfolgt durch die Benutzung der Ausleihautomaten. Ausleihregeln sind im Bibliotheksmanagementsystem hinterlegt, werden bei der Verbuchung über den Automaten geprüft und das Medium entsprechend entsichert.

Die Open Library ermöglicht zwar grundsätzlich den Zugang zu allen Medien, aber es gibt auch immer Teile des Bestands, die von der Selbstausleihe ausgeschlossen sind, z.B. Brettspiele, Tageszeitungen und [ausleihbare Geräte](#ausleihbare-geräte).

### Navigation und Lokalisierung

Damit Nutzer\*innen die im OPAC bzw. Discovery-System gefundenen (physischen) Medien auch nutzen können, müssen sie den entsprechenden Regalstandort aufsuchen. Zur Orientierung wird häufig die Aufstellungssystematik der Medien genutzt und im Rechercheergebnis steht, unter welcher Signatur, in welchem Regal, in welcher Etage ein Buch zu finden ist.

In kleineren Bibliotheken kann eine solche Standortangabe zusammen mit der Beschilderung des Gebäudes vor Ort ausreichen, um die Medien zu lokalisieren. Je größer jedoch der Bestand und die Räume einer Bibliothek sind, desto schwieriger ist diese Aufgabe.

Für das einfachere Auffinden von Regalstandorten gibt es verschiedene Visualisierungsmöglichkeiten. Im einfachsten Fall kann der genaue oder ungefähre Standort eines Mediums auf einer statischen oder sogar interaktiven Karte angezeigt werden. Eine solche Karte könnte hierbei direkt beim Suchtreffer im Katalog angezeigt werden, oder als dediziertes Terminal vor Ort vorhanden sein. Anhand der Karte können sich Nutzer\*innen dann zum Buch bewegen.

Technisch komplexer ist es nun, wenn eine solche Karte nicht nur den Standort des Buches, sondern gleichzeitig auch den aktuellen Live-Standort des/der Nutzer\*in visualisieren soll. In einem solchen Fall, der etwa einer Google-Maps Karte gleicht, muss zusätzlich technischer Aufwand betrieben werden, um den aktuellen Nutzerstandort zu ermitteln. Im Gegensatz zu Karten und Navigationslösungen in Automobilen oder unterwegs mit dem Smartphone, kann allerdings innerhalb von Gebäuden nicht auf GPS-Satelliten zur Positionsbestimmung zurückgegriffen werden, sondern es müssen andere Signale oder Ortungspunkte genutzt werden, um den Standort der Person im Raum zu ermitteln.

Mögliche Ansätze sind hierbei etwa sog. Beacons, die mit Funklösungen wie Bluetooth und Triangulation die Position eines Mobiltelefons ermitteln, oder bildbasierte Lösungen, die über die Smartphonekamera bestimmte Objekte oder Marker im Raum erkennen.

Die Standorte von physischen Medien sind über eine Zuordnung von Signaturen zu Regalen, teilweise sogar auch zu einzelnen Regalbrettern, einfach möglich. Dies bedingt allerdings eine systematische Aufstellung, die in Zeiten von mehrheitlich elektronischen Medien und immer weniger Printmedien nicht mehr hilfreich ist. So sollte ursprünglich eine Systematik am Regal den Bestand zu einem bestimmten Thema abbilden, man sollte über die benachbarten Bücher einen Überblick zu einem Thema bekommen können. Da aber bspw. E-Books nicht im Regal auftauchen, stellt sich die Frage nach der Aufstellung von gedruckten Medien in den Lesesälen neu. Eine Alternative zur systematischen Aufstellung kann die dynamische Aufstellung sein. So kann man z.B. ohne Umsignierungen temporäre Sammlungen bilden. Man kann darüber nachdenken, ob die Nutzer\*innen ausgeliehene Bücher selbst zurückgeben können, indem sie sie einfach an einen freien Platz im Regal stellen. Man kann die Nutzer\*innen die früher so ungeliebten „Nester“, also die Konzentration von Büchern an einer beliebigen nicht systematischen Stelle, bilden lassen, so dass vielleicht auch andere von dieser „eigenen“ Systematik profitieren. Dies bedingt natürlich eine technische Lösung, mittels der die Bücher anschließend auch wiedergefunden werden, also irgendwie geortet werden können (siehe auch der Abschnitt [Revision](#revision)).

### Sortiersysteme

Medien-Sortieranlagen sind in der Regel direkt an den Rückgabeautomaten angeschlossen. Nutzer\*innen legen Medien in einen Schacht, in dem die Medien erfasst, zurückgebucht und gleichzeitig gesichert werden. Fließbänder oder Rollen transportieren das Medium entsprechend vorher definierte Ziele, z.B. bestimmte Signaturbereiche, andere Zweigstellen oder vorgemerkte Medien.

Die Anschaffung einer Sortieranlage ist mit hohen Kosten verbunden, sowohl für die eigentliche Anschaffung und Installation als auch für die Wartung. Eine Kosten-Nutzen-Analyse (Anzahl Ausleihen/Rückgaben, Personaleinsatz, Platzbedarf) sollte daher vor der Entscheidungsfindung unbedingt durchgeführt werden. Insbesondere in öffentlichen Bibliotheken sollte außerdem immer bedacht werden, dass nicht alle Medien über das Sortiersystem zurückgenommen werden können (siehe [Medienarten](#medienarten)). Auch sehr kleine, leichte oder besonders große, schwere Medienarten müssen i.d.R. gesondert zurückgebucht werden.

### Revision

Die Aufstellung der Medien in einer Systematik ist die Grundlage des Auffindens dieser Medien. Durch die übliche Zirkulation (Ausleihe, Rückgabe, Wiedereinstellen) und selbst schon beim Stöbern vor Ort ist die korrekte Aufstellung latent gefährdet – es entsteht Unordnung.

Technische Mittel zur Unterstützung bzw. Vereinfachung von Inventur und Revision sind daher gern verwendete Hilfen bei der Kontrolle oder Wiederherstellung der korrekten Aufstellung.

Für die Auswahl von geeigneten Werkzeugen können zwei Fragestellungen herangezogen werden:

* Ist das (nicht ausgeliehene) Medium im Haus?
* Ist das (nicht ausgeliehene) Medium am richtigen Platz?

Die zweite Frage ist eine Konkretisierung der ersten und bereits ein Hinweis auf die unterschiedliche Leistungsfähigkeit der Hilfsmittel bzw. auf das geeignete Vorgehen bei Inventur und Revision.

Das klassische Vorgehen ist das genaue Auffinden und Prüfen der Medien mit Hilfe einer Liste. Dies ermöglicht im besten Fall die Wiederherstellung der korrekten Aufstellung. Durch Digitalisierung der Liste und, sofern vorhanden, eine Anreicherung mit Cover-Bildern, kann eine Beschleunigung dieses Prozesses erreicht werden. Mobile Geräte wie Tablets mit geeigneten Apps oder Laptops auf Rollwagen beschleunigen den Prozess nochmals.

RFID kann bei der Beantwortung der ersten Frage, ob Medien generell im Haus sind, sehr gut eingesetzt werden. Für die zweite Frage, ob Medien am richtigen Platz und auch noch in der richtigen Reihenfolge stehen, ist der Einsatz von RFID aufgrund der räumlichen Auflösung jedoch nur bis zu einem gewissen Grad geeignet. Hierbei ist die verwendete RFID Frequenz ein entscheidender Faktor. Durch die Lesereichweite von 0 bis ca. 35 cm bei der HF-Frequenz ist HF zwar gut geeignet, um ein konkretes Medium in einem kleinen Suchfeld zu finden. Durch Grenzen beim gleichzeitigen Lesen von vielen HF-Transpondern, die sich – wie in Regalen üblicherweise der Fall – nah beieinander befinden, ist es jedoch praktisch nicht besonders hilfreich.

Eine elektronisch unterstützte Inventur in HF-Bibliotheken kann auf zwei Arten durchgeführt werden. Die erste Möglichkeit ist das „Entlangwandern“ an den Regalen mit einem Tischtennisschläger-artigen Gerät, welches an den Buchrücken entlanggeführt wird. Die Genauigkeit dieser Art der Inventur ist bei weitem nicht absolut, da die Lesegenauigkeit stark vom Abstand/Winkel von Transponder zu Lesegerät abhängt (idealerweise sind Antenne und Transponder parallel ausgerichtet). Zudem stören Metallgegenstände (z.B. Metallregale) den Empfang.

Die zweite Möglichkeit ist der Einsatz sogenannter Smart-Shelves. In diesen sind aktive RFID-Komponenten verbaut, die die Detektion der auf ihnen befindlichen Medien erlauben. Nachteilig sind hier die Kosten der Regale, die umgerechnet bei mehreren Euro pro detektiertem Medium liegen. Ein großes Manko bei den oben genannten Inventursystemen ist zudem die mangelnde Integration mit dem BMS bzw. eine ausgereifte Benutzungsoberfläche.

Anders sieht es im UHF-Frequenzbereich aus. Durch die potentielle große Lesereichweite von bis zu ca. 12 Metern bei praktisch keiner Beschränkung der Anzahl der Transponder im Erfassungsfeld ist die Frage der prinzipiellen Präsenz eines Mediums gut zu beantworten. Gleichzeitig ist durch die große Lesereichweite die korrekte Aufstellung an einer konkreten Stelle nur durch einen hohen technischen Aufwand automatisiert zu prüfen.

Scan-Roboter, die regelmäßig den Bestand überprüfen (beispielsweise an der TU Dortmund), können derzeit mit einer Genauigkeit von ca. 1 Kubikmeter den Standort eines Transponders bestimmen. Die zur Zeit im Einsatz befindlichen Roboter der Firma Metralabs aus Ilmenau scannen den Bestand in sogenannten „Runs“, die etwa 30.000 Tags pro Stunde finden. Nach einem abgeschlossenen Run werden die Ergebnisse in einer CSV-Datei über einen Fileshare oder einfach per E-Mail geliefert. Enthalten sind darin u.a. Taginhalt und x,y,z-Koordinaten ausgehend von einem einmal festgelegten Nullpunkt. Diese Informationen können (und müssen) dann weiterverarbeitet werden. Denkbar ist eine Integration in den Katalog, so dass in einem Lageplan der Standort des Mediums angezeigt wird, sobald man darauf klickt. Eine Inventur über alle Daten ist möglich, wenn man die Daten abgleicht mit dem ausgeliehenen Bestand und die Fehlmenge ausweist. Eine Stellrevision ist etwas aufwändiger, da der Bestand dazu über längere Zeit überwacht werden und Abweichungen bestimmt werden müssen – wenn der genaue Soll-Standort jedes Mediums nicht in einer Tabelle erfasst ist, mit der der aktuelle Roboterdaten verglichen werden kann.

Festzuhalten ist, dass auch die von den Robotern erhobenen Daten keine absolute Genauigkeit haben, sondern immer etwas „Schwund“ enthalten. Die Gründe für die Ungenauigkeit sind vielfältig: Reflexionen durch Metall-Anhäufungen im Suchfeld, verstellte Gänge zwischen den Regalen, die die Roboter am Durchfahren und Scannen hindern, sehr eng zusammenstehende Bücher, die verhindern, dass beide Tags gescannt werden usw. Mit einer Fehlerrate von ca. 1% ist beim Robotereinsatz zu rechnen.

### Smart Library

Angelehnt an neologistische Komposita wie „smart home“, „smart industries“ (zu Deutsch ‚Industrie 4.0‘) oder „smart cities“ hielt ebenfalls die Kombination „smart“ und „library“ Einzug in die Welt der Informationseinrichtungen. Einzelne Produkte wie „smart shelves“ warben mit dem Label für neue Dienstleistungen über den Einsatz von RFID-Technologie. Insgesamt bezieht man das Konzept der „smart library“ auf Bibliotheken, die Informationstechnologien im Rahmen der Digitalisierung und Automatisierung auch in neuen Bereichen wie physischen Räumen, Nachaltigkeit und Vernetzung einsetzen (Freyberg und Wolf 2019; Seeliger 2019).

### RFID

Der folgende Abschnitt gibt einen tieferen Einblick in die Thematik „RFID“, die für viele Teile der allgemeinen Infrastruktur (Verbuchung, Mediensicherung, Sortiersysteme…) relevant ist. Insbesondere im Hinblick auf Fragestellungen einer Migration oder Einführung von RFID-basierten Technologien für einen Bibliotheksbestand werden Grundlagen und Hintergründe erläutert, Speicherung und Datenmodelle diskutiert und Unterschiede zwischen Verschiedenen RFID-Technologien aufgezeigt. Weitere Informationen zum Einsatz von RFID in Bibliotheken siehe Seeliger (2014) und Kern (2011).

#### RFID-Grundlagen

RFID (radio frequency identification) ist eine Technologie, um Objekte mittels Funksignalen zu erfassen und zu orten. Die Objekte werden hierbei durch RFID-Tags (auch Etiketten oder Transponder genannt) markiert. Ein RFID-Tag enthält zumeist einen Schaltkreis (Speicher + Prozessor) und eine Antenne. Um mit einem RFID-Tag zu kommunizieren, wird ein Sende-/Empfangsgerät benötigt („RFID-Reader“), welches die Antenne über Funkwellen mit Strom versorgt („Induktion“) und danach mit dem Schaltkreis kommuniziert. Die meisten RFID-Tags sind „passiv“, da sie über die Antenne aus der Ferne mit Strom versorgt werden. Es existieren aber auch „aktive“ Tags, die eine Batterie enthalten. Aufgrund der dafür notwendigen Größe sowie der Kosten sind aktive Tags in Bibliotheken nicht im Einsatz. Die Kosten für passive Tags bewegen sich im Bereich von wenigen Cent pro Stück, natürlich stark abhängig von der Bestellmenge.

„RFID“ ist nicht gleich „RFID“! Die Identifikation mittels Radiofrequenzen wird in verschiedenen Frequenzbereichen betrieben. Die Frequenzbereiche sind untereinander inkompatibel, Tags sind nur in einem Frequenzbereich betreibbar, die Schreib-/Lesegeräte sind ebenfalls dediziert auf den Frequenzbereich. Selten gibt es Geräte, die mehrere Frequenzbereiche lesen können, diese enthalten beide notwendigen Technologien. In europäischen Bibliotheken ist meistens ein sogenanntes RFID-HF im Einsatz, welches im Frequenzbereich von 3-30 MHz arbeitet. In der Industrie, in Bibliotheken in China und USA sowie in einzelnen Bibliotheken in Europa kommt RFID-UHF mit einer Frequenz von 868 MHz in Europa (andere Frequenzbereiche in anderen Gebieten) zum Einsatz. Der Hauptunterschied in der Anwendung zwischen den beiden Frequenzbereichen ist die Schreib-/Lesereichweite (siehe [UHF oder HF?](#uhf-oder-hf-was-sind-tags)).

Neben dem Beschreiben und Auslesen des Speicherchips aus der Ferne ist auch eine mehr oder minder präzise Ortung der Tags möglich. Die Ortung ist mit Ungenauigkeiten versehen, da in dem Frequenzbereich von RFID-UHF Reflexionen der Normalfall sind. Eine Lokalisierung ist also nur über die Kombination verschiedener Verfahren in sehr vielen Messungen und einiger Statistiken möglich und dennoch mit einer Unsicherheit versehen.

Neben Bibliotheken finden sich RFID im Alltag auch oft im Einzelhandel, hier meist als Diebstahlsicherung, zur Inventur oder auch zum Erkennen der Artikel an der Kasse. Auch hier kommt bei neueren Installationen häufig RFID-UHF zum Einsatz. Kontaktlose Zahlungsmöglichkeiten auf Bankkarten (NFC) sind ebenfalls eine Anwendungsform von RFID.

#### UHF oder HF, was sind Tags?

Im deutschsprachigen Raum ist zur Zeit (2022) der Einsatz von RFID-HF verbreitet. Einzelne Bibliotheken setzen auch schon RFID-UHF ein (Beispiele: Bibliothek der Wirtschaftsuniversität Wien oder auch die Universitätsbibliothek Dortmund sowie die Bibliothek der Burg Giebichenstein Kunsthochschule Halle). In der Industrie ist RFID-HF praktisch nicht im Einsatz, was vermuten lässt, dass diese Technologie in einigen Jahren aussterben könnte. Die größere Reichweite von RFID-UHF ermöglicht zum Beispiel eine automatisierte Inventur / Stellrevision (siehe auch Abschnitt [Revision](#revision)). Auf der anderen Seite muss man die großen Reichweiten auch einschränken, um etwa bei Ausleihautomaten nicht das zu verbuchen, was sich in einem größeren Abstand befindet. Dies erfordert Kenntnisse in der Justage der Leistung und Signalstärke der eingesetzten Antennen (Dortmund) oder Abschirmung derselben (Wien).

Ein sichtbarer Unterschied der Transponder ist bedingt durch die beiden Funkfrequenzen HF (High Frequency: 13,56 MHz) und UHF (Ultra High Frequency: 860 bis 960 MHz). Höhere Frequenzen erfordern eine andere Bauform der Empfangsantennen in den Transpondern. HF Transponder haben durchschnittlich die Ausmaße einer Chipkarte, damit auf einer Fläche von ca. 10 bis 20 Quadratzentimetern eine spiralförmige Antenne in einem quadratischen oder rechteckigen Layout untergebracht werden kann. In der Mitte der Spirale ist der Mikrochip platziert. Je größer die Antennen sowohl in den Transpondern als auch in den Antennen der Geräte sind, desto besser funktioniert der Kommunikationsprozess.

UHF Transponder hingegen haben üblicherweise Antennen in einer eher länglichen Bauform, bei der zwei symmetrische Antennenteile von einem Zentrum, in dem der Mikrochip platziert ist, in gegenüberliegende Richtungen zeigen. Hier ist weniger die Größe, sondern das für den jeweiligen Anwendungsfall am besten geeignete, konkrete Layout der Antennen für eine zuverlässige Kommunikation wichtig. Die Ausmaße eines typischerweise in Bibliotheken eingesetzten UHF RFID Tags sind ca 15 mm × 94 mm, es gibt aber auch kleinere Ausführungen.

Es existieren ebenfalls Transponder, die sowohl mit UHF, als auch mit HF ansprechbar sind, was die Kosten pro Transponder deutlich erhöht, da zwei verschiedene Antennen vorhanden sein müssen. Beide Frequenzen gleichzeitig in einer Bibliothek zu verwenden, bringt keinen Vorteil, der diesen Aufwand rechtfertigen würde.

|  |  |  |
| --- | --- | --- |
| |  |  | | --- | --- | |  |  |   Abbildung 1.1: HF Transponder und UHF Transponder |

Das unterschiedliche Layout der Transponder ist für das Einkleben in Büchern wenig relevant. Beide Arten werden bei der Herstellung auf Folie aufgebracht. Man unterscheidet verschiedene Ausprägungen der Weiterverarbeitung: „dry naked“ bezeichnet einen fortlaufenden Folienstreifen. auf den Chips und Antennen in gewissen Abständen aufgebracht sind. Hier gibt es keine Stanzung und keine Klebefläche. „wet inlay“ hingegen bezeichnet eine Weiterverarbeitung, die Klebemasse auf Trägerfolie enthält und gestanzt ist. Weitere Verarbeitung sorgt auf Wunsch für Papieroberfläche, so dass die Transponder auch bedruckt werden können. Hersteller von RFID-Transpondern (in der klebenden Version auch „Tags“ genannt) liefern oft nur „dry naked“, weiter verarbeitete Tags beschafft man von sogenannten „Konvertierern“ (nicht zu verwechseln mit dem manchmal so genannten Prozess der „Konvertierung“ des Medienidentifikators von Barcode zu RFID). Der verwendete Klebstoff genügt keiner hohen Anforderung, die in Bibliotheken mit seltenen Werken denkbar wäre. Langzeithaftung oder garantierte Unschädlichkeit des Klebstoffes für das Papiermaterial sind kein Thema für die Hersteller und nur für sehr viel Geld bei speziellen Konvertierern zu bekommen.

Relevant für die Sichtbarkeit der technischen Komponenten in einer Bibliothek ist jedoch die Gegenseite der Transponder: die Schreib-/Lese- Geräte bzw. genauer gesagt deren Antennen. Die maximale Lesereichweite wird bei HF maßgeblich durch die Größe der Antennen, bei UHF hingegen hauptsächlich durch die Feldstärke bestimmt. Dementsprechend haben HF Antennen, die für die Buchsicherung beispielsweise an Ausgängen verwendet werden, die ungefähre Größe einer mittelgroßen Person und es werden üblicherweise zwei sich gegenüberstehende Antennen zu einem „Gate“ kombiniert, nicht zuletzt um die Leseentfernung auf ca. 80-100 cm zu verdoppeln und der Kontrolle der Transponder beim passieren einer Person durch das Gate zu ermöglichen. Ein solches Gate ist sichtbar, muss in die Architektur des Eingangsbereiches einer Bibliothek integriert werden und kann ggf. auch ein Hindernis darstellen (Rollstuhlfahrer, Fluchtwege). Für andere Anwendungen ist der Größenunterschied der Antennen der beiden Frequenzbereiche weniger relevant, wenn auch nicht komplett zu vernachlässigen: Eine Inventur mit HF bedingt eine gewisse Nähe der Sende/Empfangsantenne zu den Medien, ansonsten müssten die Antennen Ausmaße von mehreren Metern haben. Bei UHF hingegen werden auch Medien in größerer Entfernung erkannt (siehe auch im Abschnitt [Revision](#revision)). Typische Bibliotheks-Vorgänge wie die Ausleihe oder Rückgabe, egal ob an einer Theke oder an einer Selbstbedienungsstation, erfordern eine Lesereichweite von 0 bis maximal ca. 30 cm. Da sich bei beiden Frequenzen für diese Entfernung eine in etwa ähnliche Größe der Antennen ergibt, ist bei diesen Anwendungen der optische Unterschied der Geräte nicht sehr offensichtlich. Bedeutsam für diese Anwendungen ist allerdings die Anzahl der gleichzeitig im Stapel gut lesbaren Medien: bei HF sind es unter optimalen Voraussetzungen ca. 3 bis 8 Medien (je nach Größe bzw. Dicke) bei UHF ist diese Zahl nicht wirklich beschränkt. Beispielsweise werden im Extremfall problemlos 20 bis 30 Medien in einem Stapel gleichzeitig erkannt.

UHF und HF unterscheiden sich außerdem durch den Speicherplatz, der auf dem Chip des Tags vorhanden ist. Faustformel ist: Je weiter gefunkt werden muss, desto energieintensiver ist der Prozess und desto weniger kann übertragen/gespeichert werden. Demzufolge ist auf einem UHF Tag im Regelfall weniger Platz als auf einem HF Tag. So wenig sogar, dass bei vielen üblichen Tags nicht beliebig lange Informationen auf einen Tag geschrieben werden können, sondern sich explizit darüber Gedanken machen muss, welche Daten gespeichert werden sollen. Es existieren allerdings auch spezielle Tags mit großem Speicher, auf denen man etwa das Dänische [Datenmodell](#datenmodelle) finden kann.

RFID-Tags können in verschiedenen Dingen eingebracht werden, so gibt es Tags in Büchern, Schlüsselanhängern sowie auch mobile devices (NFC, Smartphones). Nicht alle Arten sind untereinander kompatibel, so kann man mit NFC in Smartphones zwar RFID-HF-Tags in Bibliotheken auslesen, nicht aber RFID-UHF-Tags.

Zu der Lebensdauer der Tags lässt sich bisher recht wenig sagen – was auch eine gute Nachricht ist, da sich im Zeitraum des Praxiseinsatzes von ca. 20 Jahren bisher keine nennenswerte Abnutzung der elektronischen Eigenschaften der Tags gezeigt hat. Den Klebstoff betreffend, mit dem die Tags in die Medien eingeklebt werden, gilt in Bibliotheken Altbekanntes. Schädigung von Papier und Langzeithaltbarkeit dieser Klebstoffe ist im Zusammenhang mit Barcodeetiketten gut erforscht (Kern 2011).

#### Roboter und RFID

Roboter können im Zusammenhang mit RFID genutzt werden, um Inventuren durchzuführen oder Stellrevisionen, also das Erkennen von verstellten Medien. Wegen der geringen Reichweite ist das mit RFID-HF nur unter großen Einschränkungen möglich. So wird in der Kunstbibliothek Sitterwerk in St. Gallen eine nächtliche Inventur und Lokalisierung der Bücher vorgenommen, um eine dynamische Aufstellung zu realisieren. Dazu wird eine aktive RFID-Komponente nachts sehr dicht am Regal entlanggeführt, um die RFID-Tags auszulesen. Das funktioniert aus (mechanischen) Sicherheitsgründen nur in der Schließungszeit, die Lokalisierungsinformation der Medien ist also tagsüber bei viel Betrieb nicht sehr wertvoll, da veraltet. Weil in der dortigen Bibliothek nicht sehr viel Betrieb ist, ist das aber dort kein großer Schaden.

Für eine Bibliothek mit großem Bestand und viel Betrieb ist der Einsatz von RFID-UHF für Inventur und Lokalisierung sinnvoller. Roboter etwa der Firma Metralabs (Illmenau) sind seit einigen Jahren in Betrieben wie Adler und Conrad, aber auch in hunderten australischen Supermärkten, unterwegs, um laufende Inventuren auch in Umgebung mit Menschen durchzuführen. Seit 2020 gibt es an der UB Dortmund solche Roboter, die in einem RFID-UHF-getaggten Bestand nicht nur Lokalisierung, sondern auch Stellrevisionen durchführen. Auch werden vermisste Medien (also Medien, die nicht am erwarteten Standort auffindbar sind) damit wiedergefunden. Aufgrund der physikalischen Eigenschaften von RFID (Reflexionen der Funkwellen im Raum an vielen Materialien) ist trotz hundertfacher Sichtung eines Tags durch den Roboter in einem „Run“ (ein kompletter Scan) bisher die Genauigkeit auf ca 50 cm in jeder Raumachse beschränkt. In der Praxis bedeutet dies, dass die vom Roboter gelieferte Koordinate für den Aufenthaltsort eines RFID-UHF-Tags in einem kugelförmigen Wahrscheinlichkeitsraum mit einem Meter Durchmesser zu betrachten ist – sprich, man muss das Buch im Umkreis von einem Meter von der Ortsangabe suchen. Das ist für eine Stellrevision zwar besser als gar keine Angabe, genügt aber noch nicht, um Nutzer\*innen mit diesen Informationen genau zum Buch zu leiten. Auf diesem Gebiet wird weiterentwickelt, unter Laborbedingungen wurde bereits eine Ortungsgenauigkeit von unter 10 cm erreicht.

#### Migration RFID

Die Migration von einer Technologie zur anderen führt unweigerlich zu drei Fragen:

* Auf welche Technologie soll ich migrieren?
* Was genau muss migriert werden?
* Wie führe ich den Prozess der Migration durch?

In den letzten Abschnitten schon angeklungen ist der Unterschied von HF und UHF, wodurch sich in den meisten Fällen die Frage der Technologiewahl als erste stellt. Als Faustregel kann hier festgestellt werden, dass die schon langjährige Verbreitung von HF im DACH-Raum dazu geführt hat, dass einige Anbieter mit Standardlösungen am Markt sind und für die meisten Anwendungsfälle abgedeckt sind, inkl. dem Migrationspfad von Barcodes.

Bei der Migration an sich geht es dann zumeist darum, das Identifikationsmerkmal eines Mediums auf ein RFID-Tag abzubilden. Hier existieren mehrere Ansätze: Es kann ein RFID-Tag „von der Stange“ ohne Beschreiben in ein Buch eingebracht werden, dessen eindeutiger Identifikator ausgelesen und – zumeist im BMS – mit dem jeweiligen Medium verknüpft werden. Eine andere Möglichkeit wäre das Beschreiben des RFID-Tags mit einem Identifikator, etwa der Mediennummer, einer eindeutigen Signatur oder einer komplexeren Datentstruktur wie dem Dänischen Datenmodell (siehe folgender Abschnitt).

Schließlich muss der Prozess der Migration gewählt und durchgeführt werden. Das Einkleben und zumindest Auslesen, ggf. aber auch Beschreiben eines RFID-Tags ist eine Aufgabe, die entweder von der Belegschaft einer Bibliothek oder als eingekaufte Dienstleistung durchgeführt werden kann. Erfahrungen zeigen, dass die eingekaufte Dienstleistung in drastisch kürzerer Zeit mit der Aufgabe fertig wird, als wenn dafür nur eigenes Personal eingesetzt wird. Eine typische Zeit für das Taggen eines Buches mit außen angebrachtem Barcode (also Einkleben eines leeren RFID-Tags, Auslesen des Barcodes, Schreiben der Infos des Barcodes auf das RFID-Tag) sind 20 Sekunden inklusive Verbringungsarbeiten (Erfahrungswert der UB Dortmund, Mittelwert beim Taggen von 860.000 Medien an vier Standorten durch einen externen Dienstleister). Es bietet sich an, im gleichen Atemzug auch eine Inventur des Bestands durchzuführen.

Der Wechsel von RFID-HF auf RFID-UHF ist noch nicht erprobt, sollte aber keine größeren Probleme darstellen als der Wechsel von Barcode zu RFID. Ein RFID-HF-Tag und ein RFID-UHF-Tag können nebeneinander in einem Buch geklebt und sicher ausgelesen werden. Einige Hersteller von RFID-Hardware bieten auch hybride Geräte an, die beide Arten von Tags auslesen können.

Ein Wechsel von RFID-UHF zu RFID-HF erscheint nicht sinnvoll und wird daher hier nicht betrachtet. Grundsätzlich gilt dabei aber das gleiche wie im vorherigen Absatz.

#### Datenmodelle

Für die Inhalte von HF-Tags in Bibliotheken wurde das sogenannte „Dänische Datenmodell“ spezifiziert, welches später in der ISO-Norm ISO 28560 Teil 3 aufgegangen ist. Generell wird das Datenmodell in der ISO 28560 spezifiziert. Sie enthält heute drei Teile mit folgendem Inhalt: ISO 28560-1 enthält eine Beschreibung vielfältiger, für Bibliotheken denkbarer Datenfelder. Dies sind neben der Mediennummer auch der Titel von Büchern und weitere Daten, welche eventuell offline verfügbar auf dem Chip sein sollten. Aus den Elementen kann für jedes Land ein „Profil“« zusammengestellt werden. ISO 28560 Teil 2 basiert wiederum auf ISO 15962 und den oben genannten OIDs. Er wird in den angelsächsischen Ländern stark propagiert. In diesen Ländern sind bisher vorwiegend proprietäre Datenmodelle im Einsatz, das Dänische Modell kaum verbreitet. ISO 28560 Teil 3 entspricht zu fast hundert Prozent dem Dänischen Datenmodell. Es ist im Vergleich zum Teil 2 zwar fest kodiert, aber deutlich einfacher strukturiert (Kern2014).

Als Referenz für das Datenmodell auf einem Tag dient die Norm ISO 28560 in drei Teilen:

* Teil 1 beschreibt die erforderlichen und optionalen Datenfelder, die länderspezifisch in Profilen zusammengestellt werden können. Die Felder werden hierbei als OID bezeichnet.
* Teil 2 beschreibt die Anordnung dieser Felder (OIDs) in einem flexiblen bzw. fließenden Speicherlayout und wird primär in angelsächsischen Ländern angewandt.
* Teil 3 der ISO 28560 entspricht dem immer noch häufig im Sprachgebrauch genutzten Begriff des Dänischen Datenmodells und beschreibt die feste Struktur der Daten in verschiedenen anwendungsspezifischen Blöcken.

Diese Norm sollte ursprünglich einerseits die Interoperabilität zwischen Bibliotheken ermöglichen. Es sollte also möglich sein, ein Buch einer fremden Bibliothek mit dem eigenen System zu verbuchen. In Deutschland oder in Österreich wird dies nicht flächendeckend (vorsichtig ausgedrückt) angewendet. In der Schweiz ist das hingegen üblich und wegen der zentralen Magazinbibliothek auch notwendig. Andererseits sollte eine Interoperabilität zwischen BMS und RFID-Infrastruktur ermöglicht werden, damit einzelne Komponenten des Systems problemlos ausgetauscht werden können, ohne die Funktionalität zu gefährden.

Die Informationen, die im Dänischen Datenmodell in einem RFID-Tag gehalten werden, sind teilweise Daten, die auch schon im BMS über ein Medium vorgehalten werden. Das Dänische Datenmodell ist von seinem Format so komplex bzw umfangreich, dass es wegen des geringeren zur Verfügung stehenden Speichers nicht auf beliebige UHF-Tags geschrieben werden kann, sondern explizit UHF-Tags mit großen Speicher genutzt werden müssen. Der geringere Speicher von den meisten UHF-Tags erlaubt lediglich die Speicherung einer Bibliotheks-ID, einer ID des Mediums (auch „Mediennummer“ oder „Strichcode“ genannt) und natürlich eines Bits, welches den Sicherungsstatus speichert.

#### Exkurs: Speicherung von Informationen auf Strichcodes und RFID-Tags

Um Informationen, wie etwa eine Mediennummer, eine Signatur oder eine komplexere Datenstruktur – wie das Dänische Datenmodell – auf einem Identifikationsmerkmal, wie einem Barcode, einem QR-Code oder einem RFID-Tag in maschinenlesbarer Form abzuspeichern, muss die Information entsprechend dem jeweiligen Zielformat encodiert werden. Im Folgenden wird beispielhaft für Barcodes und RFID-Tags erläutert, wie so etwas funktioniert.

**Barcodes** sind weit verbreitet, beispielsweise auf Produkten im Supermarkt. Meistens steht unter einem Barcode im Klartext, welche Nummern- oder Buchstabenfolge sich hinter einem Barcode verbirgt. Allerdings ist nicht jeder Barcode wie ein anderer. Es existieren verschiedene Formate, die sich darin unterscheiden, durch welche Strich- und Leerplatzfolge jeweils einzelne Zeichen dargestellt werden. Je weniger Zeichen von einem Barcodeformat unterstützt werden sollen, desto weniger Striche werden pro Zeichen benötigt und desto kompakter wird der Barcode (siehe [Strichcode in Wikipedia](https://de.wikipedia.org/wiki/Strichcode)).

[Abbildung 1.2](#fig-barcode) stellt einen Barcode im Code39-Format dar, welches in Bibliotheken verbreitet ist und Zahlen, Großbuchstaben und ein paar wenige Sonderzeichen darstellen kann.

|  |
| --- |
| Abbildung 1.2: Beispiel für einen Code39 Barcode |

**QR-Codes** funktionieren analog, erweitern allerdings die Darstellung um eine zweite Dimension, sodass auf weniger Platz mehr Informationen dargestellt werden können. Gleichzeitig enthalten QR-Codes eine Prüfsumme, die Fehler bei der Erkennung, etwa bei Beschädigungen am Code, ausgleichen können.

Ein **RFID-Tag** enthält einen Speicherchip und hält daher seine Informationen in Bits und Bytes, genau so, wie Daten und Dateien auch auf einem Computer gespeichert werden. Wie sich jeweils eine Information aus Bits und Bytes interpretieren lässt (ganz so, wie man aus den Balken eines Barcodes Buchstaben und Zahlen interpretiert), hängt von der jeweiligen Encodierung ab, wobei der Speicher eines Tags auch in unterschiedliche Encodierungsformate aufgeteilt werden kann, um Platz zu sparen.

[Abbildung 1.3](#fig-speicher-rfid) zeigt einen Speicherbereich, wie er auf dem Chip eines Tags vorkommen könnte:

|  |
| --- |
| Abbildung 1.3: Beispiel für einen Speicherbereich auf einem RFID-Tag |

Hierbei ist der Speicher 32 Bit klein, was (geteilt durch 8) genau 4 Byte entspricht. Die einzelnen Bits können entweder 0 oder 1 sein und eine zweistellige Darstellung von jeweils 8 Bit als Byte nutzt die Werte von 00 bis FF im Hexadezimalsystem. Der Speicher, auf dem nun 19E9B6EA steht, wird in diesem Beispiel wie folgt encodiert: Die ersten beiden Byte sollen nach URN Code 40 gelesen werden, ein Verfahren mit dem man Text kompakt darstellen kann ähnlich der Idee hinter Code39 (IPC RFID STANDARD FOR INDENTIFYING POSTAL ITEMS BASED ON THE UPU S10 CODE, USING THE ISO/IEC 18000-63 PROTOCOL, Version 1.0, Februar 2017, Seite 33ff). Damit ergibt sich der Text „DE2“. Die letzten beiden Byte sollen hier einfach als Zahl gelesen werden. Die hexadezimale Zahl B6EA wird im Dezimalraum zu 46826. Das Beispielhafte RFID-Tag würde hier also etwa eine Präfixkennung DE2 für die Bibliothek enthalten, sowie eine fortlaufende Mediennummer. Würden wir vereinbaren, auch URN40 für die letzten zwei Byte des Speichers zu nutzen, so hätten wir nur die Möglichkeit, Zahlen von 000 bis 999 (bzw. dann auch dreistelligen Text) zu encodieren. Daher die wechselnde Encodierung.

Komplexere Datenmodelle, wie das Dänische Datenmodell, brauchen demzufolge deutlich mehr Speicherplatz, um auf ein Tag zu passen.

#### Steuerkommandos, Arten des Zugriffs

RFID-Geräte kommunizieren mit RFID-Tags über verschiedene sog. Steuerkommandos. Analog zur Barcodes, die von einem Drucker gedruckt werden (WRITE) und von einer Lesepistole eingelesen werden (READ), so gibt es ähnliche Steuerkommandos auch für RFID-Tags, wobei diese entsprechende Kommunikation zwischen RFID-Gerät und Tag auslösen. Je nach Technologie (HF, UHF bzw. sogar nach technischer Spezifikation der Tags, SLI, SLI-S, SLI-X) existieren verschiedene Kommandos, die verschiedene Aktionen mit einem Tag auslösen. Ohne in technische Details zu gehen lassen sich diese Kommandos grob wie folgt abstrahieren: READ, WRITE, SECURE, UNSECURE, KILL, PROTECT. Dies entspricht der natürlichen Interaktion mit den Tags, etwa beim Inventarisieren neuer Bücher (WRITE), dem Verbuchen (READ, dann UNSECURE) oder dem Durchschreiten eines Gates (prüfen auf SECURED).

Bei der Nutzung von Bibliotheksgeräten, die RFID einsetzen, kommt man im Normalfall nicht mit den Steuerkommandos von RFID-Hardware in Berührung. Sollte man aber mit der Hardware selber experimentieren oder testen, wird man manchmal auch roh auf die Tags schreiben müssen.

#### Sicherheit und Datenschutz

RFID-Tags können auf zwei Arten manipuliert werden: Einerseits kann man mittels geeigneter Hardware (im Falle von RFID-HF genügt ein Smartphone) den Inhalt eines nicht schreibgeschützten Tags verändern. Dies betrifft sowohl das Sicherungsbit (das Gate schlägt also nicht mehr an, wenn ein so manipuliertes Medium herausgetragen wird) als auch den gesamten Tag Inhalt, so dass das Medium von der Infrastruktur der Bibliothek nicht mehr verarbeitet werden kann. Andererseits kann man die meisten UHF-Tags und manche HF-Tags mit einem einfachen Befehl zerstören, also dauerhaft und endgültig stummschalten. Beide Arten der Manipulation kann man mit einem Passwortschutz wirkungsvoll verhindern (fun-fact: die meisten in Deutschland eingesetzten HF-Systeme enthalten diesen Passwortschutz nicht, sind also nicht vor einfachsten Manipulationen geschützt).

Bei der Einführung von RFID werden häufig Diskussionen zum Thema Datenschutz geführt. Wenn allerdings die Tags lediglich mit einer nur intern bekannten ID beschrieben werden, also einer ID, die nicht öffentlich im Katalog des BMS einsehbar ist, besteht diese Gefahr nicht. Selbst wenn jemand Medien im Rucksack eine\*r Nutzer\*in scannen würde (was technisch nicht unaufwändig ist), könnte man daraus keine Rückschlüsse auf das betreffende Medium schließen.

#### Anbindung von RFID-Infrastruktur an BMS

Damit ein Gerät einer RFID-Infrastruktur mit dem Rest der Systeme einer Bibliothek, insbesondere dem [BMS](#bibliotheksmanagementsysteme), kommunizieren kann, muss es über entsprechende Schnittstellen angebunden werden. Bei der Anschaffung sollte daher darauf geachtet werden, dass die Anlage etwa die standardisierten Schnittstellen wie SIP2 oder NCIP nutzt, um Dienstleistungen wie Rückgabe, Sortierung und Ausleihe mit dem Bibliothekssystem abwickeln zu können (Michaelis 2014).

Die Anbindung von lokalen RFID-Readern an Computerarbeitsplätzen von Mitarbeiter\*innen erfolgt im Regelfall durch das Anschließen eines solchen Gerätes direkt am Arbeitsplatz, zumeist über USB. Es existieren allerdings auch Reader, die über einen Netzwerkanschluss direkt mit dem Netz der Einrichtung verbunden werden können und dadurch weitere Flexibilität ermöglichen, da kein Gerät an einen lokalen Computer angesteckt werden muss.

Im Vergleich zu Barcode-Lesern, die den Inhalt eines gelesenen Barcodes meist einfach als Tastatureingaben an die Position des Cursors am Bildschirm ausgeben, können RFID-Geräte aufgrund der auch komplexen Inhalte von RFID-Tags (siehe [Datenmodelle](#datenmodelle)) auch mittels Programmierschnittstellen aus dem BMS oder anderen Systemen angesprochen werden. Die Logik der Interpretation des des Taginhalts liegt hierbei dann beim BMS, oder bei einer Zwischensoftware („Proxy“) die zwischen Reader und BMS vermittelt.

#### Medienarten

RFID-Transponder sind nicht für alle Medienarten geeignet. Aufgrund der Tatsache, dass RFID-Tags über Funkwellen mit Strom versorgt werden und kommunizieren, gilt, dass die Tags durch das Vorhandensein von Metall, Wasser, o.ä. beeinflusst oder abgeschirmt werden können. Selten vorkommende metallisch beschichtete Einbände von Bücher zum Beispiel verhindern effizient die Nutzung von RFID. Konkret bedeutet das, dass die Antenne eines RFID-Tags im Regelfall nicht mehr funktioniert, wenn sie in direkter Nähe zu metallischen Oberflächen ist: Eine CD, ein Laptop im Rucksack, oder sogar ein anderen RFID-Tag in einem dünnen Buchstapel können die Reichweite und Lesbarkeit einschränken. Beim Einbringen von Tags ist daher darauf zu achten, dass sie nicht auf solche Materialien aufgebracht werden. Ebenfalls hilft es nicht, ein Tag auf die Außenseite eines metallischen Gegenstands aufzubringen.

Generell sollte der Gegenstand, auf den ein Tag aufgebracht wird, von Funkwellen durchdrungen werden können, zumindest aber von der Seite, an der das Tag aufgebracht wurde.

Gleichzeitig bedeutet das auch, dass RFID-Tags durch das Vorhandensein von Wasser, also Menschen, ebenfalls abgeschirmt werden können. Eine hunderprozentige Erkennungsrate in einem Sicherheitsgate ist somit unrealistisch.

RFID-Transpoder sind natürlich nicht geeignet für Medien, bei denen eine Unwucht störend ist (Schallplatten, CDs), sie sollten dabei auf der Außenhülle angebracht weren. Bei CDs ist zusätzlich der Metallanteil der CD störend.

Dazu kommt, dass RFID-Tags unterschiedlicher Hersteller und Arten unterschiedlich auf verschiedene Umgebungen reagieren. Für UHF wird vom Unternehmen EECC eine Studie herausgegeben die die physikalischen Eigenschaften von verschiedenen Tags untersucht („UHF RFID Almanach 2022“ 2022). Um diese Studie nutzen zu können, sind allerdings fundierte physikalische Kenntnisse notwendig, alternativ kann die Firma um mit einer Empfehlung beauftragt werden, wie 2019 in der UB Dortmund geschehen.

Nach Einführung von RFID ist das Weiterführen von Barcodes zur Identifikation zwar nicht mehr zwangsläufig erforderlich, es ergeben sich aber zwei Vorteile, insofern weiterhin der Barcode mit am oder im Medium angebracht wird. Durch den Barcode kann das Medium weiterhin maschinenlesbar identifiziert werden. Falls RFID Komponenten mal ausfallen sollten, kann mit dem Barcode traditionell weiter gearbeitet werden. Wenn außerdem (wie meist üblich) unter dem Barcode auch die im Barcode codierten Zeichen mit zu sehen sind, können auch Menschen das Medium bzw. den Band problemlos eindeutig identifizieren. Barcodes können dabei auch auf ein mit Papier beschichtetes RFID-Tag aufgebracht werden.

Auf die Barcodes kann beim Einsatz von RFID auch verzichtet werden, wenn in Kauf genommen wird, dass man für die Identifikation eines Buches eine Recherche im BMS durchführen muss, wenn das Tag unlesbar oder falsch beschrieben ist.

#### Alternativen zu RFID in Bibliotheken

Fortgeschrittene Entwicklungen im Bereich OCR und generell CV (computer vision) können es ermöglichen, ganz ohne technisch lesbare Identifikationsmerkmale ein Buch zu erkennen und zu verarbeiten. Bei einer solchen Lösung wird mittels einer Kamera das Äußere eines Mediums aufgenommen und erkannt und kann somit weiterverarbeitet werden. Um das Medium einer Bibliothek zuzuordnen kann weiterhin etwa ein Aufkleber auf dem Buchen, z.B. ein Signaturetikett auf dem Rücken, erkannt werden, um das Buch der Bibliothek zuordnen zu können und ggf. mehrere Exemplare auseinander halten zu können.

Bisher ist kein solches System aktiv im Einsatz; die Erkennung von Objekten ist jedoch ein aktives Forschungsthema.

Vollständige Automatisierung ist in Bibliotheken auch möglich mit Barcodeverbuchung und EM-Sicherung. Wenn die Barcodes vorn außen am Buch angebracht sind, können sie auch von Ausleihautomaten und Rückgabeautomaten gelesen und verarbeitet werden. EM-Sicherung wird realisiert mittels magnetisierbaren Streifen, die mit doppelt klebendem Film in den Falz der Bücher geklebt wird. Diese Streifen sind über separate Geräte magnetisierbar bzw. entmagnetisierbar. Der Status ist detektierbar über spezielle Gates, so dass darüber die Buchsicherung realisierbar ist. Auch diese Sicherung ist natürlich fehleranfällig und nicht 100%ig. Schon metallene Gegenstände in derselben Tasche wie gesicherte Bücher verhindern die Erkennung. Es gibt bei dieser Methode also keinen Vorteil gegenüber eines RFID-Betriebs – im Gegenteil, diese Technologie ist veraltet, wird immer seltener genutzt und insofern immer teurer.

## Dienste für Nutzer\*innen

### Website, Landingpages und Newsletter

Eine Internetpräsenz ist für eine öffentliche Einrichtung mittlerweile unabdingbar und dient vielen Nutzer\*innen als Erstkontaktmöglichkeit, Arbeitsmittel und Informtionsplattform gleichermaßen.

Eine technische Basis für eine Onlinepräsenz wird manchmal von hiesigen Verwaltungen oder Rechenzentren gestellt, allerdings sollte auch hier darauf geachtet werden, dass die Grundfragen und -dienste mit einer solch vorgefertigten Variante umgesetzt werden können. Alternativ kann es auch sein, dass allein die technische Basis – ein (Web-)server – bereitgestellt wird und sich die jeweilige Einrichtung selbst um die Umsetzung einer Webseite kümmern muss.

Folgende Basisinformationen für Bibliotheksbesucher\*innen und Interessierte sollten zu finden sein

* Öffnungszeiten, Anmeldemodalitäten, Gebühreninformationen
* Veranstaltungshinweise
* Digitale Angebote
* Fachspezifische Angebote wie medienpädagogische Inhalte
* Verlinkung zu Bibliotheks-OPAC oder anderen Online-Katalogen

Schließlich muss darauf geachtet werden, dass nicht nur die eigentliche Webseite der Einrichtung die Basisinformationen enthält, sondern auch, dass diese in einschlägigen Suchmaschinen indexiert sind, etwa die Adresse und Öffnungszeiten in einer Google Suche und auf Google Maps. Hierzu müssen entsprechende SEO (Search Engine Optimization) Parameter eingestellt bzw. an die jeweiligen Plattformen übermittelt werden.

### Internetzugang

Der kostenlose Zugang zum Internet ist für viele Nutzer\*innen ein Grund die Bibliothek als Lern- und Arbeitsort zu nutzen. [PC-Arbeitsplätze](#Xa39a3ee5e6b4b0d3255bfef95601890afd80709) und freies WLAN gehören deshalb in den meisten Bibliotheken mittlerweile zum Standard. Letzteres ist Voraussetzung, um mit eigenen Geräten wie Notebook, Handy und Tablet arbeiten zu können. Damit in allen relevanten Bereichen WLAN mit angemessener Bandbreite verfügbar ist, sollten Bibliotheken Anforderungen an die Ausstattung des Gebäudes mit einer ausreichenden Anzahl an WLAN-Access-Points bestimmen. Das öffentliche Netz sollte vom internen Netz für Mitarbeiter\*innen der Bibliothek getrennt sein, um das Risiko eines Angriffs auf die Infrastruktur zu minimieren. Bei öffentlichen PCs sind zusätzlich Datenschutz-Maßnahmen zu treffen.

Grundsätzlich sind für die Bereitstellung von Internet zwei Fragen zu klären:

* Ist der Zugang offen oder sind ein Passwort und ggf. Registrierung notwendig?
* Erfolgt der Zugang direkt durch die Bibliothek oder per Roaming über Dritte?

In Hochschulen und Forschungseinrichtungen bietet sich die WLAN-Roaming-Infrastruktur [eduroam](https://eduroam.org), die in Deutschland vom DFN koordiniert wird und auch international an vielen Bildungseinrichtungen verfügbar ist. Der Betrieb von eduroam wird in der Regel durch das universitäre Rechenzentrum verantwortet. Nutzer\*innen müssen allerdings an einer Einrichtung registriert sein und ihre Zugangsdaten kennen, um eduroam zu verwenden.

Für den offenen Zugang kann im Idealfall mit der Trägereinrichtung z.B. der eigenen Kommune zusammengearbeitet werden. Eine weitere Möglichkeit ist das ehrenamtliche Freifunk-Projekt. Je nach Rahmenbedingung gibt es verschiedene Leitfäden und Fördermöglichkeiten zur Einrichtung offenen Internetzugangs.

|  |
| --- |
| Internet-Nutzungshinweise in den Städtischen Bibliotheken Leipzig |

Als Anbieter von öffentlichem WLAN sollten Bibliotheken auf Gefahren und mögliche Sicherheitsvorkehrungen hinweisen. Bei der Nutzung von öffentlichem WLAN muss beachtet werden, dass die Verbindungen in der Regel nicht verschlüsselt sind und somit alle, die sich im gleichen Netzwerk befinden, potenziell die übertragenen Daten mitlesen können. Zur Minimierung des Risikos sollten Webseiten möglichst nur verschlüsselt per HTTPS aufgerufen werden und Datei- und Verzeichnisfreigaben deaktiviert sein, um zu verhindern, dass andere Teilnehmer im Netzwerk auf eigene Dateien zugreifen können.

### Gruppen- und Einzelarbeitsplätze

Zur Ausstattung von Gruppen- und Einzelarbeitsplätzen gehört auch angemessene Informationstechnik. Wesentlich sind zunächst ein stabiler Internetzugang und Steckdosen. Ausstattung, Verwaltung und Unterhalt von Räumen mit Technik ist insbesondere für öffentliche Bibliotheken ressourcen- und kostenintensiv. Entsprechend sollten sich Bibliotheken an der tatsächlichen Nachfrage ihrer Nutzer\*innen orientieren und nur die Technik anschaffen, die sie selbst verwalten können.

Es gibt auch einige Bibliotheken, die in ihren öffentlichen Arbeitsbereichen sogenannte Smartboards zur Verfügung stellen. Damit sind große Monitore gemeint, die Computer enthalten. Die gewohnten Funktionen wie Webbrowsing, Textverarbeitung und andere Programme sind über Bildschirmtastatur/Touchscreen erreichbar, man kann aber auch eine Art digitales Whiteboard nutzen und die darauf erarbeiteten Ergebnisse digital weiternutzen.

In der Praxis zeigt sich jedoch bisher, dass Smartboards meistens nur als großer Monitor mit extern angeschlossenem Notebook oder integriertem Computer genutzt wird. Ein Angebot von derartigen Projektionsflächen lässt sich auch kostengünstiger realisieren.

In einigen Bibliotheken werden auch komplette Workspaces für große Gruppen angeboten. Für die gezielte Nutzung von speziell ausgestatteten Arbeitsplätzen kann das Angebot einer Platz- bzw. Raumbuchungslösung nützlich sein. Das Buchungssystem lässt sich auf der Homepage oder in einer Service-App einbinden und kann auf diesem Weg ebenso wie andere Dienstleistungen genutzt werden.

### Öffentliche PC-Arbeitsplätze

Sowohl wissenschaftliche als auch öffentliche Bibliotheken bieten PC-Arbeitsplätze für ihre Nutzer\*innen an. Auch wenn der Trend zu eigenen Geräten geht, bleiben die Nutzungszahlen bei den PC-Arbeitsplätzen besonders in öffentlichen Bibliotheken stabil.

Dienste wie die Nutzung des Internets und Textverarbeitung mit oder ohne Gebühren sind die häufigsten Einsatzzwecke für PC-Arbeitsplätze und sollten weiterhin niedrigschwellig angeboten werden. Die Gebührenabrechnung für angemeldete Nutzer\*innen erfolgt über [Bezahlsysteme](#Xa39a3ee5e6b4b0d3255bfef95601890afd80709), die an das Bibliotheksmanagementsystem angegliedert sind. Geräte mit Münz- oder Kartenzahlung bieten außerdem die Möglichkeit, diese auch ohne Bibliotheksmitgliedschaft zu nutzen.

In der Regel stehen PC-Arbeitsplätze angemeldeten Nutzer\*innen zur Verfügung. Anhand der Benutzergruppe können altersbedingte Einschränkungen vorgenommen werden. So müssen Eltern z.B. der Internetnutzung von minderjährigen Kindern zustimmen.

Bei kostenfreier und nicht reglementierter Nutzung muss der Jugendschutz besonders beachtet werden. Denkbar ist dabei das Whitelisting von Internetseiten, d.h. eine Freischaltung von Seiten, die aus Jugendschutz-Sicht als unbedenklich eingestuft wurden. Ein Blacklisting hingegen ist nicht empfohlen, da keinesfalls alle bedenklichen Seiten bekannt sein können.

Auch der Datenschutz spielt im öffentlichen Bereich eine große Rolle. PCs müssen so konfiguriert werden, dass Nutzer\*innen ausschließlich ihre eigenen Dateien sehen und keinen Zugriff auf Dateien von anderen Personen erhalten. Dies wird z.B. durch persönliche Nutzerprofile (gebunden an das Benutzerkonto) oder systemseitige Rücksetzung aller Einstellungen (Gastzugänge) erreicht. Zum Betrieb solcher „Kiosksysteme“ gibt es entsprechende Software. Auch der Einsatz von Thin-Clients ist in diesem Bereich sinnvoll.

Auch bei der Freigabe der Nutzung von Speichermedien sollten Sicherheitsvorkehrungen getroffen werden, beispielsweise vorgeschaltete automatische Virenprüfungen. Eine Sperre von externen Speichermedien wäre auch denkbar, jedoch spricht die aktuelle Nachfrage dagegen.

Hinweise zur Einrichtung eines öffentlichen WLAN in Bibliotheken gibt der Abschnitt zum [Internetzugang](#internetzugang).

### Digitalisieren/Scannen, Digitalisierung on demand

Scannen, Kopieren und Drucken sind weitere häufig genutzte Angebote in Bibliotheken. Zur Unterstützung von Nachhaltigkeit und Digitalisierung von Studium und Lehre könnte das Ausdrucken auf Papier reduziert oder gar nicht angeboten werden und stattdessen das Einscannen auf Datenträger oder Speichersysteme befördert werden.

Um Nutzer\*innen das Digitalisieren von Medien zu ermöglichen, kann eine Bibliothek Scanner zur Verfügung stellen. Im einfachsten Fall sind das Multifunktionsgeräte (MFGs), die sowohl Kopier- als auch Scan- und Druckfunktionen anbieten. Meistens sind Druck- und Kopierfunktionen kostenpflichtig, Scannen oft kostenfrei. Höherwertige Scans von größeren Vorlagen und ergonomisches Scannen sind mit sogenannten Kopfscannern möglich. Bei diesen Scannern liegt das Medium offen, mit dem Druckbild nach oben auf der Vorlagefläche. Mit Fingerdruck oder Fußschalter wird der Scanvorgang ausgelöst, anschließend kann ohne Umdrehen der Vorlage, wie es bei einem herkömmlichen Kopierer notwendig wäre, umgeblättert werden.

Ein weiterer Anwendungsfall ist das Einscannen von Einzelblattvorlagen. z.B. Vorlesungsmitschriften. Hier sind Scanner mit automatischem Papiereinzug ideal, die Vorder- und Rückseite gleichzeitig einscannen können.

Das Digitalisat kann in der Regel auf einem USB-Stick gespeichert werden. Komfortabler sind eine Netzwerkverbindung und eine Anmeldemöglichkeit für Nutzer\*innen. Alternativ kann auch die Eingabe einer Mailadresse mit anschließendem Versand eines Links auf das Dokument angeboten werden. Das Digitalisat selbst per E-Mail zu verschicken ist i.d.R. aufgrund der Dateigröße nicht möglich.

### Ausleihbare Geräte

Als „Bibliothek der Dinge“ wird die Möglichkeit bezeichnet, in Bibliotheken auch Gegenstände wie Werkzeuge, Sportgeräte und Musikinstrumente ausleihen zu können. Für die Ausleihe von Kunstwerken oder Spielen sind auch die Begriffe „Artothek“ bzw. „Ludothek“ üblich. Für diese Gegenstände ist in der Regel eine besondere Form der [Mediensicherung](#mediensicherung) notwendig. Für die Ausgabe von Tablets gibt es beispielsweise spezielle Automaten. Durch Verbindung mit dem Bibliotheksmanagementsystem ist es auch möglich, die Freigabe an ein Mindestalter zu knüpfen und verschiedene Profile auf den Tablets anzulegen.

### Makerspace

Ein Makerspace ist ein Bereich in einer Bibliothek, in dem Hardware und Software zum Ausprobieren zur Verfügung gestellt wird. Ziel ist das Angebot eines niedrigschwelligen Zugangs zu neuen (technischen) Entwicklungen, die der breiten Öffentlichkeit – in der Regel aus Kostengründen – sonst nicht zur Verfügung stehen. Beispiele für Angebote in Makerspaces sind:

* 3D Drucker
* Geräte zur Holz/Metallverarbeitung
* Stickmaschine/Nähmaschine
* Ton- und Videotechnik
* Repaircafé

Zusätzlich zur Bereitstellung der Technik bieten viele Bibliotheken Einführungs- und Expertenkurse an, die jedoch auch stark von den vorhandenen Personalressourcen abhängig sind. Makerspaces sind vor allem in größeren öffentlichen Bibliotheken verbreitet. Auch in einigen wissenschaftlichen Bibliotheken gibt es inzwischen entsprechende Angebote, wobei hier der Fokus mehr auf dem Einsatz in Lehre und Lernen liegt, zum Beispiel das Dortmunder [Hybrid Learning Center (hylec)](https://hylec.tu-dortmund.de/).

### Bibliotheks-App

Bei Überlegungen zum Einsatz einer App für die Dienstleistungen der Bibliothek sollten verschiedene Aspekte betrachtet werden. Eine App wird um so häufiger installiert, je mehr wichtige und häufig genutzte Funktionalitäten damit nutzbar sind. Eine Integration der Dienstleistungen der Bibliothek in eine bestehende App der übergeordneten Institution ist also der Eigenentwicklung vorzuziehen – sofern das möglich ist.

In einer App können grundsätzlich alle Dienstleistungen der Bibliothek angeboten werden. Ein Mehrwert entsteht dann, wenn man den mobilen Charakter des Endgerätes berücksichtigt. Beispiele: Navigation in der Bibliothek mit Wegweisung am Endgerät, Buchung des Gruppenarbeitsraums, vor dem man gerade steht (z.B. über einscannbare QR-Codes).

Eine vollständige Nutzung der Dienstleistungen der Bibliothek ist nur dann möglich, wenn man sich auch digital anmelden kann. Es sollte also eine Form eines digitalen Ausweises geben, Accountname/Passwort im einfachsten Fall, komplett digitaler Ausweis über die App im besten Fall.

Wildau als Beispiel mit UNIDOS hat eine Integration des Bibliothekskontos, mit der Möglichkeit der Anzeige aller entliehenen Medien und der Verlängerung dieser. Bewerkstelligt wird die Funktion via SIP2-Schnittstelle. Zusätzlich können Discovery-Systeme sowie eine Raumbuchung verlinkt bzw direkt via App ermöglicht werden.

### Technische Beratung und Schulungen

Technische Beratung erfolgt oft in dem Umfang, der für lokale Bibliotheksdienste sinnvoll ist. Bietet eine Bibliothek z.B. die Onleihe als Dienst an, werden sich Nutzer\*innen bei Fragen direkt an die Bibliothek wenden und nicht an den Dienstleister.

Somit müssen sich auch die Mitarbeiter\*innen in der Bibliothek stetig fortbilden, um ihren Nutzer\*innen einen guten Service zu bieten.

Beispiele:

* Erklärung und Dokumentation zu Diensten, z.B. Ebook-Leihe, Streaming-Dienste, E-Learning-Ressourcen
* Ebook–Reader Beratung zur Unterstützung der Ebook-Ausleihe
* Beratung zu App-Nutzung, die als digitale Inhalte angeboten werden

Werden neue Dienste eingeführt, bedarf es neben der Werbung auch einer Einführung oder dem Angebot einer Schulung, in erster Linie für Mitarbeiter\*innen. Viele Anbieter unterstützen dabei mit eigenem Schulungsmaterial, was unter Umständen je nach Zielgruppe angepasst werden muss.

## Dienste für Mitarbeiter\*innen

Die folgenden IT-Dienstleistungen dienen der Unterstützung der täglichen Arbeit, insbesondere im Hinblick auf verteilte Arbeitsumgebungen und mobiles Arbeiten. In vielen Fällen werden sie von der übergeordneten Einrichtung einer Bibliothek bereitgestellt. Neben dem Zugang zu Arbeitsmitteln dienen die Dienste vor allem der internen [Kommunikation und dem Wissenmanagement](#kommunikation).

|  |
| --- |
| Info |
| **Wissensmanagement** besteht aus [Prozessen](#wissensmanagement) zur Erfassung und Weitergabe von Wissen innerhalb einer Organisation, das oft nur implizit in den Köpfen der Mitarbeitenden vorliegt. |

### Mobiles Arbeiten

Für mobiles Arbeiten müssen Endgeräte transportabel sein (Notebook, Tablet…) und zum anderen müssen die Dienste, die für das Arbeiten notwendig sind, vom jeweiligen Standort aus erreichbar sein (siehe [VPN](#vpn)). Für dauerhaftes Arbeiten von anderer Stelle als dem Büro (Homeoffice) ist aus Ergonomiegründen ein fester Arbeitsplatz mit Tastatur, Maus, Bildschirm und ggf. Anschlussmöglichkeit für mobile Geräte („Dock“) vorzuziehen oder vorgeschrieben.

Der Begriff *bring your own device* (BYOD) bezeichnet die Nutzung von privaten Endgeräten in der Infrastruktur des Arbeitgebers. Dies ist allerdings mit einigen Herausforderungen verbunden. So kann nicht zentral sichergestellt werden, dass das Endgerät frei von Schadsoftware ist. Auch die Sperrung des Gerätes aus der Ferne ist nicht möglich (auf eigenen Geräten des Arbeitgebers kann eine solche Software installiert werden, die im Falle eines Diebstahls aktiviert werden kann). Letztendlich liegt also die Verantwortung dafür, ob dienstliche Daten über das private Gerät in falsche Hände gelangen, beim Arbeitnehmer.

### VPN

Ein „Virtuelles Privates Netzwerk“ dient dazu, über einen authentifizierten Zugriff das Endgerät der Mitarbeiter\*innen bzw. Nutzer\*innen virtuell in das interne Netzwerk (Intranet) der Institution einzubinden. Das ermöglicht die Nutzung von Diensten, die auf der Basis der Netzwerkadresse (IP-Adresse) entscheiden, ob der Zugriff ermöglicht wird. Viele Dienste einer Institution werden über eine *Firewall* aus Sicherheitsgründen dem gesamten Internet verborgen und sind nur über ein VPN auch außerhalb der Institution zum Beispiel aus dem Homeoffice zugreifbar.

### Werkzeuge zur Kommunikation

Wie im [Kapitel zur Kommunikation](#kommunikation) beschrieben, gibt es verschiedene Werkzeuge zur synchronen und asynchronen Kommunikation von **Telefon** über **E-Mail** und **Chats** bis zu **Videokonferenzsystemen**.

Neben der reinen Kommunikation dienen [bereitgestellte Werkzeuge](#werkzeuge) wie **Wikis** und **Ticketsystem** auch der Zusammenarbeit und Erstellung von Inhalten.

### Fileservices

Dienstlich genutzte Dateien sollten an zentraler Stelle abgelegt werden, damit sie in eine Backup-Lösung eingeschlossen werden können und damit die Möglichkeit besteht, sie mit anderen Menschen auszutauschen. Entweder wird dazu einen klassischer Fileserver wie zum Beispiel *Windows Server* genutzt, oder es kommt eine Cloudlösung wie *Nextcloud* zum Einsatz. Zusätzlich oder alternativ kann ein [Dokumentenmanagementsystem](#dokumentenmanagementsysteme) genutzt werden.

### Dokumentenmanagementsystem (DMS)

Dokumentenmanagementsysteme sind Multi-User Softwaresysteme mit Anbindung an eine hinreichend großen und ausfallsicheren Datenspeicher. Sie lösen in der Regel drei Anforderungen:

* langfristige ggf. auch revisionssichere Ablage digitalisierter oder rein digitaler Dokumente die einer Aufbewahrungsfrist unterliegen (Archivierung, Versionierung)
* Unterstützung der Datenverarbeitung für Prozesse zwischen verschiedenen Akteuren (Workflows)
* Strukturierung und Pflege der Dokumente institutions-relevanter interner und externer Prozesse (Aktenplan)

Ab und an erhalten Bibliotheken den Auftrag, die Originale von digitalisierten und in einem DMS abgelegten Dokumenten physisch zu archivieren oder auch den Digitalisierungsprozess zu verantworten. Hierbei ist es empfehlenswert, zwischen Unternehmens- bzw. institutionskritischen Dokumenten, die nicht für die Bibliotheksnutzer verfügbar sein sollen, und Dokumenten mit bibliothekarischem Bezug zu unterscheiden, denn eine Bibliothek ist im Allgemeinen kein Archiv.

## Zusammenfassung und Ausblick

Die technische Infrastruktur bildet die Grundlage für die Dienste einer Bibliothek. Während sich die grundlegenden Dienste für Mitarbeiter\*innen nicht wesentlich von anderen Einrichtungen unterscheiden, sind viele Dienste für Nutzer\*innen an die Verwaltung physischer Medien gekoppelt.

# IT-Management

|  |
| --- |
| Zusammenfassung |
| Dieses Kapitel beschreibt die Einführung und den [Betrieb](#betriebsmodelle) von IT-Systemen über den gesamten [Lebenszyklus](#lebenszyklen-von-it-systemen). |

## Einleitung

IT-Systeme sind selten statisch sondern folgen einem [Lebenszyklus](#lebenszyklen-von-it-systemen) von der Planung bis zu ihrer Ablösung. Während des [Betriebs](#betriebsmodelle) der Systeme müssen mögliche [Risiken](#betriebssicherheit-und-risikomanagement) beachtet und [rechtliche Rahmenbedingungen](#rechtliche-rahmenbedingungen) eingehalten werden. In Bibliotheken sind daher entsprechende [IT-Kompetenzen](#kompetenzen) und ein [organisatorischer Rahmen](#organisation) für die [Digitale Souveränität](#digitale-souveränität) gefordert. Um diesen Anforderungen begegnen zu können, gibt es Möglichkeiten zur [Aus- und Weiterbildung](#aus--und-weiterbildung).

## Lebenszyklen von IT-Systemen

Alle Software-Systeme folgen einem Lebenszyklus, der mit ihrer Einführung beginnt und irgendwann mit ihrer Abschaltung endet ([Abbildung 2.1](#fig-it-zyklus)). Die wesentlichen Phasen im klassischen Lebenszyklus eines IT-Systems werden im Folgenden näher betrachtet. Darüber hinaus wird erläutert wie Änderungen an IT-Systemen in Institutionen im Rahmen des [Change Management](#change-management) begleitet werden sollten.

Die konkrete Abfolge vor allem der ersten Phasen kann je nach der angewendeten Projektmanagement-Methode (agil vs. klassisch) variieren. Eine Diskussion von agilen und klassischen Methoden liegt außerhalb des Fokus dieses Handbuchs.

|  |
| --- |
| Abbildung 2.1: Softwareentwicklungs-Lebenszyklus |

### Planung und Analyse

Grundlage für die Umsetzung eines Softwareprojekts, egal ob es sich um individuell erstellte Software oder die Anpassung eines existierenden IT-Systems handelt, ist ein gemeinsames Verständnis für das Ziel und die Anforderungen des Projektes. Dieses gemeinsame Verständnis, insbesondere der Anforderungen, sollte bei allen Projektmitgliedern und den weiteren Stakeholdern vorhanden sein. Die Anforderungen werden idealerweise vor und während der Entwicklung [unter Einbeziehung von Nutzer\*innen](#anforderungen) ermittelt und angepasst.

Zur Planungs- und Analysephase gehört neben einer grundsätzlichen Machbarkeitsanalyse des Projekts die Zusammenstellung eines geeigneten Teams, die Bestimmung der Stakeholder sowie die Klärung finanzieller und rechtlicher Rahmenbedingungen.

*Siehe auch ausführlicher zum* [*Entscheidungsprozess bei der Einführung eines Bibliotheksmanagementsystem*](#entscheidungsprozess)

|  |
| --- |
| Info |
| Zuweilen kommt es vor, dass die Entscheidung für ein IT-System bereits getroffen ist, bevor geklärt wurde, welches Problem damit gelöst werden soll. Auch in diesem Fall ist es sinnvoll, die Einführung mit einer offenen Planung und Anforderungsanalyse zu beginnen, und danach zu prüfen, welche Anforderungen das System tatsächlich abdecken kann. |

### Design/Prototyping

Während der Design- bzw. Prototyping-Phase entwickeln Designer\*innen und Entwickler\*innen erste Prototypen. Ziel ist es dabei, Feedback der verschiedenen Stakeholder zu erhalten, um gemeinsam ein besseres Verständnis der Anforderungen zu erhalten bzw. diese zu präzisieren. Das Kapitel [Anforderungen an Bibliotheks-IT](#anforderungen) geht gesondert auf die Bedeutung dieser Einbeziehung und damit verbundener Methoden ein.

### Implementierung

Aufbauend auf einem gemeinsamen Verständnis der Anforderungen überführen Entwickler\*innen Prototypen in lauffähigen Code. Wird im Rahmen des Projekts ein bestehendes System implementiert, werden die Prototypen zunächst in ein Testsystem und in der Folge in das produktive System überführt.

In klassischen Projekten sieht man in dieser Phase zuerst ein Produkt mit den gewünschten Features, während nutzer\*innenorientierte Vorgehensmodelle (siehe Kapitel [Anforderungsanalyse](#anforderungen)) hier auf einen iterativen Prozess setzen, welcher Produktiterationen häufiger bereitstellt und evaluiert.

Grundsätzlich unterscheidet sich die Implementierung von Informationssystemen in und für Bibliotheken nicht von der Softwareentwicklung in anderen Bereichen. Unabdingbar ist der Einsatz eines Versionskontrollsystems, ein Issue-Tracker und möglichst automatische Tests und Deployment (kontinuierliche Integration), so dass Änderungen am Quellcode direkt zu einer Aktualisierung der Test- und/oder Produktiv-Instanz der installierten Software führen.

### Test und Integration

Als letzte Lebensphase vor der Produktivschaltung werden Abnahmetests und die Integration des entwickelten bzw. erworbenen Systems in die Zielumgebung durchgeführt. Im Falle der Inanspruchnahme eines Dienstleisters wird hier auch dessen Leistung final abgenommen, wenn das System erfolgreich produktiv in Betrieb genommen werden kann.

Während der Tests wird korrekte Umsetzen der Anforderungen sowie die Umsetzung der Anforderungen geprüft.

#### Wartung

Die Wartungsphase folgt auf die Produktivsetzung des IT-Systems. In dieser Lebensphase wird das System nicht mehr grundlegend weiterentwickelt, es werden jedoch Fehler (Bugs) entfernt und Anpassungen der Funktionsweise im Sinne der Parametrisierung oder die Optimierung der Programmabläufe vorgenommen.

Typischerweise finden sich IT-Systeme, die grundlegende Geschäftsprozesse abbilden oder die nach individuellen Anforderungen erstellt wurden, viele Jahre in dieser Phase.

|  |
| --- |
| Beispiel |
| [Abbildung 2.2](#fig-verweildauer) illustriert die Lebensspanne einiger ausgewählter Nachweissysteme der Staatsbibliothek zu Berlin, die zum Zeitpunkt der Erstellung dieses Textes erst teilweise abgelöst wurden.   |  | | --- | | Abbildung 2.2: Lebenszeit (in Jahren) von Bibliothekssystemen in der Wartungsphase am Beispiel der Staatsbibliothek zu Berlin (Stand 2022) | |

### Ablösung

Die Ablösung eines Systems kann eine Vielzahl an Gründen haben. So entwickeln sich die technischen Möglichkeiten und die Anforderungen der Nutzer\*innen kontinuierlich weiter. Eine Ablösung kann aber auch durch technische Obsoleszenz erzwungen werden, wenn zugrundeliegende Software-Komponenten wie das Betriebssystem oder ein Datenbankmanagementsystem nicht mehr sicher betrieben werden können.

Die konkrete Ablösungsplanung sollte mit genügend zeitlichem Vorlauf begonnen werden. Dies gewährleistet die Arbeitsfähigkeit in der Ablösungsphase. So können in der Vorphase beispielsweise notwendige Daten migriert werden, die vom Altsystem vorgehalten werden.

Mit dem frühzeitigen Beginn der Ablösungsplanung noch in der Wartungsphase können zudem vermeidbare Risiken minimiert werden. Dies ist insbesondere deshalb wichtig, da man als Betreiber eines IT-Systems nicht alle Faktoren kontrolliert, welche eine kurzfristig notwendig werdende Ablösung des Systems verursachen können. Darunter fallen zum Beispiel:

* die Abschaltung wegen technischer Obsoleszenz (s.o.),
* der Ausfall des Systems durch Hardware-Ausfälle,
* die Ankündigung von Wartungsarbeiten und Sicherheits-Patches durch den Hersteller oder die Insolvenz des Herstellers (insb. bei proprietärer Software) oder
* die De-Facto-Unwartbarkeit durch den Wegfall geeigneten Personals mit Spezialkenntnissen (z.B. veralteter Programmiersprachen), siehe dazu auch den Abschnitt [Ressourcenplanung](#ressourcenplanung)

Letztlich führen all diese Punkte zur Abschaltung eines IT-Systems aufgrund von IT-Sicherheitsproblemen, da diese Einbrüche in die Systeme (Hacks) begünstigen. Hinzu kommt das Risiko von Datenverlusten, entweder durch physischen Verlust im Falle eines Hardware-Defekts oder durch den logischen Verlust, da z.B. proprietäre Datenformate nicht mehr gelesen werden können.

Der Weiterbetrieb eines IT-Systems ohne Ablösungsplanung birgt hohe Risiken in sich und kann eine Organisation folglich in ernsthafte Schwierigkeiten bringen, insbesondere wenn geschäftskritische Prozesse betroffen sind.

### Change Management

Die Umgestaltung von Prozessen und Tools in Bibliotheken ist ein Wandel, der die Organisation und Kultur des Hauses in ihrer Ganzheit berührt. Das Change Management – also die planvolle Steuerung von Veränderungsprozessen – spielt dabei eine zentrale Rolle. Es gilt, sowohl technologische Aspekte als auch menschliche Faktoren sorgsam in Betracht zu ziehen, um nachhaltige, akzeptierte und effektive Lösungen zu implementieren. Dies betrifft nicht nur die Auswahl und Einführung von Werkzeugen und Plattformen, sondern auch die Entwicklung von Kompetenzen, die Gestaltung von Arbeitsprozessen und die Förderung einer Kultur der Offenheit und Zusammenarbeit:

* Realen Bedarf ermitteln, Ziele und Zielgruppen definieren, keine „Solutions looking for a Problem“ einführen: welches (kommunikative) Problem will ich lösen?
* Einsatz bestehender Tools prüfen: brauche ich wirklich etwas Neues, oder kann ich mit gewissen Abstrichen leben und dafür ein bestehendes Tool kreativ einsetzen? Die Gefahr beim Einsatz einer großen Vielzahl von Tools ist, dass Information / Wissen zu stark in verschiedene Systeme fragmentiert wird. Die Gefahr beim Einsatz von zu unspezifischen Tools ist, dass die Lösung dem eigentlichen Problem nicht gerecht wird.
* Auf Interoperabilität und offene Standards achten; Vendor Lock-In vermeiden
* Die angestrebten Veränderung frühzeitig und offen kommunizieren, Unterstützung in Form von Dokumentation, Schulungen und Support anbieten, Ängste nehmen (sowohl bei Kolleg\*innen als auch Gremien). Der Trend hin zu kollaborativen Tools sorgt für eine höhere Sichtbarkeit, das ist vielen nicht ganz geheuer. Oft ist es eine große Hemmschwelle, auf einmal für alle sichtbar in ein Wiki oder Forum zu schreiben
* Bereit sein, Verfahren und Werkzeuge auch wieder abzuschaffen („alte Zöpfe auch abschneiden“). Die Einführung zusätzlicher Werkzeuge sorgt sonst auch oft für Unverständnis und Frust.
* Akzeptanz regelmäßig prüfen (siehe [Evaluation](#evaluation))

## Betriebsmodelle

Insbesondere serverbasierte Software, wie zum Beispiel das [Bibliotheksmanagementsystem](#bibliotheksmanagementsysteme), kann auf verschiedene Arten betrieben werden. Die Betriebsarten unterscheiden sich bezüglich Installation, Kosten, Pflege und Wartung sowie Backup und Support. Die Wahl eines Betriebsmodells ist nicht nur eine betriebswirtschaftliche Entscheidung sondern beeinflusst auch strategisch die [Digitale Souveränität](#digitale-souveränität) der Einrichtung.

### Lokale Installation

Bis etwa 2010 war diese Betriebsart der Normalfall: Eine Einrichtung erwarb die Lizenz für eine (Server-)Software, entweder als Einzelkauf oder im Abo, und installierte diese auf eigenen Servern, z.B. im Serverraum der Bibliothek. Im Fachjargon spricht man auch von einer „on-premise“ Installation.

In diesem Modell kümmert sich die Einrichtung selbst um Installation und Updates. Folglich erfordert dieses Modell höheren Personaleinsatz und kann dazu führen, dass bei einem personellen Engpass eine Software länger betrieben bzw. nicht aktualisiert wird, als eigentlich ratsam wäre. Auch muss sich die Einrichtung um grundlegende Dinge, wie Backups und Ausfallsicherheit selbst Gedanken machen.

Auf der anderen Seite bietet dieses Modell der Einrichtung die meiste Kontrolle über die eingesetzte Software - etwa hinsichtlich nötiger Erweiterung oder Anpassung - und macht sie damit weitgehend unabhängig von äußeren Einflüssen.

### Hosting

In diesem Betriebsmodell wird die Ebene der Rechenkapazität bzw. Serverhardware an einen Dienstleister ausgelagert. Der Dienstleister kann hierbei etwa das Rechenzentrum einer Universität oder des angeschlossenen Bibliotheksverbundes sein, oder ganz allgemein jeder kommerzielle Betreiber eines Rechenzentrums, bei dem Kapazitäten erworben werden.

Sämtliche Betriebsfragen, wie Backups und Ausfallsicherheit der eingesetzten Hardware können an diesen Anbieter delegiert werden. Im Falle des Hostings durch einen Bibliotheksverbund entfallen möglicherweise auch Einrichtung, Installation und Upgrades. Die Betriebskosten müssen beim Verbund kalkuliert werden, was jedoch durch das Hosting für mehrere Einrichtung besser skaliert.

### Cloud

Bei diesem Betriebsmodell, das manchmal auch als SaaS (Software as a Service) bezeichnet wird, liegt der fachlich und technische Betrieb beim Anbieter bzw. Dienstleister der Software und die Einrichtung nutzt lediglich eine für sie konfigurierte Installation („Instanz“). Dies ist insbesondere bei webbasierten Anwendungen die bevorzugte Betriebsart, stellt aber erhöhte Anforderungen an die Anbindung lokaler Endgeräte wie z.B. [Ausleihautomaten](#selbstverbucher-ausleihautomaten)), weil dabei eine sichere und stabile Verbindung zwischen den lokalen Automatisierungsgeräten und dem entfernt gehosteten System hergestellt werden muss.. Die Einrichtung ist weder für die Wartung der eingesetzten Hardware noch für die Pflege der genutzten Software zuständig.

In der Praxis kann sich ein solches Betriebsmodell als komfortabel erweisen, da keine Personalressource für allgemeine Tätigkeiten des IT-Betriebs oder spezielle Bibliotheks-IT-Tätigkeiten benötigt werden. Gerade für kleine Einrichtungen kann dies ein guter Weg sein, möglichst personalsparend serverbasierte Software einzusetzen. Eine Kostenersparnis ist bei einer Vollkostenrechnung aber nicht unbedingt zu erwarten. Je nach Größe der Einrichtung oder basierend auf der Anzahl der Endnutzer\*innen führt ein solches Betriebsmodell meist zu Abonnementkosten.

### Digitale Souveränität

Digitale Souveränität bedeutet die selbstbestimmte Gestaltung und Kontrolle über die eigenen digitalen Ressourcen, Technologien und Daten, um unabhängige, nachhaltige und nutzerzentrierte Services gewährleisten zu können.

Sie bietet Bibliotheken die Möglichkeit, sowohl ihren Service als auch ihre interne und externe [Kommunikation](#kommunikation) auf eine nachhaltige, unabhängige und innovative Weise zu gestalten und weiterzuentwickeln, während gleichzeitig die Prinzipien der Offenheit und Gemeinschaftlichkeit im Umgang mit Wissen und Technologie gefördert werden.

Digitale Souveränität repräsentiert für Bibliotheken damit nicht nur einen ethischen Imperativ, sondern führt auch zu einer überzeugenden strategischen Ausrichtung:

* **Unabhängigkeit von Dienstleistern:** Durch die Selbstverwaltung und -steuerung der eingesetzten digitalen Werkzeuge können Bibliotheken spezifische Anpassungen vornehmen und sind nicht auf externe Anbieter angewiesen, wodurch auch die Datenkontrolle intern bleibt.
* **Aufbau von Know-how im Haus:** Die kontinuierliche Arbeit mit und Entwicklung von digitalen Werkzeugen befähigt das Personal, ein vertieftes technisches Verständnis und Fertigkeiten zu entwickeln, was langfristig zur Selbstständigkeit und Innovationskraft beiträgt.
* **Stärkung von Kooperation durch Open Source:** Gemeinsame Entwicklungsprojekte im Open-Source-Bereich unterstützen nicht nur den Wissensaustausch und die Innovationskraft unter Bibliotheken, sondern fördern auch die Entstehung robuster und anpassbarer Technologien, die gemeinschaftlich weiterentwickelt werden können.
* **Öffentliches Geld für öffentliche Software:** Diese Philosophie unterstreicht das Bestreben, mit öffentlichen Mitteln finanzierte Software auch der gesamten Gemeinschaft zugänglich zu machen, wodurch der Grundsatz der Transparenz und des gemeinschaftlichen Nutzens betont wird.
* **Finanziell attraktiv:** Durch die Vermeidung von Lizenzkosten und die Möglichkeit, Software nach Bedarf anzupassen und weiterzuentwickeln, eröffnen Open-Source-Werkzeuge nicht nur kosteneffiziente, sondern auch zukunftssichere Investitionspfade.
* **Datenschutz und -sicherheit:** Eigenkontrolle über die Systeme kann zu besseren Sicherheits- und Datenschutzpraktiken führen, da Bibliotheken direkten Einfluss auf die Verwaltung und den Schutz von Benutzerdaten haben.
* **Inklusion und Zugänglichkeit:** Die Fähigkeit, digitale Werkzeuge selbst zu gestalten und anzupassen, ermöglicht es Bibliotheken, gezielt auf inklusive Praktiken und die Schaffung barrierefreier Dienste und Ressourcen hinzuwirken.

## Betriebssicherheit und Risikomanagement

Neben den Problemen der Ablösungplanung gibt es weitere Risiken des Betriebs von IT-Systemen, von denen einige im nachfolgenden Abschnitt vorgestellt werden.

### Vendor-Lock-In

Ein nicht zu unterschätzendes Risiko, welches sich aus der Einführung eines proprietären IT-Systems ergibt, ist der sogenannte Vendor Lock-In. Dieser beschreibt die Abhängigkeit von Produkten oder Dienstleistungen eines Anbieters durch die der gleichzeitige Einsatz von anderen Produkten oder der Wechsel zu anderen IT-Systemen erschwert wird. Durch den Einsatz von Systemen mit etablierten Standards, offenen Datenformaten und Schnittstellen sowie geeigneter Ablösungsstrategien kann das Risiko eines Vendor Lock-Ins verringert werden.

Der Begriff des Vendor Lock-ins kann noch auf den Bereich der Fehlerbehebung und die Wartung von Software ausgedehnt werden. Im Fall von proprietärer Software, welche ohne Zugriff auf den Quellcode betrieben wird, ist die Fehlerbehebung ausschließlich Sache des Herstellers. Fällt dieser, wie oben beschrieben, aus, kann ein Betrieb aus IT-Sicherheitsperspektive nicht mehr verantwortet werden. Hinzu kommt, dass das sogenannte Reverse Engineering bzw. das Dekompilieren dieser Software in der Regel verboten ist. Mit einem [Grundsatzurteil des EuGH aus dem Jahr 2021](https://curia.europa.eu/juris/document/document.jsf?text=&docid=247056&pageIndex=0&doclang=DE&mode=lst&dir=&occ=first&part=1&cid=9912038) wird dieses Verbot jedoch aufgeweicht. So ist es nun rechtmäßigen Erwerbern erlaubt, Fehler in einem Computerprogramm zu beheben und dafür auch proprietäre Software zu dekompilieren.

In der Praxis sollte dieses Notfallszenario aber nicht in die Planung einbezogen werden, da die Fehlerbehebung innerhalb fremder Software unter dem Rückgriff auf Dekompilierung besondere Kenntnisse seitens des zuständigen IT-Personals voraussetzt.

### Software-Abhängigkeiten

Sowohl der Betrieb von proprietärer als auch von Open-Source-Software ist vom Funktionieren einer Vielzahl weiterer Software-Komponenten abhängig. Diese Abhängigkeit lässt sich mit einem vereinfachten Schichtmodells des Betriebs eines IT-Systems illustrieren:

|  |
| --- |
| Abbildung 2.3: Schichtmodell-Bild (Platzhalter) |

Aus [Abbildung 2.3](#fig-schichtmodell) wird deutlich, dass moderne Software-Systeme zum Beispiel auf einem Betriebssystem oder weiteren Subsystemen wie einem Datenbankmanagementsystem basieren. Um das gesamte IT-System betreiben zu können, müssen die Einzelkomponenten zusammen spielen. Fällt eines der Systeme, beispielsweise das Betriebssystem, aufgrund von Obsoleszenz aus, so ist es unter Umständen möglich, die darüber liegenden Schichten auf ein neues Betriebssystem zu migrieren, jedoch ist dies nicht garantiert.

Das Risiko erhöht sich, wenn im Rahmen eines Wartungsvertrags durch den Hersteller festgelegt wurde, dass zum Beispiel nur bestimmte Kombinationen aus Betriebssystem und weiterer Komponenten zugelassen sind. In diesem Fall kann ein IT-System aus der Wartung fallen, obwohl es vorerst betreibbar bleibt. Mit dem Ausfall der Wartung entfallen auch Software-Updates etc. Damit ist der mittel- bis langfristige Weiterbetrieb des Systems ohne Gefährdung der Betriebssicherheit aller IT-Systeme der Organisation nicht möglich.

### Rechtliche Rahmenbedingungen

Die meisten Bibliotheken befinden sich in öffentlicher Hand und sind deshalb bestimmten Gesetzen und Verordnungen unterworfen. Von besonderer Bedeutung sind dabei Anforderungen an die Software-Ergonomie und die Barrierefreiheit (Accessibility) von IT-Systemen.

### Software-Ergonomie

Die gesetzliche Unfallversicherung fordert z.B. die Berücksichtigung ergonomischer Grundsätze bei der Entwicklung von Software. Moderne grafische Anwendungen müssen ebenso wie Internetseiten diese Anforderungen erfüllen:

Die Software muss gebrauchstauglich sein, das heißt, sie sollte gewährleisten, dass Benutzer festgelegte Ziele in einem bestimmten Nutzungskontext effektiv, effizient und zufriedenstellend erreichen können. Dies setzt voraus, dass die Grundsätze der Dialoggestaltung nach *DIN EN ISO 9241-110*, wie *Aufgabenangemessenheit, Selbstbeschreibungsfähigkeit, Steuerbarkeit, Fehlertoleranz, Erwartungskonformität, Individualisierbarkeit, Lernförderlichkeit* beachtet und realisiert werden.

– (Gesetzliche und Unfallversicherung e.V. (DGUV) 2019)

Die Erreichung dieser Ziele wird im Kapitel [Anforderungsanalyse](#anforderungen) thematisiert.

### Barrierefreiheit

Neben dem Befolgen der Anforderungen an ergonomisch bedienbare Software, liegt es auf der Hand, dass IT-Systeme für eine Vielzahl von Anwender\*innen nutzbar sein sollte. Diese grundlegende Anforderung bezeichnet man als Barrierefreiheit bzw. Accessibility.

Während Barrierefreiheit häufig mit einem sehr engen Behinderungsbegriff assoziiert wird, wie z.B. die Rampe für Rollstuhlfahrer\*innen, ist dieser Begriff mittlerweile aufgrund der gesetzlichen Grundlagen in Deutschland wesentlich weiter zu fassen (siehe [§3 Behindertengleichstellungsgesetz](https://www.gesetze-im-internet.de/bgg/__3.html)). So leiten sich z.B. auch Anforderungen an Angebote in Leichter Sprache o.ä. aus diesem weiten Behinderungsbegriff ab.

Aufbauend auf der einschlägigen Gesetzgebung regelt die [BITV 2.0](https://www.gesetze-im-internet.de/bitv_2_0/index.html) (Verordnung zur Schaffung barrierefreier Informationstechnik nach dem Behindertengleichstellungsgesetz) die konkrete Gestaltung barrierefreier IT-Systeme und Webangebot. Hierbei greift sie auf die aktuell gültigen Web Content Accessibilty Guidelines (WCAG) zurück, welche die Anforderungen der Barrierefreiheit anschaulich mit vielfältigen Beispielen illustriert.

Die gesetzliche Anforderung Barrierefreiheit umsetzen zu müssen trifft dabei nicht nur auf Organisationen der öffentlichen Verwaltung, wie es viele Bibliotheken sind, zu. Vielmehr müssen sich alle Stellen, die europäisches Vergaberecht anwenden müssen (siehe EU Richtlinie 2016/2102), z.B. im Rahmen von Drittmitteln oder Zuwendungen, nach diesen Vorgaben richten.

Wichtig ist hierbei zu beachten, dass die BITV 2.0 nicht zwischen internen und externen Nutzer\*innen unterscheidet. Das heißt, dass sowohl rein bibliotheksintern genutzte Systeme als auch nach außen gerichtete Anwendungen, wie z.B. Discovery-Systeme, barrierefrei zu gestalten sind.

Neben der Ermöglichung der digitalen Teilhabe für den Großteil der Bevölkerung sollen abschließend noch drei weitere positive Aspekte der Beachtung von Barrierefreiheitsanforderungen genannt werden:

1. Viele Umsetzungen der Grundsätze der Barrierefreiheit, erhöhen auch die Usability für Menschen ohne Behinderung,
2. ebenso wird zumeist die Nachnutzbarkeit auf Mobilgeräten erhöht und
3. die Gestaltung barrierefreier Anwendungen schlägt sich positiv im Suchmaschinenranking nieder, da [Suchmaschinen diesen Aspekt zur Bewertung heranziehen](https://developers.google.com/search/blog/2020/05/evaluating-page-experience).

## Management der Bibliotheks-IT

Die Auswahl und Implementierung sowie der Betrieb von digitalen Diensten ist ein stetig wachsender Aufgabenbereich für Bibliotheken. Durch den Verlust ihres früheren Monopols auf die Versorgung mit Informationen ist Ende der 1990er Jahre in den Bibliotheken ein starker Innovationsdruck entstanden. In der Folge wurden neue Dienstleistungen wie fachliche Portale, Dokumentenserver etc. im Rahmen von Projekten realisiert. Die notwendigen Kenntnisse haben vielerorts eigens dafür eingestellte Mitarbeiter\*innen eingebracht. Eine systematische Ausweitung von Kenntnissen zu IT-Systemen, Metadatenmanagement, Web-Standards, Usability und User Experience bleibt jedoch für die meisten Bibliotheken eine große Herausforderung.

### Kompetenzen

Der Einsatz von IT-Systemen erfordert dezidierte Kenntnisse und Fähigkeiten. Veränderungen der Systemlandschaft, z.B. durch einen Systemumstieg oder die Einführung neuer Systeme, erfordern daher eine regelmäßige Analyse der vorhandenen und benötigten IT-bezogenen Kompetenzen.

Bei der Analyse sind folgende Bereiche zu unterscheiden:

* Basisinfrastruktur für Hard- und Software: Infrastruktur Netze/Hardware, Netzdienste, Server, Basis-Software, Storage, Backup
* Systemadministration: Installation und Betrieb
* Inbetriebnahme und Individualisierung: Metadatenmanagement, Konfiguration und Parametrisierung, Software-Entwicklung
* Projektmanagement

Für die Systemadministration werden Kenntnisse zu grundsätzlichen Architekturen webbasierter Systeme benötigt. In der Regel handelt es sich dabei um die Installation und Administration von Datenbanken, PHP- und Java-Systemen auf Linux-Servern, einschließlich der jeweiligen Wartung durch Minor und Major Updates, Patches etc.

Für die Inbetriebnahme und Individualisierung von Systemen sind zunächst Kenntnisse bei der Umwandlung von Daten in unterschiedliche Formate notwendig, wenn eine Datenmigrationen aus Altsystemen durchgeführt wird. Die Systeme müssen dann konfiguriert und parametrisiert werden, das heißt auf die konkreten Nutzungsszenarien angepasst. Hier ist zwischen einer so genannten Out-of-the-Box-Verwendung, bei der das System im Rahmen seiner Möglichkeiten genutzt wird, und einer weitergehenden Individualisierung durch eigene Software-Entwicklung zu unterscheiden. In beiden Fällen ist ein Zusammenspiel von bibliotheksfachlichen und Software-technischen Kompetenzen erforderlich, um das Verständnis von bibliothekarischen Geschäftsgängen und den Prozessen und Funktionalitäten des Systems zusammen zu bringen. Daher werden in der Regel Implementierungsteams aus Anwender\*innen und Software-Betreuer\*innen bzw. Entwickler\*innen gebildet

Die Arbeit dieser Implementierungsteams sollte idealerweise nach Grundlagen des Projektmanagements und des [Lebenszyklus von IT-Systemen](#Lebenszyklen-von-IT-Systemen) erfolgen, also unter Berücksichtigung klarer Strukturen für die Planung, die Kommunikation und die Kontrolle. Siehe dazu auch das Kapitel [IT-Entwicklung].

|  |
| --- |
| Tipp |
| Niemand kommt als IT-Expert\*in auf die Welt und es ist praktisch unmöglich bei allen Entwicklungen auf dem Laufenden zu bleiben. Versuchen Sie ihre Kompetenzen realistisch einzuschätzen und scheuen Sie sich nicht Kolleg\*innen um Rat zu fragen! |

### Organisation

Größere Bibliotheken haben in der Regel eigene IT-Abteilungen, die folgende Aufgaben übernehmen:

* Anwendungsbetreuung bei Hard- und Software sowie Peripheriegeräten,
* Betreuung von eigenen Servern,
* Konfiguration und Parametrisierung von Systemen,
* ggf. Software-Entwicklung.

Alle Punkte außer der Konfiguration und Parametrisierung von Systemen werden in der Regel von informatisch ausgebildetem Personal vorgenommen. Die Konfiguration und Parametrisierung von Systemen ist klassischerweise eine Aufgabe so genannter Systembibliothekar\*innen, die teilweise in Ergänzung zu ihrer bibliothekarischen Ausbildung auch über formalisiert erworbene Zusatzqualifikationen verfügen und auf diese Weise eine ideale Schnittstelle zwischen den bibliothekarischen Anwender\*innen und den Systemen sind.

Die Betreuung eigener Server wird häufig von übergeordneten Einrichtungen wie universitären, städtischen oder regionalen Rechenzentren oder externen Dienstleistern übernommen. Bei kleinen Bibliotheken werden oftmals auch andere Tätigkeiten von externen Dienstleistern durchgeführt.

Der Umgang mit dem Mangel an IT-Fachkräften wird für die Ressourcenplanung des IT-Managements in Bibliotheken zur Herausforderung werden. Dabei wird auch Open-Source-Software die in der Community entwickelt und unterstützt wird, eine größere Rolle spielen, ebenso wie externe Dienstleister und Software as a Service. Eine umfassende Bedarfsanalyse bei IT-Systemen wird daher zukünftig noch stärker berücksichtigen müssen, wie eine längerfristige Betreuung von eingesetzten Systemen ggf. auch ohne eigenes Personal gewährleistet werden kann.

#### Ressourcenplanung

Veränderungen des Status Quo durch einen Systemwechsel, neue Funktionalitäten oder auch personelle Änderungen durch neue Mitarbeitende, Berentung o.ä. beeinflussen die Personal- und Ressourcenplanung. Eine kontinuierliche Beschäftigung mit diesem Thema ist notwendig, um das [Betriebsrisiko von IT-Systemen](#betriebssicherheit-und-risikomanagement) zu minimieren. Die Benutzung von gut etablierter und weit verbreiteter freier Software verursacht Kosten durch Kompetenzaufbau oder -einkauf, macht unabhängig in Bezug auf Datenhoheit, Datensicherheit und vor allem die Weiterentwicklung. Die Lizenznahme eines kommerziellen Produktes lässt grundsätzlich weniger Individualisierung zu und verlagert die Verantwortung für die Verfügbarkeit und Betriebssicherheit auf den jeweiligen Anbieter.

Für die Planungen muss entsprechend regelmäßig der Bedarf analysiert werden:

* Systeme: Welche Kompetenzen erfordert der Betrieb der Systeme?
* Personal: Wie viel Personal steht mit welchen Kompetenzen zur Verfügung?
* Wie verteilen sich die Kompetenzen auf das vorhandene Personal?
* Wie hoch ist die Übereinstimmung bei vorhandenen und benötigten Kompetenzen?
* Sind Weiterbildungen sind erforderlich?

Das Thema Aus- und Weiterbildung sowie die Personalgewinnung wird im Folgenden ausführlicher betrachtet.

### Aus- und Weiterbildung

In der Einleitung wird Cody Hanson (2015) zitiert: „Most importantly, all library staff must understand that our software is our library, and is everyone’s responsibility.“ Bezogen auf Einarbeitung und Weiterbildung bedeutet das, dass sich Mitarbeitende mit der (Weiter-)Entwicklung von Software ebenfalls weiterbilden und weiterentwickeln. Nur so kann die Verantwortung von allen Mitarbeitenden mit Bezug zur Bibliotheks-IT gemeinsam getragen werden.

Nachfolgend werden aktuelle Beispiele zur Aus- und Weiterbildungen mit Bezug bibliothekarischen IT-Bereich aufgeführt. Nicht betrachtet werden Szenarien wie die Einarbeitung von Anwender\*innen von IT-Systemen bei der Einführung oder dem Wechsel von Systemen.

#### Ausbildungsmöglichkeiten und Zusatzqualifizierung

Historisch gibt es keine formalisierte Ausbildung für die erwähnten Systembibliothekar\*innen. Die notwendigen Kenntnisse werden klassischerweise im Rahmen von „Training on the Job“ erworben.

Allgemeine Ausbildungen und Studiengänge im Bereich IT und Data Science bieten eine gute Grundlage, decken aber bibliotheksspezifische IT-Themen nur unzureichend ab. Stand 2022 gibt es mehrere spezielle Ausbildungsangebote für die Arbeit in der Bibliotheks-IT mit unterschiedlichen Schwerpunkten:

* berufsbegleitende Master-Studiengänge:
  + [Bibliotheksinformatik an der TH Wildau](https://www.th-wildau.de/studieren-weiterbilden/studiengaenge/bibliotheksinformatik-msc-berufsbegleitendes-studium/)
* Vollzeit-Studiengänge:
  + [Masterstudiengang Digitales Datenmanagement an der FH Potsdam und HU Berlin](https://www.fh-potsdam.de/studium-lehre/studiengaenge/digitales-datenmanagement-m)
* Kurse
  + [Zertifikatskurs Data Librarian an der TH Köln](https://www.th-koeln.de/weiterbildung/zertifikatskurs-data-librarian_63393.php)
  + [Zertifikatskurs Forschungsdatenmanagement an der TH Köln](https://www.th-koeln.de/weiterbildung/zertifikatskurs-forschungsdatenmanagement_82048.php)

Diese Zusatzqualifizierungsmöglichkeiten sind eine sehr gute Möglichkeit, um vorhandene Mitarbeiter\*innen systematisch weiterzuentwickeln und die IT-Kenntnisse in der Bibliothek zu verbreiten. Der Erwerb dieser Abschlüsse ist jedoch zeitaufwändig und passt nur für wenige Lebenssituationen.

#### Weiterbildung

Bibliothekarische Ausbildungsstätten sowie Verbundzentralen sind wichtige Akteure bei der Weiterbildung von Bibliothekspersonal und machen teilweise entsprechende punktuelle Weiterbildungsangebote. Dabei handelt es sich in der Regel um ein- oder halbtägige Angebote, die durchaus im Einzelnen Hilfestellung bieten. Für Mitarbeitende mit Bezug zur Bibliotheks-IT sollten ausdrücklich zeitliche und ggf. finanzielle Ressourcen für die Nutzung dieser Angebote bereitgestellt werden.

Auch der Besuch von Konferenzen ist eine wichtige Säule der Weiterbildung. Im Kontext der Bibliotheks-IT hervorzuheben sind hier

* Jahrestagungen der Verbundzentralen
* Bibliothekstage und -konferenzen
* Tagung der [European Library Automation Group](https://elag.org) (ELAG)
* [Code4Lib-Konferenzen](htts://code4lib.org/conference/) in den USA
* [Access Conference](https://accessconference.ca/) in Kanada
* [UKSG Annual Conference](https://www.uksg.org/events/annualconference) in Großbritannien
* Jahrestreffen der [UXLibs](http://uxlib.org/) in Großbritannien

Die wichtigste Rolle bei der Qualifizierung für die Aufgaben im Bereich Bibliotheks-IT dürfte die informelle Weiterbildung spielen. Informelle Weiterbildungsformen sind

* Anwendungstreffen: z.B. jährlich für DSpace, VuFind, Koha, FOLIO, Kitodo, OPUS, …
* Library Carpentries
* Mailinglisten, Foren und andere Kommunikationskanäle
* persönliche Kontakte, Gruppen wie UX Roundtable
* Fachpublikationen: Code4Lib Journal, Weave Journal
* Soziale Medien: Weblogs, Twitter, Discord

#### Personalgewinnung

Die Gewinnung von Personal für die Aufgaben im Bereich der digitalen Dienste ist neben der Aus- und Weiterbildung eine zweite Herausforderung. Die Gehaltsstruktur im öffentlichen Dienst ist für informatisch ausgebildetes Personal nicht unbedingt wettbewerbsfähig, so dass viele ausgewiesene IT-Stellen nur schwer besetzt werden können. Eine unmittelbare Reaktion darauf kann sein, die Vorteile der Beschäftigung im öffentlichen Dienst besser herauszuarbeiten (unkommerzielles Umfeld, gesellschaftliche Relevanz der Tätigkeiten).

Dennoch ist es erwartbar, dass Aufgaben im Bereich Bibliotheks-IT künftig stärker an Verbundzentralen oder externe Dienstleister outgesourct werden müssen.

Die Ausrichtung der bibliothekarischen Studiengänge wird die Bedarfe bei den digitalen Diensten noch stärker berücksichtigen und Studierende mit einem erhöhten Interesse an den Aufgaben in der Bibliotheks-IT rekrutieren müssen.

## Zusammenfassung und Ausblick

Auch nach Auswahl eines Systems ist eine permanente Beobachtung des [Lebenszyklus](#lebenszyklen-von-it-systemen) erforderlich. Es empfiehlt sich immer eine frühzeitige Reaktion auf sich ändernde Anforderungen. Das Wissen um das System als auch um seine Anwendung müssen ebenfalls aktuell gehalten werden, z.B. durch entsprechende [Fortbildungen oder Schulungen](#aus--und-weiterbildung). Sollte sich ein System-Umstieg abzeichnen, sind vor allem die internen Arbeits-Prozesse zu berücksichtigen: das Wissen der Systemanwendenden und -betreuenden ist somit unverzichtbar, denn nur dadurch kann auf eine Ablösung bzw Anpassung des Systems effektiv reagiert werden.

# Anforderungsanalyse

|  |
| --- |
| Zusammenfassung |
| Dieses Kapitel behandelt die Anforderungsanalyse als Kernbestandteil bei der Entwicklung von IT-Systemen. Der Schwerpunkt liegt dabei auf der [Einbeziehung von Nutzer\*innen](#einbeziehung), um die tatsächlichen Bedürfnisse möglichst passend zu erfüllen. |

## Einleitung

Im Kapitel [Management von IT-Systemen](#management) wurde bereits auf Themen wie Barrierefreiheit und software-ergonomische Anforderungen sowie den permanenten Anpassungsbedarf an Systeme im Laufe ihrer Lebenszeit eingegangen.

Betrachtet man sein persönliches Nutzungsverhalten im digitalen Bereich wird klar, dass sich auch die eigenen Präferenzen bezüglich der Nutzung von Apps oder Webseiten ändern. Ursachen dafür sind beispielsweise Veränderungen an Lebens- oder Arbeitskontexten, Erwartungen an die Bedienbarkeit von Systemen oder durch die digitale Transformation beziehungsweise technischen Fortschritt möglich gewordene neue Nutzungsformen von Medien.

IT-Entwicklung sollte sich daher auch an den [Bedürfnissen](#bedarfsermittlung) von [Nutzer\*innen](#einbeziehung) ausrichten. Es gibt verschiedene Methoden, die entsprechenden Bedarfe und Anforderungen zu ermitteln und sie in die Entwicklung einzubeziehen. Dazu gehören unter anderem der Einsatz von Personas, Use Cases oder Storyboards. Weitere [Methoden](#methoden) sind zum Beispiel Storyboards, Wireframes oder auch Prototypen.

## Nutzer\*innenorientierten Gestaltung

Nutzer\*innenorientierte Gestaltung heißt, die Bedürfnisse von Nutzenden in den gesamten Entwicklungsprozess einzubeziehen. Das bedeutet, dass deren Bedarfe nicht nur als Quelle von initialen Anforderungen dienen, sondern kontinuierlich in den Entwicklungsprozess einbezogen werden. Hierbei ist es besonders wichtig, die Fähigkeiten und Bedürfnisse der Nutzenden sowie ihre Arbeitskontexte und -aufgaben in den Entwurf von IT-Systemen einzubeziehen. Diese Aspekte finden sich auch in den zugrundeliegenden Definitionen, wie der Usability (siehe Abschnitt [Was beeinflusst den Nutzungseindruck?](#was-beeinflusst-den-nutzungseindruck)) wieder.

Beim nutzer\*innenorientierten Design oder dem User-Centered Design (UCD) handelt es sich nicht im formale Methoden im engeren Sinn, sondern um eine Sammlung i.d.R. empirisch abgesicherter Techniken mit drei Kernideen (Gould und Lewis 1987).

1. Fokussierung auf Nutzer\*innen und deren Aufgaben von Beginn der Entwicklung an
2. deren kontinuierliche Einbeziehung und Auswertung von Nutzer\*innen-Feedback sowie Performance-Messung
3. Nutzung eines iterativen Design-Prozesses

Kling und Leigh Star ergänzen, dass die ganz individuellen Fähigkeiten der Nutzenden in Betracht gezogen werden müssen (Kling und Star 1998), was allein schon aus Gründen der digitalen Teilhabe sinnvoll erscheint.

Generell zielt UCD darauf ab, interaktive Systeme zu entwickeln, welche einfach zu nutzen und nützlich sind. Hierbei wird ein Fokus auf Aspekte wie Effektivität, Effizienz, Benutzerzufriedenheit und Zugänglichkeit gelegt (Deutsches Institut für Normung e. V. (DIN) 2020). Diese Aspekte werden in Abschnitt [Was beeinflusst den Nutzungseindruck?](#was-beeinflusst-den-nutzungseindruck) weiter erläutert.

Das Central Digital and Data Office des Vereinigten Köngreichs fasst die zentral zu bearbeitenden Arbeitspunkte im nutzer\*innenzentrierten Gestaltungsprozess und den Weg dahin prägnant in seinen „Government Design Principles“ zusammen „Government Design Principles. GOV.UK“ (2012):

1. Start with user needs
2. Do less
3. Design with data
4. Do the hard work to make it simple
5. Iterate. Then iterate again
6. This is for everyone
7. Understand context
8. Build digital services, not websites
9. Be consistent, not uniform
10. Make things open: it makes things better

Unter der oben genannten Website des Central Digital and Data Office finden sich auch umfangreiche Hinweise, wie sich die einzelnen Punkte praktisch umsetzen lassen.

### Was beeinflusst den Nutzungseindruck?

Gut bedienbare, interaktive Systeme sollen Zufriedenheit auslösen und zugänglich sein. Die Erreichung dieser Ziele und zentrale Begriffsdefinitionen sind Teil des Arbeits- und Forschungsgebiets der Software-Ergonomie und finden sich in den einschlägigen Normen wie der DIN EN ISO 9241-11 wieder (Deutsches Institut für Normung e. V. (DIN) 2020).

Von zentraler Bedeutung sind dabei zwei Kernbegriffe: die Usability (Gebrauchstauglichkeit) und die User Experience (Nutzer\*innenerfahrung)

|  |
| --- |
| Tipp |
| **Usability** ist das „Ausmaß, in dem ein System, ein Produkt oder eine Dienstleistung durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um festgelegte Ziele effektiv, effizient und zufriedenstellend zu erreichen“ DIN EN ISO 9241-11  **User Experience** bezeichnet die „Wahrnehmungen und Reaktionen einer Person, die aus der tatsächlichen und/oder der erwarteten Benutzung eines Produkts, eines Systems oder einer Dienstleistung resultieren“ DIN ISO 9241-210:2011 |

Bei Usability handelt es sich um eine Eigenschaft eines Systems, die während der konkreten Interaktion mit diesem relevant wird und beispielsweise angibt, inwiefern Hürden bei der Bedienung auftreten ([Abbildung 3.1](#fig-usux)). Zur Vermeidung von Usability-Problemen existieren ein Vielzahl von Heuristiken, die in den einschlägigen Normen skizziert werden bzw. durch Autoren wie Shneiderman in seinen „8 golden rules“ (Shneiderman und Plaisant 2005) oder Nielsen mit seinen „[10 Heuristics](https://www.nngroup.com/articles/ten-usability-heuristics/)“ benannt werden.

Die User Experience hingen bezieht sich auf die Wahrnehmung der Nutzenden sowohl vor, nach und auch während der Interaktion. Sie bezeichnet sozusagen die Positionierung gegenüber einem System und hat damit Auswirkungen darauf, ob Nutzende ein System erneut benutzen werden oder, z. B. aufgrund von schlechter Bedienbarkeit, d. h. schlechter Usability, vor einer zukünftigen Nutzung zurückschrecken. Es reicht folglich nicht aus, einzelne Aspekte einer Nutzer\*innenschnittstelle zu optimieren. Vielmehr muss der gesamte angebotene Service aus Sicht der Nutzenden optimiert werden, damit sich ein positives Nutzungserlebnis einstellt. Diese Optimierung beschränkt sich dabei nicht nur auf die digitalisierten Anteile eines Services sondern bezieht alle Arbeitsschritte, egal ob analog oder digital, mit ein.

|  |
| --- |
| Abbildung 3.1: Zusammenhang zwischen Usability und User Experience |

## Wie beziehen wir unsere Nutzer\*innen ein?

Mit der Einführung einer neuen IT-Lösung werden bestimmte strategische Ziele verfolgt wie die Ablösung eines veralteten Systems, die Einführung einer neuen Dienstleistung und dergleichen. Die konkrete Ausgestaltung dieser strategischen Ziele sollte unter Einbeziehungen der beabsichtigten Nutzenden erfolgen. Die konsequente Bedarfsorientierung sichert die Qualität der Dienste und verhindert, dass eigene Bedürfnisse und Einschätzungen von Expert\*innen die Entwicklung dominieren. Für die Einbeziehung von Nutzer\*innen gibt es verschiedene Methoden, die im Folgenden kurz dargestellt werden sollen.

### Bedarfsermittlung

#### Klassische Methoden zur Bedarfsermittlung

Zu den in Bibliotheken auch jenseits der Entwicklung von digitalen Diensten häufig genutzten Methoden der Bedarfsermittlung gehören qualitative und quantitative Befragungen sowie Beobachtungen. Diese Methoden sind aus der empirischen Sozialforschung entlehnt. Für viele Software-Projekte sind groß angelegte Befragungen zu aufwändig, allerdings ist es empfehlenswert, sich über Studien aus vergleichbaren Projekten zu informieren und daraus nach Möglichkeiten Ableitungen für eigene Zielsetzungen zu entwickeln.

Beobachtungen können sehr flexibel angelegt und geplant werden. Dadurch können valide Ergebnisse mit vertretbarem Aufwand produziert werden und die Studie bei Bedarf gut skaliert werden. Der Fokus bei solchen Studien liegt darauf, Nutzer\*innen in ihrem Arbeitsalltag zu beobachten, um ihre Herangehensweise bei der Lösung von Aufgaben und Problemen zu ermitteln. Übertragen auf digitale Dienste kann das zum Beispiel im Rahmen eines Usability-Tests passieren, in dem eine oder mehrere Personen ein System nutzen. Typischerweise werden während des Tests nicht nur Notizen oder Aufnahmen gesichert, sondern die Tester\*innen nutzen das Think-Aloud-Protokoll. Dabei sollen Nutzende in Echtzeit laut kommentieren, was sie denken, sehen und tun (siehe Abschnitt [Methoden](#methoden)).

Fokusgruppen dagegen sind eine qualitative Methode, in der Vertreter\*innen verschiedener Zielgruppen gemeinsam an einem bestimmten, vorher formulierten Thema arbeiten. Das können sowohl Diskussionen über Anforderungen und Wünsche an ein bestimmtes System sein, als auch die Planung von Einsatzszenarien oder Workflows. Durch die freie Wahl von Themen und Mitgliedern, z. B. Nutzende ohne Vorerfahrungen und/oder Expert\*innen, sind Fokusgruppen ebenfalls eine sehr flexible, breit anwendbare Methode.

#### Bedarfsermittlung mit Personas und Use Cases

Personas sind fiktive Persönlichkeiten, die stellvertretend für einzelne Zielgruppen eines Dienstes entwickelt werden. Die Beschreibungen enthalten vielfältige Informationen über die Persona und laden damit dazu ein, den zu entwickelnden Dienst aus der Perspektive der jeweiligen Persona zu beurteilen, jenseits von abstrakten Anforderungen. Darüber hinaus helfen Personas dabei, Prioritäten zu setzen und die Zielerreichung zu überprüfen. Es empfiehlt sich, für jedes strategische Ziel eine Persona zu erstellen, mindestens drei bis fünf Personas insgesamt.

Abgeleitet von solchen Personas fällt es häufig leicht, konkrete Use Cases für die Interaktion mit einem System zu definieren. Ein Use Case beschreibt dabei eine Reihe von Aktionen, die eine Person in bzw. mit einem System durchführen kann. Das kann beispielsweise in einem Fließtext passieren, in dem ein Szenario beschrieben wird.

|  |
| --- |
| Abbildung 3.2: Vorgehensmodell zum szenariobasierten Design (https://opus4.kobv.de/opus4-bib-info/frontdoor/index/index/docId/3526) |

Außerdem kann es sich lohnen, solche Use Cases zu visualisieren. Dabei können Start, Ende, mögliche Verzweigungen, alternative Aktionen und mehr mit verschiedenen Formen modelliert werden. Dafür können formalisierte Systeme wie die Unified Modelling Language (UML) zum Einsatz kommen. Sie bietet ein Set verschiedener Formen, um Start, Ende, Verzweigungen, Alternativen und mehr visuell zu beschreiben. Aber auch Skizzen können Nutzungsszenarien bereits verdeutlichen und als Diskussionsgrundlage dienen, z. B. in Form von Storyboards, die in einem eigenen [Unterkapitel zu dieser Methode](#storyboards-als-frühe-methode) noch beschrieben werden.

Use Cases können sowohl als Grundlage für den Entwicklungsprozess dienen als auch für die Evaluation eines Systems (siehe Abschnitt [Evaluierung](#evaluierung)). Für die Nutzenden-Personas einer Bibliothek kann eine breite Palette von Use Cases existieren. Manche sind dabei eher allgemein zu verstehen, andere bibliotheksspezifisch und natürlich sind alle je nach Einrichtung bzw. Anforderungen beliebig erweiterbar. Zu beachten ist, dass sowohl Personas als auch Use Cases zwingend auf der Grundlage vertrauenswürdiger Daten wie denen aus der Bedarfsermittlung erstellt werden sollten. Solche Methoden ohne Kenntnisse der Zielgruppen anzuwenden kann nur zur Reproduktion der eigenen Meinung führen.

### Methoden

Testaufgaben für Usability-Tests werden erstellt, um typische Nutzungsszenarien mit Hinblick auf die Usability des Systems hin zu überprüfen. Die folgenden Methoden können relativ einfach umgesetzt werden, generieren jedoch bereits wertvolle Erkenntnisse.

#### Think-Aloud-Protokolle

Die zentrale Idee bei Think-Aloud-Protokollen ist, dass Proband\*innen während der Interaktion mit dem zu evaluierenden System ihre Meinungen, Gedanken und Gefühle laut aussprechen.

Dadurch wird es den Beobachter\*innen ermöglicht, zuvor unsichtbare, kognitive Prozesse der Proband\*innen zu beobachten sowie einen Einblick in typische Nutzungsweisen zu gewinnen. Durch die Verbalisierung und Beschreibung des Systems durch die Nutzenden lernt man zeitgleich die Nutzer\*innenterminologie für bestimmte Sachverhalte kennen, die teils erheblich von der Fachsprache abweichen wird. Die Ergebnisse der Methode können z. B. durch Notizen oder Audioaufnahmen festgehalten werden.

#### Co-Discovery Learning

Die Kernherausforderung bei der Erstellung von Think-Aloud-Protokollen ist es, die Proband\*innen kontinuierlich zu motivieren, selbst kleinste Gedanken zu verbalisieren. Beim Co-Discovery Learning arbeiten zwei Testpersonen gleichzeitig an einem System und helfen sich gegenseitig bei der Erfüllung der Aufgaben. Dadurch entstehen Gespräche und gewissermaßen automatisch ein Think-Aloud-Protokoll beider Personen.

Die Methode bildet einerseits eine realistische Arbeitssituation des gegenseitigen Helfens ab und normalisiert andererseits das laute Aussprechen von Gedanken innerhalb einer Dialogsituation.

#### Quantitative Methoden

Beobachtungsmethoden generieren primär qualitative Daten, ebenso wie viele Inspektionsmethoden. Aus Managementsicht werden jedoch oft Entscheidungen auf Grundlage von quantitativen Daten bevorzugt, da diese häufiger als Fakten wahrgenommen werden.

Einfache, relativ leicht zu erhebende quantitative Metriken im Rahmen von Usability-Tests sind z.B.:

* Nutzungsfehler pro Zeiteinheit,
* Anzahl nicht benötigter Befehle (Menus, Icons, Links)
* Benötigte Zeit für den Abschluss einer Arbeitsaufgabe (insbesondere im Vergleich mit einer vorherigen Iteration)
* Benötigte Anzahl an Klicks/Links, um an ein bestimmtes Ziel zu kommen.

Der „Benutzungsfragebogen ISONORM 9241/10“ bietet einen interessanten Kompromiss zwischen qualitativen und quantitativen Daten, da er qualitative Aussagen bezüglich der Usability eines Systems (z.B. Aufgabenangemessenheit und Selbstbeschreibungsfähigkeit) mithilfe einer siebenstufigen Likert-Skala abbildet. Der [Fragebogen](https://people.f3.htw-berlin.de/Professoren/Pruemper/instrumente.html) ist frei im Internet verfügbar. Beachtet werden muss, dass für belastbare quantitative Daten die Größe der Testgruppe deutlich steigen muss, um Verfälschungen durch Einzelpersonen zu vermeiden.

### Einbeziehung von Nutzenden in die Entwicklung

Als Grundlage für Personas oder Use Cases und alle weiteren Schritte ist die Einbeziehung von tatsächlichen Nutzenden in die Entwicklung also bereits in einem frühen Stadium möglich und sinnvoll. Diese Einbeziehung sichert ab, dass wesentliche Ziele der Nutzenden erreicht werden und in mitunter komplexen Entwicklungsprozessen die richtigen Schwerpunkte gesetzt werden. Dafür stehen verschiedene Methoden zur Verfügung.

Nachfolgend werden drei Ansätze vorgestellt:

* Storyboards - Skizzierung von Interaktionskonzepten
* Wireframes und Mock-Ups - Skizzen der Oberflächen
* Prototypen - erste funktionsfähige Iterationen

#### Storyboards als frühe Methode

Ein Storyboard illustriert, wie ein User Interface (UI, Nutzer\*innenoberfläche) auf Eingaben reagiert ohne das Interface visuell perfekt darzustellen. Es kann genutzt werden, um in Use Cases bestimmte Aktionen zu illustrieren.

|  |
| --- |
| Abbildung 3.3: Beispiel Storyboard |

Die Visualisierung von Interaktionsideen kann Beteiligten helfen, mögliche Abläufe nachzuvollziehen. Storyboards sind dabei oft leichter verständlich als z. B. technische Diagramme mit der oben genannten UML. Trotzdem ist darauf zu achten, dass Ideen und Konzepte für Stakeholder und Nutzende klar beschrieben werden, um Missverständnisse zu vermeiden. IN dieser Form lassen sich Storyboards nutzen, um z.B. verbale Beschreibungen oder Nutzungsszenarien zu ergänzen.

Durch die noch vage Darstellung der Idee können dann Diskussionen angeregt werden. Beispielsweise können Storyboards in Fokusgruppen vorgestellt und diskutiert oder auch in Einzelgesprächen mit verschiedenen Stakeholdern analysiert werden. Möglichst alle Fragen und Ideen sollten dabei ohne Limitierungen behandelt werden können und die Ergebnisse festgehalten werden.

Vor- und Nachteile von Storyboards im Überblick:

| Vorteile | Nachteile |
| --- | --- |
| leicht verständlich, für alle Stakeholder geeignet | nicht jeder Use Case oder jede Interaktionsmöglichkeit ist darstellbar |
| bereits im frühen Entwurfsprozess einsetzbar | digitale, nichtlineare Produkte (z.B. Websites) sind schwer darstellbar |
| schnelle Erstellung ohne Vorkenntnisse möglich | ggf. Unklarheiten bei der Nutzung (z.B. durch unklare Symbole) |

#### Wireframes und Mock-Ups

Wireframes und Mock-Ups werden vor allem dazu genutzt, erste Skizzen für Struktur, Layout und Funktionalitäten eines Interface vorzustellen. Ähnlich wie Storyboards dienen sie als einfach zu erstellende Diskussionsgrundlage, mit deren Hilfe ein Abgleich der Vorstellungen von einem System und der Gestaltungsmöglichkeiten durchgeführt werden kann.

|  |
| --- |
| Abbildung 3.4: Beispiel für Wireframes |

Ein Wireframe („Drahtmodell“) ist eine noch undetaillierte („low-level“) Ausarbeitung eines Interfaces, v. a. darauf ausgerichtet, die Positionierung der einzelnen Elemente zu planen. Daher sind z.B. Bilder oder Buttons als Kästchen dargestellt, Text als Striche und ähnliches (siehe [Abbildung 3.4](#fig-5)). Ein Mock-Up ist, im Kontext Design, eine ausgereifte („high-level“) Version des Interfaces mit realistischen Farben, Schriftarten und Elementen. Sowohl Wireframes als auch Mock-Ups sind also rein statische Entwürfe des zukünftigen Produkts im Gegensatz zu Prototypen, die interaktiv sind und damit echte Funktionalitäten enthalten.

#### (Interaktive) Prototypen

Die nächsthöhere Form eines geplanten Produkts, (interaktive) Prototypen, besitzen bereits erste Funktionen des geplanten Interfaces. Auch hier gibt es eine Spanne von rudimentären, low-level bis hin zu ausgereiften, high-level Prototypen, die durch Iterationen schrittweise erreicht werden. Üblich ist außerdem die Unterteilung in „vertical slice“, die qualitativ hochwertige Umsetzung nur eines bestimmten Teils des Produkts, und „horizontal slice“, die prototypische Umsetzung einer möglichst großen Bandbreite des späteren Systems.

Erste Prototypen müssen dabei noch nicht zwingend programmiert werden, sondern können durch entsprechende Prototyping Software, wie [Figma](https://www.figma.com) oder [Axure](https://www.axure.com), umgesetzt werden. Diese besitzen eine Art Bausystem für Interfaces mit mehreren Ansichten, die über Aktionen wie den Klick auf einen Button verbunden werden können. So kann Nutzenden gewissermaßen ein Produkt vorgetäuscht werden, das dann mit rudimentären Funktionen bereits getestet werden kann.

Während des eigentlichen Softwareentwicklungsprozesses wird der anfängliche Prototyp mit jeder Iteration hochwertiger und nimmt mehr den Charakter eines vollen Systems an. Es empfiehlt sich, nach Iterationen regelmäßig zu evaluieren, ob neue Funktionen oder Änderungen noch für die Zielgruppen geeignet sind.

### Evaluierung

Die vorangegangenen Abschnitte haben herausgestellt, wie wichtig es ist, regelmäßig Feedback der Nutzenden zu erhalten. Eine zentrale Datenquelle dafür ist die Begleitung eines Projekts durch Evaluierungen. Ein Beispiel für eine lebendige Evaluierungskultur ist das “[User Research Center” der Harvard Library](https://urc.library.harvard.edu/), das regelmäßig verschiedene Methoden anwendet, um Angebote gemeinsam mit Nutzenden zu evaluieren und diese öffentlich in einem [Wiki](https://wiki.harvard.edu/confluence/pages/viewpage.action?pageId=232199222) teilt.

Im Rahmen der Usability-Evaluierung entscheidet man dabei grob zwei Methoden: Beobachtungs- und Inspektionstests ([Abbildung 3.1](#fig-usux)). Während erstgenannte unter Einbeziehung von Nutzer\*innen durchgeführt werden, werden Inspektionstests häufig durch Usability-Expert\*innen realisiert.

Als Vorteil der Beobachtungstests erweist sich aus der Praxissicht, dass diese auch ohne eine formale Usability-Ausbildung durch engagierte Mitarbeiter\*innen durchgeführt werden können. Im Folgenden soll deshalb das prinzipielle Vorgehen bei einem Beobachtungstest skizziert werden.

#### Testgruppen

Die Testgruppe muss die potentielle Nutzungsgruppe bestmöglich repräsentieren, jedoch nicht sehr groß sein. Die Erfahrung zeigt, dass ca. fünf Testpersonen ausreichen, um die wichtigsten Usabilityprobleme eines Systems zu identifizieren Jakob Nielsen (2000). Statt eines einzigen Tests mit vielen Teilnehmenden bieten sich daher schnell durchzuführende Tests mit wenigen Teilnehmenden an, um ein Produkt iterativ zu verbessern. Möchte man jedoch verschiedene Typen von Nutzer\*innen analysieren oder quantitative Ergebnisse sammeln, muss die Gruppengröße entsprechend wachsen.

Neben den typischen Streuungsmerkmalen wie demographischen und kulturellen Faktoren (z.B. Bildungshintergrund) bietet es sich an, Nutzer\*innen auszuwählen, die über ein unterschiedliches Maß an Vorwissen über das zu entwickelnde oder verwandte Produkte verfügen. Außerdem sollten Personen integriert werden, welche von Einschränkungen betroffen sind, die in Abschnitt [Accessibility](#accessibility) thematisiert wurden.

#### Testablauf und Vorbereitungen

Nach der Rekrutierung repräsentativer Nutzer\*innen und der Vorbereitung der benötigten Materialien und der Testumgebung bietet sich ein Pilottest mit Proband\*innen an. Dieser dient der Validierung der eigenen Annahmen über die Testaufgaben (s.u.) und die Machbarkeit des Ablaufs.

Die Testumgebung sollte eine entspannte und natürliche Arbeitsumgebung vermitteln. Diese ist in jedem Fall einer künstlichen Laborumgebung vorzuziehen. Während der Beobachtungstests ist sicherzustellen, dass keine Unterbrechungen, z.B. in Form von Telefonanrufen erfolgen, damit die Proband\*innen das zu evaluierende System konzentriert testen können.

Nach dem Beobachtungstest sollte es den Proband\*innen ermöglicht werden, die Testergebnisse zu erhalten. Außerdem ist es neben dem obligatorischen Dank für die Teilnahme üblich, eine Aufmerksamkeit - je nach Dauer z. B. Kaffee, Süßes, Gutscheine - auszuhändigen, um die eigene Wertschätzung für das zeitliche Investment der Proband\*innen auszudrücken. In einer Erklärung zum Datenschutz ist die anonyme Datennutzung zuzusichern.

#### Testaufgaben

Wie die Testgruppen müssen auch die Testaufgaben repräsentativ für den späteren Einsatzzweck des Systems sein. Die von den Proband\*innen zu bearbeitenden Testaufgaben müssen realistische Aufgaben in bzw. mit dem System sein und in der gegebenen Zeit absolvierbar sein. Dabei ist zu beachten, dass sich die Arbeitsaufgaben an tatsächlichen Use Cases orientieren und nicht trivial sind.

Die Formulierung der Arbeitsaufgaben muss unmissverständlich für die Proband\*innen sein und auf deren (mitunter variierendes) Vorwissen eingehen. Ein Pilottest hilft, dies zu überprüfen.

Die Gestaltung der einzelnen Aufgaben sollte einer Dramaturgie folgen, um die Proband\*innen während des gesamten Tests zu motivieren. Das heißt konkret, dass die ersten Teilaufgaben leicht zu lösen sein sollten und deren Schwierigkeit dann kontinuierlich zunimmt, um durch komplexere Aufgaben belastbare Aussagen zu erhalten.

## Zusammenfassung und Ausblick

Es gibt verschiedenste Methoden mit denen Bedarfe ermittelt und Nutzende in die Entwicklung von Software einbezogen werden können - je nach Umfang des Produkts und des Anwender\*innenkreises. Usertests erfordern ein anderes Zeitmanagement als die Entwicklung von Personas. Auch der Anwendungsfall nimmt Einfluss auf die Methodenauswahl. So kann für die Entwicklung eines neuen Designs die Verwendung von Wireframes und Mockups bei der Bedarfsermittlung hilfreich sein. Wird ein Portal mit neuen Interaktionsmöglichkeiten entwickelt, empfehlen sich Prototypen, mit denen auch die Interaktionen getestet werden können.

# Sicherheit & Datenschutz

|  |
| --- |
| Wichtig |
| Dieses Kapitel befindet sich in einer ersten öffentliche Entwurfsversion. Feedback siehe Links im rechten Menü und [Hinweise zur Mitarbeit](#mitarbeit). |

|  |
| --- |
| Zusammenfassung |
| Dieses Kapitel gibt einen Überblick über verschiedene Strategien und Maßnahmen für wichtige Aspekte von IT-Sicherheit und Datenschutz und geht auch auf die bibliotheksspezifischen Besonderheiten ein. |

## Einleitung

In den letzten Jahren ist die Zahl der Angriffe auf Bildungseinrichtungen, insbesondere Hochschulen und ihre Bibliotheken, deutlich gestiegen. Laut der Hochschulrektorenkonferenz (HRK) waren bis Januar 2023 insgesamt 24 Hochschulen bzw. Universitäten solchen cyberkriminellen Angriffen ausgesetzt und nahmen Schaden (MDR 2023). Informationseinrichtungen aller Sparten werden zudem mit sicherheits- und datenschutzrelevanten Aufgaben konfrontiert.

Die Entwicklung hin zur Digitalisierung, Automatisierung und Virtualisierung zieht nicht nur die Technisierung eigener Geschäftsgänge und Dienstleistungen nach sich. Diese Entwicklung erfordert auch eine größere Sensibilität und Aufmerksamkeit hinsichtlich der Sicherheit der eigenen Systeme. Safety und Security sind hierfür wichtige, zu unterscheidende Grundprinzipien.

|  |
| --- |
| Wichtig |
| **Security** beinhaltet alle Maßnahmen zum Schutz vor Diebstahl oder Beschädigung von Soft- und Hardware. **Safety** meint den sicherheitsbewussten Umgang mit Netzwerken und Daten (Holländer 2023). |

Bibliotheken verstehen sich als Orte, die ihre Informationen und vielschichtigen Dienstleistungen i.d.R. einer großen Nutzendenschaft zur Verfügung stellen. Sie tragen durch ihre Arbeit zur Umsetzung der Openness-Strategie in Gesellschaft und Wissenschaft bei. Offenheit und Transparenz sind schützenswerte Haltungen von Bibliotheken, welche gleichzeitig aber auch eine größere Angriffsfläche ermöglichen.

Doch was macht Bibliotheken und Hochschulen so interessant für Angriffe? An diesen Orten werden für Cyberkriminelle interessante Daten verwaltet. Dazu gehören personenbezogene Daten, schützenswerte Forschungsdaten, Bewegungsdaten, Lizenzdaten, Daten zur Nutzung von Literatur, etc. Die dabei abgegriffenen Daten können u.a. für Identitätsdiebstahl, Offenlegung privater Informationen oder für die Erstellung eines detaillierten Nutzerprofils herangezogen werden (Holländer 2023).

IT-Sicherheit ist auch immer eine Frage der Zuständigkeit. Je nach Größe, Art und Organisationsstruktur der Bibliothek muss die IT-Sicherheit mit den beteiligten Instanzen wie Stadtverwaltung, Verbundzentrale, den Rechen- und IT-Zentren der Institution und den IT-Sicherheitsbeauftragten und Datenschutzbeauftragten in Bezug auf die Verantwortlichkeiten abgestimmt werden. Grundsätzlich liegt die Aufgabe der Sensibilisierung und Schulung des Bibliothekspersonals im direkten Verantwortungsbereich der Bibliothek.

|  |
| --- |
| Abgrenzung Datenschutzbeauftragte\*r IT-Sicherheitsbeauftrage\*r (Rehm (2023), Schmidt (2023)) |

Wenig ist zum Thema IT-Sicherheit in Bibliotheken veröffentlicht worden (Kost u. a. 2022). Deshalb soll dieses Kapitel zur Vorbeugung nicht nur für das Thema IT-Sicherheit sensibilisieren, eine Beschreibung des Sicherheitsvorfalls und einen Überblick über die Richtlinien und Vorgaben geben, sondern auch für Präventivmaßnahmen werben und Handlungsempfehlungen geben.

## Sicherheitsvorfall

### Einfallstore

Prominente Vorfälle in verschiedenen wissenschaftlichen Einrichtungen haben in den letzten Jahren Schwachstellen in IT-Systemen offenbart, die ernstzunehmende Sicherheitslücken darstellen. Dabei ist zu beachten, dass sowohl Maschinen als auch Menschen verantwortlich für diese Lücken sein können. Folgende sind die häufigsten Angriffsmethoden:

* **Social Engineering:** Unter Social Engineering versteht man Methoden, die zum Vertrauensgewinn eingesetzt werden, um anschließend an sensible Informationen, Zugangsdaten für Systeme oder finanzielle Mittel zu gelangen. Social Engineering-Attacken können über persönliche Kontakte oder über Webdienste wie E-Mails oder Webseiten durchgeführt werden.
* **Phishing:** Hierbei werden seriös wirkende E-Mails versendet, um den Empfänger\*innen Kennwörter oder Zahlungsinformationen zu entlocken, um so an sensible Daten zu gelangen. Die „Qualität“ derartiger gefälschter E-Mails variiert stark. Es zeigt sich jedoch, dass sie kontinuierlich professioneller umgesetzt werden und schwieriger zu identifizieren sind. Häufig werden Phishing-Attacken mit Links auf präparierte Webseiten (sog. Watering Holes) kombiniert. In diesem Fall dienen sie etwa auch als Haupteintrittstore für Ransomware, Malware, Viren und Würmer.
* **Ransomware:** Bei dieser Art des Angriffs wird der Zugriff auf Daten eingeschränkt oder unterbunden, z.B. durch Verschlüsselung. Für die Wiederherstellung des Zugriffs wird ein Lösegeld gefordert. Ransomware hat von allen Angriffsmethoden das größte Schadenspotential. Neben Lösegeldforderungen entstehen hohe Ausfalls- und Wiederherstellungskosten.  
  Als Beispiel: Für die direkten Kosten des Ransomware-Angriffs 2019 auf die JLU Gießen werden mit Stand 2023 ca. 1,7 Mio. EUR für Schadensanalyse und -behebung kalkuliert. Zusätzliche Kosten für Aufwände und Workarounds sind in dieser Analyse nicht berücksichtigt und lassen sich schwer beziffern (Kost u. a. (2022)).
* **Denial-of-Service-Attacken (DoS):** Bezeichnet einen Angriff, um die Verfügbarkeit eines Netzwerkes zu beeinträchtigen und gleichzeitig Daten abzugreifen und umzuleiten.

### Bisherige Erfahrungen

Bisherige Erfahrungen zeigen die weitreichenden Auswirkungen eines cyberkriminellen Angriffs auf Hochschulbibliotheken. In den meisten Fällen ist jedoch nicht nur die Bibliothek alleine betroffen, sondern die gesamte Hochschule. Kommt es zu einem Angriff, ist es meist notwendig, alle IT-Dienste herunterzufahren. Eine Abschottung einzelner Dienste kann kaum vorgenommen werden. Man kann sich dies als einen „harten Cut“ und ein Herunterfahren aller eigenen Server und virtuellen Maschinen zur Eingrenzung des schadhaften Fremdzugriffs vorstellen. Diese harte Maßnahme wird vorgenommen, da man nicht abschätzen kann, wo es im System schon zu welchen Schäden gekommen ist. Die Folgen: es funktioniert schlimmstenfalls NICHTS mehr. (W)LAN, Netzlaufwerke, Identity-Management und Anmeldedienste, E-Mail-Dienste, Terminverwaltungstools, Personal-Verwaltungssysteme, Türschließmechanismen, Zeiterfassung, IP-Telefonie, Lüftungs- und Beleuchtungssysteme, etc. Alle netzbetriebenen Dienste sind ggf. für mehrere Tage, Wochen oder sogar Monate außer Betrieb.

Im Ernstfall können die vorhandenen Dienstgeräte und Kommunikationsdienste nicht mehr verwendet werden. Aufgrund dessen kann man zur Kommunikation auf private Geräte (Notebooks, Smartphones, etc., sofern die Mitarbeiter\*innen bereit dazu sind) und alternative E-Mail-Dienste ausweichen. Hier ist zu bedenken, dass diese Adhoc-Lösungen nicht unbedingt den datenschutzrechtlichen Ansprüchen entsprechen. Von besonderer Bedeutung ist in diesem Fall auch das Identity Management System (IDMS), das nach einem Angriff gegebenenfalls alternativ aufgebaut werden muss.

Hinsichtlich der Bibliotheksnutzung kann das Bibliotheksmanagementsystem (BMS) z.B. auf geprüften Laptops verwendet werden, die ggf. über alternative (mobile) Netzwerke (sofern das BMS wie z.B. im GBV üblich, zentral gehostet wird und nicht ebenfalls vom Angriff betroffen ist), sodass die Ausleihe und Rückgabe wieder zeitnah ermöglicht werden können.

Zu guter Letzt darf auch nicht unterschätzt werden, welche zeitlichen Ausmaße ein Angriff einnehmen kann und welche psychischen und sozialen Auswirkungen er verursacht. Mitunter muss man mit monatelangen Einschränkungen rechnen und man kann auch nicht davon ausgehen, dass Daten überhaupt oder gar vollständig wiederhergestellt werden können.

### Wo bleibt die Künstliche Intelligenz der IT-Sicherheit?

Im Zuge der heutigen digitalen Entwicklungen wirft dieses Thema natürlich auch die Frage auf, welche Auswirkungen KI auf die IT-Sicherheit hat. Nach jetzigen Erfahrungen hat KI bislang noch nicht dazu geführt, dass Systeme anfälliger für Angriffe werden. Denkbar ist, dass Phishing-Angriffe vereinfachen kann. Im Gegenzug könnten cyberkriminelle Angriffe durch Monitoring unserer Systeme mit KI erschwert werden.

## Richtlinien und Vorgaben

Wie in jedem Bereich unseres Lebens finden sich auch in der IT-Sicherheit mehrere Richtlinien und Vorgaben sowohl auf nationaler als auch auf europäischer Ebene.

Auf nationaler Ebene liegt die Zuständigkeit unter anderem beim Bundesamt für Sicherheit in der Informationstechnik ([BSI](https://www.bsi.bund.de/DE/Das-BSI/Auftrag/auftrag_node.html)), welches seit 1991 mit der Aufgabe betraut ist, das Regierungsnetz und die kritischen Infrastrukturen (KRITIS) zu schützen. Auf Bundesebene wurde es eine zentrale Anlaufstelle für Sicherheitsstandards, sowie Meldestelle bei IT-Krisen. Es stellt unterschiedliche Normen zur IT-Sicherung zur Verfügung. Wann welcher Standard greift, hängt von der Komplexität des Einzelfalls ab.

Für Bibliotheken sind hierbei auch die „Checklisten zum [IT-Grundschutz-Kompendium](https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Standards-und-Zertifizierung/IT-Grundschutz/IT-Grundschutz-Kompendium/it-grundschutz-kompendium_node.html)“ nach dem BSI-Standard 200-2 sehr empfehlenswert, da in diesen der aktuelle Stand der Integration des Grundschutzes überprüft werden kann. In diesem Zusammenhang wird auch auf die ISO/IEC-27000-Familie hingewiesen. In dieser Normenreihe sind neben definierten und stets aktuell gehaltenen Standards, die Anforderungen an Information Security Management Systems (ISMS), Empfehlungen für Kontrollmechanismen, als auch Best-Practices-Empfehlungen zu Aufbau und Organisation von Informationsfreiheit enthalten.

Weitere Institutionen, die einen entscheidenden Einfluss auf die Vereinheitlichung des Datenschutzes in der Europäischen Union durch den Erlass der Datenschutz-Grundverordnung ([DSGVO](https://www.bmwk.de/Redaktion/DE/Artikel/Digitale-Welt/europaeische-datenschutzgrundverordnung.html)) von 2016 hatten, sind das Europäische Parlament und der Europäische Rat. Die DSGVO gilt seit 2018 unmittelbar in allen der EU zugehörigen Länder. Sie regelt unter anderem, dass jede Person das Recht auf Schutz der sie betreffenden personenbezogenen Daten hat. So dürfen Daten nur für einen bestimmten Zweck erhoben (Zweckbindung) und auch nicht für andere Vorhaben weiterverarbeitet werden (gem. Art 5 II DSGVO). Ferner wird in Art 37 DSGVO die Ernennung eines Datenschutzbeauftragten geregelt.

Ergänzt wird diese Grundverordnung durch die einzelnen nationalen gesetzgebenden Instanzen. In Deutschland geschieht dies sowohl auf Bundes- als auch auf Landesebene, in Form des Bundesdatenschutzgesetzes (BDSG) sowie der 16 Landesdatenschutzgesetze.

In Österreich wird die DSGVO ergänzt durch das Datenschutzgesetz (DSG) und das Netz- und Informationssystemsicherheitsgesetz (NISG).

Ähnliche Gesetze wurden auch außerhalb der EU erlassen. So trat Anfang September 2023 in der Schweiz die neue Verordnung über Datenschutzzertifizierungen (VDSZ) in Kraft, welche sich an den Grundzügen der DSGVO orientiert.

In ihrer besonderen Rolle als Institutionen für Informationsversorgung und Bereitstellung von Wissensinfrastruktur greifen Bibliotheken auch auf die Dienste externer Anbieter zurück, z.B. durch Verträge mit Wissenschaftsverlagen über digitale Literaturangebote. Hierbei ist es wichtig, dass Bibliotheken in diesen Verträgen darauf bestehen, dass das Tracking des Nutzungsverhaltens der Forschenden ausgeschlossen wird („Tracking in der Wissenschaft: So können Bibliotheken Daten und Wissenschaftsfreiheit schützen“ (2022)), um die Wissenschaftsfreiheit und die informationelle Selbstbestimmung zu schützen (DFG 2021).

Aufgrund der stetigen Weiterentwicklung von Software und neuen Technologien hat die Europäische Kommission einen Vorschlag für ein neues Gesetz zur Widerstandsfähigkeit von Cyberangriffen, den [Cyber Resilience Act](https://digital-strategy.ec.europa.eu/en/policies/cyber-resilience-act), auf den Weg gebracht. Gemäß diesem Vorschlag sollen Verbraucher\*innen und Unternehmen beim Kauf von Produkten und Software mit digitalen Komponenten geschützt werden, indem verbindliche Cyber-Sicherheits-Anforderungen für diese Leistungen durch die Hersteller eingeführt werden sollen. Im nächsten Schritt werden nun das Europäische Parlament und der Europäische Rat über diesen Vorschlag beraten.

## Präventivmaßnahmen

### Passwortsicherheit

Zu einem umfassenden IT-Sicherheitskonzept gehört, dass alle Beteiligten einen bewussten Umgang mit Passwörtern praktizieren. Passwörter sollten zum Beispiel nicht ohne Weiteres für Dritte zugänglich auf Papier oder einem anderen Medium (z. B. unverschlüsselt auf dem Computer) festgehalten werden. Bei Verdacht auf Angriff sollten alle betroffenen Passwörter umgehend geändert werden. Bei einer großen Menge an Passwörtern empfiehlt es sich, einen Passwort-Manager zu verwenden. Für die Wahl des Passworts sollten die allgemein gültigen Empfehlungen beachtet werden (beispielsweise die [Passwort-Empfehlungen des BSI](Tipps%20zur%20Erstellung%20sicherer%20Passwörter:%20https://www.bsi.bund.de/DE/Themen/Verbraucherinnen-und-Verbraucher/Informationen-und-Empfehlungen/Cyber-Sicherheitsempfehlungen/Accountschutz/Sichere-Passwoerter-erstellen/sichere-passwoerter-erstellen_node.html)).

### Authentifizierung und Autorisierung

Unter Authentifizierung verstehen wir grundsätzlich das eindeutige Erkennen eines Zugriffs auf eine Ressource wie z.B. einen Dienst oder einen Computer, aber auch auf ein physisches Objekt wie z.B. einen Drucker.

In der Regel erfolgt heutzutage der Zugriff über ein Netzwerk, d. h., auch die Authentifizierung muss über das Netzwerk erfolgen. Für Geräte (Computer) kann dies durch technische Merkmale wie den Media Access Code (MAC Adresse) oder - flexibler - über die IP-Adresse durchgeführt werden. Diese Merkmale authentifizieren jedoch nur das Gerät und lassen noch keine Aussage darüber zu, wer dieses Gerät gerade benutzt.

Eine Authentifizierung einer Person erfolgt meistens über eine eindeutige Kennung und ein zugehöriges Passwort. Wenn möglich, sollten verschiedene Authentifizierungsverfahren miteinander kombiniert werden. Es wird dann auch von einer 2-Faktor-Authentifizierung (2FA) bzw. einer Mehrfaktor-Authentifizierung (MFA) gesprochen.

Ist ein Zugriff eindeutig durch entsprechende Authentifizierung erkannt, kann auf Basis verschiedener Informationen je nach Bedarf entschieden werden, ob der Zugriff auch berechtigt ist. D. h., der Zugriff muss autorisiert werden. Dies kann sehr individuell erfolgen (Konto A hat Zugriff, Konto B nicht) oder anhand der Zugehörigkeit eines Kontos zu einer bestimmten Personengruppe.

Alle Komponenten, die zur Authentifizierung und Autorisierung in einer Einrichtung notwendig sind, werden auch als Authentifizierungs- und Autorisierungsinfrastruktur (AAI) bezeichnet. Wurde in einer Einrichtung eine AAI aufgebaut, auf welche jede Anwendung zurückgreifen kann, kann man festlegen, dass nach einmaligem Login ein Zugriff auf alle Anwendungen möglich und ein separates Einloggen nicht nötig ist. Dies wird auch als **Single-Sign-On (SSO)** bezeichnet.

|  |
| --- |
| Herangehensweisen für Single-Sign-On |

**Links:** Benutzer\*in meldet sich auf einem **Portal** an und bekommt Zugriff auf alle eingebundenen Dienste. **Mitte:** Benutzer\*in speichert alle Anmeldedaten auf einem Datenträger oder im Netzwerk. Ein **lokal**es Programm meldet ihn\*sie separat bei jedem Dienst, Portal oder Ticketing-System ein. **Rechts:** Benutzer\*in meldet sich bei einem der Dienste an und bekommt ein **Ticket** für den gesamten „Kreis der Vertrauten“.

Zentrale Komponenten sind hierbei der **Identity Provider (IDP)**, der auf Basis des dahinterliegenden IDMS eine digitale Identität inkl. notwendiger Benutzer\*innenattribute bereitstellt und an dem die einmalige Anmeldung stattfindet. Die genutzten Anwendungen werden allgemein als **Service Provider (SP)** bezeichnet.

Eine gute Möglichkeit, das Internet sicher zu nutzen, ist das **Virtual Private Network (VPN)**. Das VPN verschlüsselt die Identität von Internetnutzer\*innen. Damit wird es Dritten erschwert, die Nutzer\*innen im Internet zu verfolgen und Daten abzugreifen. VPN nutzt dabei eine Echtzeitverschlüsselung. Informationen werden in eine unleserliche Form umgewandelt und können nur mit einem Schlüssel wieder in die ursprüngliche leserliche Form gebracht werden.

### Updates und Backups

Eine sichere Hard- und Software ist unabdingbar. Rechner sollten ausreichend geschützt sein. Dazu gehören unter anderem aktuelle Virenscanner und automatische Updates des Virenscanners, des Betriebssystems und jeder auf dem Rechner laufenden Software. Fehlende Updates können als Einfallstore genutzt werden, um Schadsoftware auf den Computer und damit auch in das Netzwerk einzuschleusen.

Falls es zu einem Angriff oder Ausfall des Systems kommt, muss auf Backups zurückgegriffen werden. Dabei wird eine Sicherungskopie angelegt, meist auf einem anderen Medium, wie z. B. einer externen Festplatte. Server werden mittels eines RAID-Verfahrens gespiegelt, d. h. die Daten werden auf einen zweiten Server übertragen. Da es bei einem Backup auch zu Datenverlust oder -beschädigung kommen kann, empfiehlt es sich, mehr als eine Kopie anzulegen, z. B. auf Magnetbändern.

|  |
| --- |
| Wichtig |
| Ein **Server** ist ein zentraler Rechner (virtuell oder physikalisch), der Daten zur Verfügung hält und diese auf Anfrage durch einen Client (PC, mobiles Gerät) zur Verfügung stellt. |

### Schulungen

Regelmäßige Weiterbildungen ermöglichen es den Mitarbeiter\*innen von Bibliotheken, ihre Fähigkeiten auszubauen und die Sensibilität für Themen wie IT-Sicherheit zu erhöhen. Die im Abschnitt Richtlinien und Vorgaben [interne Verlinkung] erwähnten Rahmenbedingungen schaffen dabei ein Grundgerüst für Schulungsinhalte und können für verpflichtende IT-Sicherheitsschulungen - wie sie bereits von einigen Bibliotheken durchgeführt werden - verwendet werden.

Es sollten zwei Arten von Schulungen durchgeführt werden: für Mitarbeiter\*innen und Nutzer\*innen. Mitarbeiter\*innen sollten dahingehend sensibilisiert werden, dass sie mit schützenswerten Daten arbeiten. Es ist bspw. wichtig, daran zu denken, dass gerade Geräte in Büros im Sichtfeld von Personen durch sinnhafte Positionierung oder Blickschutzfilter vor unbefugtem Sehen geschützt werden. Auf die Wichtigkeit sicherer Passwörter wurde bereits weiter oben in diesem Abschnitt hingewiesen. Es sollte zudem auf die Angriffsmöglichkeiten im Abschnitt „Sicherheitsvorfall“ [interne Verlinkung] aufmerksam gemacht werden. Beispielhafte Erklärungen möglicher Angriffsszenarien und Möglichkeiten (z. B. Tools) sich zu schützen, sollten aufgezeigt werden.

## Handlungsempfehlungen

Sind öffentliche Einrichtungen (Hochschulen, Kommunen) von Cyberattacken betroffen, hat dies allumfassende Folgen. Im schlimmsten Fall sind sämtliche Dienste nicht mehr verfügbar und die Ausfallzeit und Schadenshöhe weisen ins Ungewisse. Dann können Notfall- und Sicherheitskonzepte teils detaillierte Anleitungen für die Ersthilfe bieten. Generell gilt, dass Cyberattacken nicht nur eine kommunikative, sondern häufig auch eine Organisationskrise nach sich ziehen.

|  |
| --- |
| Übersicht über mögliche Gefahrenquellen und Maßnahmen zur Prävention und im Fall eines Schadenfalles |

Zum einen muss intern (mit vorbereiteten Notfallgeräten oder ggf. privaten Endgeräten) und extern (mit Dienstleister\*innen, Bibliothekspartner\*innen, Bibliotheksnutzer\*innen) kommuniziert werden u.a. über den Vorfall, und ggf. weiterlaufende Dienste (cloudbasierte und von Drittanbieter\*innen). Besonders cloudbasierte Dienste können im Notfall evt. gute Alternativlösungen bieten, das gilt zum einen für alternative Loginverfahren z.B. bei extern gehosteten Plattformen für E-Ressourcen, wenn das IDMS nicht zur Verfügung steht.

Ist das BMS einer Bibliothek cloudbasiert, kann ein Ausleihbetrieb mit Notfallgeräten schneller wieder aufgenommen werden, da die Daten vor dem Hacker besser geschützt sind. Ähnlich kann man mit eigenen Dokumenten vorgehen, in dem das Active Directory des eigenen Systems ein tägliches Update erstellt und z.B. im CSV-Format an eine Cloud exportiert. Im einfachsten Fall betrifft es nur Notfalldokumente, die einem Krisenstab zugänglich sind. Tritt ein Notfall ein, kann man über die Standard-Benutzeroberfläche der Cloud die gesicherten Skripte starten. Der Cloudbetreiber sollte Erfahrungen mit Notfall-Lösungen und Umgebungen haben. Trotzdem sollte die Bibliothek ihre Datensicherung gut planen und cloudbasierte oder andere netzwerkunabhängige Backups länger vorhalten (mindestens 14 Tage), um auf nicht infizierte Sicherungen zurückgreifen zu können. Es ist ja nicht auszuschließen, dass die Angreifer die Zugangsdaten für die Datensicherung abgreifen konnten. Daher sollte auch über eine Verschlüsselung der Datensicherung in Erwägung gezogen werden. Genauso gilt es, die Wiederherstellung der Daten zu planen und zu üben. Unerlässlich ist auch eine gute Dokumentation der gesamten Notfall-Architektur.

Und zum anderen sind alternative Abläufe aufzubauen, um wieder an der digitalen Arbeitswelt teilzunehmen und den Bibliotheksbetrieb weitestgehend wieder anbieten zu können.

Die kommunikativen Anforderungen sind wie die organisatorischen als sehr hoch einzuschätzen, da Informationseinrichtungen mit ihren analogen Begegnungsmöglichkeiten Schnittstellen für eine ganze Hochschule sein können. Für beide Aufgaben können andere Bibliotheken und externe Firmen Unterstützung bieten.

Stehen Plan und eine Notfallumgebung, ist es unerlässlich, regelmäßige Notfallübungen durchzuführen. Damit trainiert man die Abläufe und kann diese regelmäßig überprüfen und optimieren (Müller 2023). Die dadurch entstehende Routine kann Sicherheit für den Ernstfall geben. Dennoch darf man nicht unterschätzen, dass ein Angriff auch persönliche Auswirkungen auf alle Betroffenen hat.

## Fazit

Lässt man die Informationen in diesem Kapitel noch einmal Revue passieren, bleibt als abschließendes Fazit frei nach Ringelnatz nur zu sagen: „Sicher ist, dass nichts sicher ist. Selbst das nicht“. Trotzdem ist man der Gefahr nicht hilflos ausgeliefert. Durch Präventivmaßnahmen, Notfall-Systeme und Pläne kann man einen Angriff wohl nicht verhindern. Man kann jedoch versuchen, den Schaden und die daraus folgenden Belastungen so gering wie möglich zu halten. In vielen Bibliotheken entstehen derzeit individuelle Sicherheitskonzepte und neue Sicherheitsmaßnahmen. Die nächsten Monate und Jahre werden zeigen, mit welchen neuen Strategien sich Bibliotheken vor Cyberangriffen schützen und sich auf diese vorbereiten können.

# Daten & Metadaten

|  |
| --- |
| Zusammenfassung |
| Die Beschreibung von Ressourcen mit standardisierten Metadaten bildet eine zentrale Voraussetzung für die meisten an Bibliotheken angebotenen IT-Dienste. Das vorliegende Kapitel gibt einen Überblick über die Thematik, indem es [grundlegende Begriffe](#grundlegende-begrifflichkeiten) einführt, wichtige [Metadatenstandards](#metadatenstandards) vorstellt und [Prozesse der Datenverarbeitung](#Xc0d2b380f537523ab05a0d07c2c628226492086) in Bibliotheken erläutert. |

## Einleitung

Zur Sammlung und Bereitstellung von Informationen werden von Bibliotheken Ressourcen unterschiedlichster Form (Bücher, Filme, Forschungsdaten, …) nachgewiesen. Zur Verwaltung dieser Ressourcen werden diese mit **Metadaten** beschrieben. Neben diesen Metadaten beinhalten bibliothekarische Informationssysteme zunehmend auch Dokumente selbst als **digitale Inhalte** wie sogenannte Volltexte, Digitalisate und Forschungsdaten (siehe Kapitel [Digitalisierung](#digitalisierung) und [Forschungsnahe Dienste](#forschungsnahe-dienste)). Viele der im Folgenden beschriebenen Grundlagen zu Eigenschaften, Arten und Verarbeitung von Daten gelten sowohl für Metadaten als auch für digitale Inhalte.

## Grundlegende Begrifflichkeiten

### Daten

Im Wesentlichen bestehen Daten im Sinne dieses Buchs aus einer Folge von Bits. Abgesehen von ihrer Anzahl in Bytes lässt sich auf dieser Ebene allerdings nichts weiter über Daten sagen. Uns interessiert daher mehr, für was die Daten stehen – beispielsweise für eine Jahreszahl, ein Bild oder für den Titel eines Dokumentes. Dabei besteht ein Unterschied zwischen

* der Struktur von Daten (**Syntax**)
* und der Bedeutung von Daten (**Semantik**).

Zur Interpretation von Daten dienen **Kodierungen** in Form von Datenformaten und Identifikatoren. Wo genau jeweils die Grenze zwischen Syntax und Semantik liegt, hängt davon ab, auf welcher Ebene und mit welcher Kodierung Daten betrachtet und verarbeitet werden (siehe [Tabelle 5.1](#tbl-daten-ebenen)):

Tabelle 5.1: Beispiele für Syntax und Semantik von Daten auf verschiedenen Ebenen

| Datenebene | Bedeutung | Kodierung |
| --- | --- | --- |
| 1100001 | Die Zahl 97 (64+32+1) | Byte als Zahl |
| 97 | Der Buchstabe „a“ | ASCII oder Unicode |
| a | Unterfeld für Haupttitel | Feld 021A im PICA+ K10plus-Format |
| a | Unterfeld für Umfangsangabe | Feld 300 im MARC21 Format |

Ein Großteil der [Datenverarbeitung](#datenverarbeitung) besteht darin, Daten, zum Beispiel im Rahmen von [ETL-Prozessen](#etl-prozess), von einer Kodierung in eine andere Kodierung zu überführen, um sie anschließend leichter interpretieren zu können. Bei der Konvertierung von Daten von einem in ein anderes Datenformat reduziert sich das implizite Datenmodell der Konvertierung auf den kleinsten gemeinsamen Nenner beider Formate.

Zur Beschreibung von Daten dienen

* formale **Schemata** auf Ebene der Syntax
* und **Datenmodelle** auf Ebene der Semantik.

Leider liegen beide oft nicht explizit vor, sondern müssen anhand von Beispielen, Anwendungen und Dokumentation mühsam ermittelt werden. Im Idealfall entsprechen Daten einem klar definierten Datenformat.

### Datenformate

Datenformate definieren eine Struktur, die sich in einer oder in mehreren austauschbaren Syntax-Varianten ausdrücken lässt und deren Bedeutung durch ein Datenmodell festgelegt ist. Beispielsweise definiert der Unicode-Standard eine Menge von Schriftzeichen (Buchstaben, Sonderzeichen, Emojis, …) und verschiedene Verfahren, um Zeichenketten in Bytes zu kodieren (UTF-8, UTF-16, …). Syntax-Varianten werden auch als **Serialisierung** bezeichnet. Die meisten Datenformate haben nur eine Serialisierung, so dass Format und Syntax meist synonym verwendet werden. Einzelne Syntax-Elemente entsprechen Bestandteilen im Datenmodell (siehe [Tabelle 5.2](#tbl-xml)), daher werden in der Beschreibung von Daten auch diese beiden Ebenen meist nicht sauber getrennt.

Tabelle 5.2: Einige Bestandteile des XML-Formats

| XML-Syntax | XML-Modell |
| --- | --- |
| <name /> oder <name>...</name> | XML-Element |
| name="Inhalt" | XML-Attribut |
| <!-- ... --> | Kommentar |

Als Faustregel kann gelten, dass bei statischer Betrachtung von Daten der Bezug auf ihre Syntax sinnvoll ist, während zur Verarbeitung von Daten eher auf die Bestandteile ihrer Modelle Bezug genommen werden sollte.

Datenformate lassen sich grob in zwei Kategorien unterteilen:

* **Strukturierungssprachen** wie CSV, XML, JSON und RDF ermöglichen es, Daten in kleinere Einheiten zu unterteilen und miteinander in Beziehung zu setzen. Die Sprachen basieren auf allgemeinen Ordnungsprinzipien (Felder, Tabellen, Hierarchien, Netzwerke, …) und ihre Modelle haben darüber hinaus keine eigene Semantik. Die einfachste Strukturierungssprache ist das Prinzip der Zeichenkette.
* **Anwendungsformate** legen die Struktur von Daten für konkrete Arten von Inhalten fest ([Metadatenformate](#metadatenformate) zur Beschreibung von Dokumenten, Bildformate für Bilder…). Ihre Modelle verweisen letztendlich auf reale Objekte und Eigenschaften. Viele Anwendungsformate sind ihrerseits mittels einer Strukturierungssprache kodiert, zum Beispiel basiert das DataCite-Format zur Beschreibung von Forschungsdaten auf dem XML-Modell.

Darüber hinaus gibt es besondere Formate, deren Anwendung in der Verarbeitung von Daten liegt. Neben Programmiersprachen, die selbst nicht als Datenformate betrachtet werden, sind dies folgende Sprachen:

* **Schemasprachen** wie XML Schema, JSON Schema und nicht zuletzt *reguläre Ausdrücke* dienen der formalen Beschreibung der Syntax von Datenformaten. Dabei bezieht sich jede Schemasprache auf eine Strukturierungssprache (*XML Schema* für XML-Formate, Avram für feldbasierte Formate, …).
* **Abfragesprachen** dienen dem Verweis auf einzelne Teile von Datensätzen. Sie beziehen sich ebenfalls immer auf eine Strukturierungssprache (zum Beispiel XPath für XML, JSON Path für JSON, …) und sind für die Verarbeitung von Daten notwendig.
* **Modellierungssprachen** helfen bei der Beschreibung von Datenmodellen. Die häufigsten Modellierungssprachen basieren auf dem Entity-Relationship-Modell. Da zwischen Syntax und Semantik irgendwann die reine Datenebene verlassen werden muss, sind die wichtigsten Mittel zur Datenmodellierung allerdings Diagramme und Beschreibungen in natürlicher Sprache.

Die Verwendung von Schema-, Abfrage- und Modellierungssprachen hilft, viele häufige Fehler bei der (Meta-)Datenverarbeitung zu vermeiden. Ein Beispiel hierfür ist das *Resource Description Framework* (*RDF*) mit dazugehörigen Schemasprachen (*SHACL*/*ShEx*), Abfragesprachen (*SPARQL*) und Modellierungssprachen (*RDFS*/*OWL*). In anderen Fällen wird aus Mangel an Werkzeugen und Kenntnissen auf spezielle Datensprachen verzichtet und stattdessen auf allgemeine Programmiersprachen zurückgegriffen.

|  |
| --- |
| Info |
| **Reguläre Ausdrücke** sind das gängigste Mittel zur Beschreibung der Syntax von Daten. Gleichzeitig können mit ihnen Zeichenketten nach Mustern durchsucht werden. Ein regulärer Ausdruck für die Syntax einer ISBN-13 mit optionalen Trennstrichen ist beispielsweise:  (97[89])-?([0-9]{1,5})-?([0-9]+)-?([0-9]+)-?[0-9]  Üblicherweise decken Schemasprachen nicht alle Aspekte eines Datenformats ab. So lässt sich die Korrektheit der abschließenden Prüfziffer ([0-9]) nicht mit einem regulären Ausdruck überprüfen. |

Eine ausführlichere Beschreibung von Datenformaten mit bibliothekarischem Schwerpunkt bietet die Datenbank [format.gbv.de](https://format.gbv.de/). Grundlagen von Metadaten und Ontologien vermitteln Assfalg (2023) und Rölke und Weichselbraun (2023).

In der Praxis werden Daten in einem Datenformat zusätzlich durch anwendungsspezifische Auslegungen und Einschränkungen geprägt, darunter Format-Varianten, Metadatenprofile bzw. Anwendungsprofile, Erfassungsregeln und die jeweilige Erfassungspraxis.

### Identifikatoren

Ein wesentlicher Teil von Daten besteht aus Identifikatoren (IDs wie Nummern, Codes…) zum Verweis auf externe Objekte oder an anderer Stelle verwaltete Informationen. Identifikatoren ermöglichen die eindeutige Referenzierung gleicher Dinge in unterschiedlichen Kontexten, so dass Daten aus verschiedenen Quellen miteinander abgeglichen und kombiniert werden können.

Neben eher intern genutzten Datensatz-Identifikatoren (z.B. die *PPN* des Bibliothekssystems PICA oder die ZDB-ID der Zeitschriftendatenbank) sind vor allem international standardisierte Identifikatoren von Bedeutung. Beispiele für solche Identifier-Systeme mit Relevanz für bibliothekarischer Daten sind die nachfolgenden:

* Die *International Standard Book Number* (**ISBN**) wird von Verlagen für Bücher vergeben. Seit 2007 ist die 13-stellige ISBN Teil des EAN-Barcode-Systems.
* Die *International Standard Serial Number* (**ISSN**) identifiziert Zeitschriften und Schriftenreihen.
* Der *Digital Object Identifier* (**DOI**) identifiziert digitale Publikationen in elektronischen Zeitschriften und Repositorien.
* Der *International Standard Identifier for Libraries and Related Organisations* (**ISIL**) referenziert Bibliotheken, Archive, Museen und verwandte Einrichtungen.
* Die *Open Researcher and Contributor ID* (**ORCID**) identifiziert Autor\*innen von wissenschaftlichen Publikationen.
* Der *Uniform Resource Locator* (**URL**) dient als Adresse einer digitalen Ressource im Web und wird teilweise gleichzeitig als deren Identifikator eingesetzt.
* Das System der *Uniform Resource Identifier* (**URI**) ermöglicht die Vereinigung verschiedener Identifier-Systeme und bildet die Grundlage von *RDF* und Linked Open Data (*LOD*).

Gemeinsam ist den Identifikatoren, dass sie jeweils eine definierte Syntax haben (z.B. XXXX-XXXY im Falle der ISSN, wobei X für eine Ziffer und Y für eine Prüfziffer steht), deren Bestandteile hierarchisch von einer zentralen Instanz festgelegt werden. Nach dem Prinzip des Namensraums kann dabei die Vergabe von Teilen an untergeordnete Organisationen delegiert werden. Beispielsweise werden ISIL für Bibliotheken in Deutschland beginnend mit dem Präfix „DE-“ durch die [ISIL-Agentur an der Staatsbibliothek zu Berlin](https://sigel.staatsbibliothek-berlin.de/) verwaltet.

Völlig dezentrale Identifikatoren gibt es zur Identifizierung von digitalen Objekten nur in Form von Prüfsummen, die sich automatisch aus vorhandenen Daten berechnen lassen (SHA-Summe, IPFS-Adresse, Prüfziffer…).

### Normdaten

Einfache kontrollierte Vokabulare bestehen aus normierten Listen von eindeutigen Benennungen – beispielsweise könnte in einem Gemüse-Vokabular festgelegt sein, dass immer „Karotte“ statt „Möhre“ verwendet werden muss. Wird jeder Eintrag mit einem künstlichen Identifikator versehen, muss die Benennung selbst nicht eindeutig sein. Existiert eine Datenbank zum Nachschlagen dieser IDs, so wird diese auch als *Normdatei* bezeichnet. Ihre Datensätze dienen als *Normdaten* der eindeutigen Identifizierung von Personen, Organisationen, geographischen oder administrativen Einheiten, Themen oder anderen Entitäten an verschiedenen Stellen. Ein Normdatensatz besteht mindestens aus einem Identifikator und einer Vorzugsbenennung als primärer Name. Oft gibt es weitere identifizierende Merkmale wie alternative Benennungen, Lebensdaten von Personen, Ortsangaben u.Ä. sowie Verknüpfungen zwischen verschiedenen Entitäten.

Umfang und Komplexität von Normdateien reichen von einfachen Listen bis zu komplexen [Ontologien und Wissensgraphen](#datenmodelle-und-rdf-formate). Ein prominentes bibliothekarisches Beispiel einer Normdatei ist die Gemeinsame Normdatei (*GND*), in der neben Personen auch Körperschaften, Veranstaltungen, Geografika, Werke und Sachschlagwörter miteinander vernetzt sind.

Zur Anreicherung von Daten mit Normdaten mittels *Entity Recognition* (Erkennung von Entitäten in Daten) und *Entity Linking* (Abgleich von Entitäten mit IDs anderer Normdateien) sollte unterschieden werden zwischen:

* Normdateien mit Entitäten wie Personen (ORCID), Publikationen (DOI, ISBN), Sprachen (ISO 639) etc., die sich grundsätzlich eindeutig unterscheiden lassen sowie
* Normdateien, deren abstrakte Entitäten von Kontext und Modellierung abhängen (Klassifikationen, Thesauri, …).

Zur Verwaltung von Normdaten gibt es einige Datenformate wie *MARC 21 for Authority Data* und *ISAAR (CPF)*. Als gemeinsamer Nenner auch außerhalb des Bibliotheksbereichs gilt das RDF-basierte *Simple Knowledge Organization System* (*SKOS*), für das mit JSKOS auch eine JSON-Variante existiert. Vor allem kleinere oder speziellere Normdateien (z.B. Codelisten im Rahmen der Katalogisierung) liegen allerdings selten in maschinenlesbarer Form vor.

Das [*Basic Register of Thesauri, Ontologies & Classifications*](https://bartoc.org/) (BARTOC) erfasst Informationen zu Normdateien aller Art, darunter auch Verfahren zum technischen Zugriff.

## Metadatenstandards

Neben allgemeinen [Datenformaten](#datenformate) sind für die Bibliotheks-IT vor allem Metadatenformate zur Beschreibung von Dokumenten relevant. Die meisten der im Folgenden beschriebenen Metadatenformate spielen außerhalb von Kultureinrichtungen keine wesentliche Rolle. Für digitale Objekte (*METS/MODS*, *LIDO*, *CDWA*, *EN 15907*, *EAD*, …, siehe Kapitel [Digitalisierung](#digitalisierung)) und für Forschungsdaten (DataCite, siehe Kapitel [Forschungsnahe Dienste](#forschungsnahe-dienste)) gibt es darüber hinaus spezielle Formate.

### Arten von Metadaten

Folgende Arten von Metadaten können nach ihrer Funktion unterschieden werden:

* **Deskriptive** (=beschreibende) **Metadaten** zur Identifizierung und inhaltlichen Beschreibung wie Titel, Verfasser\*in, Schlagwörter etc.
* **Administrative** (=Verwaltungs-) **Metadaten** wie Angaben zu Herkunft, Speicherung, Zugriffsrechten, Verwaltung etc.
* **Strukturelle Metadaten** über den Aufbau von Dokumenten wie die Einteilung in einzelne Kapitel, Abschnitte, eingebundene Medien etc. Die Grenze zwischen digitalen Inhalten und ihrer Struktur ist allerdings mitunter fließend.
* **Technische Metadaten** zu Merkmalen wie Umfang, verwendeten Datenformaten etc.

Je nach Anwendung gibt es spezielle Metadatenformate oder es werden verschiedene Arten von Beschreibungen in einem Format zusammengefasst.

### Feldbasierte bibliothekarische Metadatenformate

**Machine-Readable Cataloging (MARC)** ist das älteste und noch immer wichtigste Format für den Austausch von Daten zwischen Bibliotheken. Die aktuell relevante Variante ist MARC 21, insbesondere das Format MARC 21 für bibliographische Daten. Neben der binären Kodierung kann MARC 21 auch in XML und JSON kodiert werden. Viele Eigenheiten und Probleme des Formats sind historisch bedingt, eine Alternative konnte sich bislang nicht durchsetzen.

**PICA** ist das von MARC inspirierte Datenformat der Katalogisierungssysteme CBS und LBS (Voß 2022). Das wichtigste Anwendungsprofil ist das K10plus-Format des BSZ/GBV.

**MAB** und **allegro** sind ebenfalls an MARC angelehnte, feldbasierte Formate aus dem deutschsprachigen Raum, die allerdings nur noch sporadisch verwendet werden.

### XML-basierte Datenformate

* **METS** und **MODS** sind zwei zusammen im Bereich [Digitalisierung](#digitalisierung) eingesetzte Formate für strukturelle und administrative (METS) sowie bibliografische Metadaten (MODS). Strukturdaten in METS ermöglichen granulare Gliederung und Verlinkung von Objekten wobei mögliche Typen und Beziehungen in Regelsätzen definiert sind.
* [*Encoded Archival Description*](https://wiki.deutsche-digitale-bibliothek.de/pages/viewpage.action?pageId=19010182) (**EAD**) ist der zentrale dokumentarische XML-Standard zur Beschreibung von archivischen Findmitteln.
* **LIDO** ist ein etabliertes Austauschformat für den Museumsbereich.
* **DataCite** ist ein bibliographisches Datenformat insbesondere zur Beschreibung von Forschungsdaten (siehe Kapitel [Forschungsnahe Dienste](#forschungsnahe-dienste)).

### Datenmodelle und RDF-Formate

|  |
| --- |
| Definition |
| Eine **Ontologie** ist ein Datenmodell, das verschiedene Klassen und Eigenschaften in RDF definiert und so die einheitliche Kodierung und Verknüpfung verschiedener Datenquellen als Linked Data bis zu umfangreichen Wissensgraphen ermöglicht.  Im Gegensatz zu einfacheren Formen von [Normdaten](#normdaten) geht es bei Ontologien nicht nur um die eindeutige Identifizierung (Beispiel: ist mit „Bank“ das Gleiche wie „Sitzbank“ oder ein Geldinstitut gemeint?) sondern auch um Eigenschaften und Beziehungen (Beispiel: mögliche Größe, Material und Ort von Bänken). **Wissensgraphen** enthalten neben Ontologien auch konkrete Daten über Instanzen der Ontologie-Klassen (Beispiel: Liste konkreter Sitzbänke an einem Ort). |

**Dublin Core** bzw. das Dublin Core Metadata Element Set (**DCMES**) hat als kleinster gemeinsamer Nenner der meisten Metadatenstandards die größte Verbreitung. Es besteht aus 15 Basiselementen wie „creator“, „title“, „date“ und „description“ und Erweiterungen mit den DCMI Metadata Terms wie „Alternative Title“, „Date Created“ und „Date Available“.

Die *Functional Requirements for Bibliographic Records* (**FRBR**) sind ein sehr abstraktes Metadatenmodell. Sie beinhalten insbesondere eine Einteilung von bibliographischen Entitäten in die Beschreibungsebenen „work“, „expression“, „manifestation“ und „item“.

Die **BIBFRAME**-Ontologie wurde entwickelt, um MARC auf Grundlage von RDF zu ersetzen. Die wesentlichen Elemente sind „work“, „instance“ und „item“ sowie damit verbundene Eigenschaften und Entitätstypen (siehe [Abbildung 5.1](#fig-bibframe)).

**Schema.org** ist eine allgemeine Ontologie für strukturierte Daten in Webseiten.

|  |
| --- |
| Abbildung 5.1: Hauptbestandteile des Datenmodell BIBFRAME |

### Verlagsdaten und Literaturangaben

Die Formate *ONIX*, *JATS*, *BITS* und *CrossRef XML* stammen aus dem Verlagsbereich zur Beschreibung von Zeitschriftenartikeln und Büchern. Sie basieren alle auf XML und sind für Bibliotheken für den Datenimport relevant. Datenformate für Literaturangaben (*BibTeX*, *RIS*, *Endnote*, *CSL-JSON*…) werden dagegen zum Export von Katalogdaten bereitgestellt. Zitationsregeln für Literaturangaben und Ansetzungsregeln von *ISBD* sind dagegen für den Datenaustausch eher unbrauchbar. Learning Object Metadata (*LOM*) dient in verschiedenen lokalen Anpassungen der Beschreibung von Lerneinheiten.

## Datenverarbeitungsprozess in Bibliotheken

### Datenerfassung

Traditionell werden bibliothekarische Metadaten durch Katalogisierung erstellt. Die Verwaltung der Katalogdaten erfolgt entweder lokal oder gemeinsam in einer **Verbunddatenbank**. Der Vorteil der Verbundkatalogisierung liegt darin, dass jedes Dokument nur einmal zentral beschrieben werden muss, während bei lokaler Katalogisierung durch *Fremddatenübernahme* nur zum Teil auf vorhandene Kataloge zurückgegriffen werden kann.

Im Idealfall sollte die Erfassung nach *Autopsie*, also auf Grundlage des vorliegenden Werkes, durch geschultes Personal und nach etablierten Regelwerken (*Katalogisierungsrichtlinien*) erfolgen. Um möglichst viele Publikationen zu erfassen, wird jedoch zunehmend auch auf anderweitig erfasste Metadaten von Verlagen, Repositorien und aus anderen Quellen zurückgegriffen. Dazu müssen Daten unterschiedlicher Erschließungstiefe und -qualität im Rahmen von [ETL-Prozessen](#etl-prozess) gesammelt, analysiert und mit vorhandenen Daten vereinheitlicht werden. In jedem Fall muss beachtet werden, dass sich Regeln und Umstände, nach denen Daten erfasst werden, mit der Zeit ändern können (beispielsweise der Umstieg der Erfassungsregeln von *RAK* auf *RDA*) und dass das Ergebnis auch davon abhängt, wie gut überprüft werden kann, was die Anforderungen an die Daten sind.

Darüber hinaus gibt es Verfahren zur automatischen Erstellung von Metadaten aus vorhandenen Dokumenten, beispielsweise zur Erkennung und Auswertung von Literaturangaben und zur thematischen Einordnung von Dokumenten. Mit diesen Verfahren lassen sich zwar größere Mengen von Daten erfassen, es muss aber immer mit einer gewissen Fehlerrate gerechnet werden.

Welche Art und welcher Umfang von Fehlern und Uneinheitlichkeiten bei der Datenerfassung tolerierbar sind, hängt letztlich davon ab, wozu die Daten erfasst werden. So gelten beispielsweise für eine historische Bibliographie andere Maßstäbe als für einen Suchindex.

Nicht zuletzt sollte bedacht werden, dass Geschwindigkeit und Qualität von Datenerfassung auch von der Usability der Werkzeuge abhängen, mit denen Daten erstellt, bearbeitet und *analysiert* werden können.

|  |
| --- |
| Info |
| Mehr zur bibliothekarischen Datenerfassung in den *Grundlagen der Informationswissenschaft* (2023), *Teil B*. |

### ETL-Prozess

Da sich die IT in- und außerhalb von Bibliotheken über verschiedene Organisationen und Systeme erstreckt, müssen an vielen Stellen Daten von einer oder mehreren Quellen in ein anderes Informationssystem übertragen werden. Der grundsätzliche Prozess der Datenintegration, der Quell- und Zielsysteme verbindet, wird als **ETL-Prozess** bezeichnet. Der Prozess aus drei zentralen Schritten „Extract“, „Transform“ und „Load“ stammt ursprünglich aus dem Bereich des Data Warehousing und findet sich auch in anderen Anwendungsfällen. Im Folgenden wird er am Beispiel der Integration von Metadaten in ein Discovery-System beschrieben. [Abbildung 5.2](#fig-etl) illustriert den generellen ETL-Prozess mit einigen exemplarischen Arbeitsschritten.

|  |
| --- |
| Abbildung 5.2: Beispiel eines ETL-Prozess |

#### Extraktion

Ziel der **Extraktion** (Extract) ist die Auswahl und der Abzug relevanter Daten aus verschiedenen Datenquellen. Hierbei handelt es sich primär um einen technischen Vorgang, das sogenannte **Harvesting**, welcher automatisiert oder manuell gestartet werden kann. Der Aufwand und die Qualität des Harvesting können je nach Datenquelle sehr unterschiedlich ausfallen. Denkbare Datenquellen sind Dateien, Datenbanken bzw. Datenbankabzüge, [Schnittstellen](#schnittstellen) oder eher unstrukturierte Quellen wie Websites, die zunächst mittels Screenscraping erschlossen werden müssen.

Der Extraktionsvorgang erfolgt bei Bedarf regelmäßig, um die Daten im Zielsystem aktuell zu halten. Mögliche Aktualisierungsintervalle sind:

* **periodisch**, das heißt in zeitlich regelmäßigen Abständen unabhängig von der jeweiligen Aktualisierung der Daten in den Quellsystemen
* **ereignisgesteuert**, immer wenn bestimmte Bedingungen wie zum Beispiel die Änderung von Daten in den Quellsystemen eintreten
* **manuell**, beispielsweise wenn Daten aus Quellsystemen ad hoc importiert werden sollen. Manuelle Aktualisierungen bieten sich vor allem an, wenn der Inhalt der Datenquellen weitestgehend statisch ist, da die dauerhaft verlässliche Aktualisierung im Zielsystem ohne Automatismus nicht garantiert ist.

Der Extraktionsvorgang ist technisch relativ einfach handhabbar, wenn strukturierte [Datenformate](#datenformate) und/oder [Schnittstellen](#schnittstellen) existieren — die wesentlichen Aufwände finden sich dann im nachfolgenden Transformationsschritt. Anders sieht es aus, wenn beispielsweise Daten manuell eingesammelt werden müssen oder Screenscraping notwendig ist. Beim *Screenscraping* müssen aufwändig Extraktionsskripte erstellt werden, um Daten aus Webseiten in ein strukturiertes Format zu überführen. Diese Skripte sind zudem sehr fehleranfällig und müssen jedes Mal angepasst werden, wenn die Betreiber\*innen der Datenquelle Veränderungen vornehmen.

Die extrahierten Daten werden in einem sogenannten Arbeitsbereich abgelegt und dort im nächsten Prozessschritt aufbereitet.

#### Transformation

Daten aus verschiedenen Quellsystemen liegen zumeist in unterschiedlichen Formaten mit unterschiedlichen Datenmodellen vor. Neben Unterschieden in der Syntax können gleiche Sachverhalte auch auf semantischer Ebene unterschiedlich beschrieben sein, da die Daten mitunter für abweichende Anwendungsfälle erfasst wurden. So müssen beispielsweise in einem Discovery-System Metadaten zur einfachen Beschreibung so aufbereitet werden, dass sie auch erweiterte Suchstrategien unterstützen.

Ziel der **Transformation** ist es, alle Daten in ein einheitliches Format mit gemeinsamen Datenmodell zu überführen. Dieses Zielformat wird beim ETL-Prozess auch als *Schema* bezeichnet. Die Vereinheitlichung des Schemas (*Mapping*) ist ein wesentlicher Schritt jeder Datenkonvertierung. Zur Minimierung des Transformationsaufwands dienen gemeinsame Standards wie *MARC21* als Austauschformat oder die einheitliche Verwendung von RDF-Ontologien.

Über die einfache Konvertierung hinaus sind im Rahmen der Transformation oft weitere Aufbereitungen zur Vereinheitlichung und Verbesserung der Datenqualität notwendig. Beispiele hierfür sind:

* Prüfung und **Vereinheitlichung** der Zeichencodierung auf normalisierten Unicode
* formale **Anpassungen** von Daten wie die Vereinheitlichung von Datumsformaten, Ländercodierungen etc.
* Erkennung und Eliminierung von **Duplikaten**
* Abgleich, Vereinheitlichung und **Konsistenzprüfung** von Aussagen über dieselben Objekte aus verschiedenen Datenquellen
* **Anreicherung** oder Korrektur von Datensätzen mittels Zusatzinformationen aus Normdaten oder anderer zusätzlicher Datenquellen

Der Aufwand der Transformation sollte nicht unterschätzt werden, da die Datenqualität verschiedener Datenquellen stark variieren kann und Qualitätsprobleme oft erst spät entdeckt werden. Zudem können sich Datenquellen und ihre Qualität zwischen Aktualisierungen ändern. Sofern die Datenübernahme nicht nur einmalig stattfinden soll (*Konversion*), ist die Betreuung des Transformationsschrittes eine Daueraufgabe.

Da im Transformationsschritt regelmäßig Massendaten analysiert und modifiziert werden müssen, ist der Einsatz von IT-gestützten [Werkzeugen](#werkzeuge) und Verfahren der [Datenanalyse](#datenanalyse) unerlässlich. Die aus der Analyse gewonnenen Erkenntnisse müssen wiederum kontinuierlich in die Anpassung des Schema-Mappings einfließen, damit der Transformationsprozess nicht ins Stocken gerät.

#### Laden

Auf die Transformation folgt beim **Laden** (load) die Überführung der vereinheitlichten Daten in das Zielsystem – beispielsweise in den Suchindex eines Discovery-Systems. Dabei dürfen nur Datensätze in Produktivsysteme übernommen werden, die den Transformationsschritt erfolgreich durchlaufen haben, während für Test- und Entwicklungssysteme andere Regeln möglich sind.

Das Laden selbst ist ein technisch beherrschbarer Schritt, welcher optimalerweise darauf abzielt, das Zielsystem ohne Ausfallzeiten aktuell zu halten. Bei Änderungen des Schemas muss deshalb besonders darauf geachtet werden, gleichzeitig entsprechende Anpassungen im Zielsystem vorzunehmen. Daneben ist es bei absehbaren Änderungen ratsam, das Schema und die Verarbeitung von Daten im Zielsystem von Vornherein flexibel zu gestalten.

#### Umsetzung des ETL-Prozess

Die Umsetzung von Datenkonvertierung und ETL-Prozessen erfolgt im bibliothekarischen Umfeld oft über selbst entwickelte Skripte für das Harvesting und die Transformation. Das entsprechende Know-How ist in der Regel auf wenige Köpfe verteilt und kann auch nicht leicht durch das Hinzuziehen externer Expertise verfügbar gemacht werden.

Es gibt einige kommerzielle ETL-Komplettlösungen mit Data-Warehouse- oder Business-Intelligence-Hintergrund. Angesichts der teils erheblichen Einstiegskosten und zur Vermeidung von *vendor-lock-in* sind für Bibliotheken möglichst einfache und allgemeine [Werkzeuge zur Datenverarbeitung](#werkzeuge) jedoch meist die bessere Wahl. Die Vorteile etablierter ETL-Werkzeuge liegen in Schulungsmöglichkeiten und der Verfügbarkeit externer Expertise. Mit *Catmandu*, *Metafacture* und *OpenRefine* gibt es mehrere Open Source- ETL-Frameworks, deren eingeschränkter Funktionsumfang und verbesserungswürdige Usability durch Anpassungen für bibliothekarische Datenformate und Schnittstellen möglicherweise aufgewogen werden.

Grundsätzlich lassen sich die kontinuierlich anfallenden Aufwände der Transformation und Qualitätssicherung durch ein ETL-Werkzeug nicht vermeiden, sondern nur besser handhabbar machen. Dazu sollten die einzelnen Arbeitsschritte zentral verwaltet und dokumentiert werden, beispielsweise durch ein Versionskontrollsystem. Durch den Fokus auf Usability kann mittels ETL-Werkzeugen bibliothekarisches Personal stärker eingebunden werden. Da hier insbesondere im Bereich der Programmierung die Einstiegshürden niedriger sind, ist es Fachpersonal leichter möglich, direkt Änderungen und Optimierungen am ETL-Prozess vorzunehmen. Bei einer reinen Verankerung von ETL im IT-Bereich sind solche Eingriffe dagegen nur in Zusammenarbeit von IT und Metadaten-Expert\*innen umsetzbar.

In jedem Fall gehen mit der Einführung von ETL-Werkzeugen in die bibliothekarische Arbeit immer auch individuelle Anpassungen im Prozess von Extraktion, Transformation und Laden einher. Dieser Aufwand kann sowohl gegen die Einführung solcher Werkzeuge sprechen als auch dafür, vorhandene „Bastellösungen“ zu evaluieren und zu konsolidieren.

### Werkzeuge

Für die Verarbeitung von Daten im Rahmen bibliothekarischer IT-Systeme werden grundsätzlich die gleichen Werkzeuge verwendet wie außerhalb des Bibliothekswesens, daher wird an dieser Stelle auf eine allgemeine Einführung in die Datenverarbeitung verzichtet. Ganz allgemein sind hier als Werkzeuge

* Mittel zur **Dateiverwaltung** und ein **Texteditor** unabdingbar,
* allgemeine **Kommandozeilenprogramme** (curl, sort, grep…) sehr zu empfehlen
* und **Programmiersprachen** vor allem für komplexere Aufgaben hilfreich.

Material und Kurse für die praktischen Grundkenntnisse mit Bibliotheksbezug finden sich insbesondere in den Bereichen *Data Librarianship* und *Data Science*.

Werkzeuge für konkrete Datenformate orientieren sich an den zugrunde liegenden [Strukturierungssprachen](#daten). So gibt es beispielsweise eigene Editoren oder Editor-Plugins für XML-Daten und JSON-Daten und entsprechende Kommandozeilentools wie xmlstarlet für XML und jq für JSON. Für tabellarische Daten eignet sich etwa eine Tabellenkalkulation oder das tabellenbasierte Werkzeug *OpenRefine* ([openrefine.org](https://openrefine.org/)).

Für bibliothekarische Datenformate und Schnittstellen gibt es darüber hinaus einige speziellere Werkzeuge:

* Programmierbibliotheken wie *MARC4J* und *YAZ* erleichtern die Datenverarbeitung im Rahmen eigener Programme.
* Anwendungsprogramme wie *WinIBW* und *BibControl* erfordern zwar weniger IT-Kenntnisse, sind dafür aber nur eingeschränkt und/oder nur für sehr spezielle Aufgaben verfügbar.
* Freie Werkzeuge zum Metadatenmanagement wie die Frameworks *Catmandu* ([librecat.org/Catmandu](http://librecat.org/Catmandu/)) und *Metafacture* ([metafacture.org](https://metafacture.org/)).

Für einzelne Anwendungen und Formate gibt es einige weitere Werkzeuge und es kann sich lohnen, solche Werkzeuge selbst zu entwickeln und als Open Source zur Verfügung zu stellen. Für das PICA-Format sind solche Programme in der *Einführung in die Verarbeitung von PICA-Daten* (Voß 2022) aufgeführt.

### Schnittstellen

Eine API (Application Programming Interface, auch Programmierschnittstelle) ist eine definierte Methode zur Abfrage und/oder Änderung von Daten in einem Informationssystem. Wie die Daten innerhalb des Systems verwaltet werden, ist dabei nebensächlich. Dieses Prinzip ermöglicht die Kombination unterschiedlicher Softwarekomponenten. Wenn möglich, sollten produktunabhängige, offen dokumentierte APIs verwendet werden. Im Bibliotheksbereich sind insbesondere folgende APIs relevant:

* **Z39.50** wurde vor Erfindung des Web zur Suche in Bibliotheksdatenbanken entwickelt. Nachfolger ist das XML-basierte *Search/Retrieve via URL* (**SRU**) mit der zugehörigen Abfragesprache *Contextual Query Language* (**CQL**).
* Das [*Open Archives Initiative Protocol for Metadata Harvesting*](https://www.openarchives.org/pmh/) (**OAI-PMH**) dient dem Abruf von Metadaten aus Repositorien. Die Daten können nach Datum und Teilmengen gefiltert und so in Suchmaschinen und Portalen wie [BASE](https://www.base-search.net/) und der [Deutschen Digitale Bibliothek](https://www.deutsche-digitale-bibliothek.de/) (DDB) zusammengeführt werden.
* Das *NISO Circulation Interchange Protocol* (**NCIP**), das *Simple Library Network Protocol* (**SLNP**) und das *Standard Interchange Protocol* (**SIP2**) sind interne APIs für Ausleihe und Fernleihe. Sie werden zwischen Bibliotheken, Fernleihservern und zur Anbindung von Verbuchungsautomaten eingesetzt (Michaelis 2014).
* Die [*Patrons Account Information API*](https://gbv.github.io/paia/) (**PAIA**) ist eine offene Schnittstelle zum Zugriff auf Ausleihkonten.
* Die [*Document Availability Information API*](https://gbv.github.io/daia/) (**DAIA**) ist eine offene Schnittstelle zur Abfrage der Verfügbarkeit von Medien.
* Mit **unAPI** können einzelne Datensätze in verschiedenen Formaten abgerufen werden.
* Die [**Reconciliation Service API**](https://reconciliation-api.github.io/specs/draft/) ermöglicht den Abgleich mit Normdaten zur eindeutigen Referenzierung (siehe Abschnitt [Identifikatoren und Normdaten](#identifikatoren-und-normdaten)).
* Die APIs des [*International Image Interoperability Framework*](https://iiif.io/) (**IIIF**) ermöglicht die Referenzierung und Nutzung digitalisierter Werke in externen Werkzeugen durch gezielte Verlinkung auf einzelne Bestandteile.
* Über **SPARQL**-Schnittstellen können RDF-Daten aus Wissensgraphen wie zum Beispiel Wikidata abgerufen werden.
* Verschiedene Schnittstellen zur Authentifizierung und Autorisierung wie *LDAP*, *Shibboleth* und *OAuth*.

Darüber hinaus bieten die meisten Anwendungen eigene, meist interne Schnittstellen, zum Beispiel die Solr-API der Suchplattform *Apache Solr*. Besonders im Bereich [Forschungsnaher Dienste](#forschungsnahe-dienste) gibt es weitere, spezialisierte Schnittstellen.

### Datenanalyse

Im Gegensatz zu physischen Objekten ist Daten ihre Beschaffenheit nicht direkt anzusehen. Lediglich der Umfang von Daten in Bytes und ggf. die Anzahl von Dateien und Datensätzen kann einen ersten Anhaltspunkt liefern. Weitere Einschätzungen, insbesondere darüber, ob Daten vollständig oder fehlerhaft sind, setzen eine konkrete Analyse der Daten voraus. Dies beinhaltet auch die Visualisierung von Daten zur Exploration, Kommunikation und Diskussion (Jetter 2023).

Die Auswertung von Daten ist nicht nur für das Qualitätsmanagement relevant (Voß 2021), beispielsweise um im Rahmen des ETL-Prozesses Verteilungen und Ausreißer zu erkennen, sondern auch um aus Daten weitere Erkenntnisse zu gewinnen. So können beispielsweise Ausleihzahlen nach Medien gruppiert für die Bestandsplanung eingesetzt werden. Voraussetzung dafür ist, dass Daten überhaupt vorliegen, ebenso wie Mittel und Kenntnisse zu ihrer Auswertung. Bei fehlenden oder zu umfangreichen Daten können Stichproben erhoben werden, wobei auf die Zufälligkeit der Stichprobe und das Konfidenzintervall des Ergebnisses geachtet werden muss.

Neben rudimentären Statistik-Kenntnissen helfen bei der Datenanalyse Werkzeuge wie die [im vorigen Abschnitt](#werkzeuge) beschriebenen Mittel zur Datenverarbeitung. Für allgemeine Analysen eignet sich vor allem eine Tabellenkalkulation. Für komplexere Analysen gibt es spezielle Statistik-Programme und Programmiersprachen wie *SAS* und *R*. Für explorative Analysen und um bei geänderter Datenlage automatisch aktuelle Ergebnisse zu bekommen, bieten sich interaktive Umgebungen wie *Jupyter Notebooks* oder *Observable* an. Für spezielle, wiederkehrende Analysen und Aufbereitungen kann es auch sinnvoll sein, eigene Anwendungen zu entwickeln bzw. entwickeln zu lassen. Beispiele hierfür sind *BibControl*, das *Metadata Quality Assessment Framework* (*MQA*), die Deutsche Bibliotheksstatistik sowie Statistikfunktionen als Teil anderer Programme (zum Beispiel [Statistik und Reporting](#statistik-und-reporting) als Teil des BMS).

## Künstliche Intelligenz

Unter den Begriffen Künstlichen Intelligenz, Machine Learning und Deep Learning werden verschiedene Verfahren der Datenverarbeitung zusammengefasst, die Aspekte menschlicher Intelligenz imitieren können. Auf Grundlage statistischer Verfahren lassen sich so Aufgaben automatisieren, für die bisher Personal notwendig gewesen wäre. Der Umfang der verarbeitbaren Daten wird lediglich durch die zur Verfügung stehende Rechenleistung begrenzt. Die Qualität der Ergebnisse hängt stark von der jeweiligen Aufgabenstellung und den eingesetzen Verfahren ab. Im besten Fall lassen sich die Ergebnisse praktisch nicht mehr von menschlich erzeugten Daten unterscheiden. Beim Einsatz von KI-Technologien sind allerdings auch ethische Aspekte zu beachten.

Für die bibliothekarische Datenverarbeitung lassen sich grob zwei Arten von KI-Anwendungen unterscheiden:

* Verfahren zur Analyse und Anreicherung von Daten, beispielsweise die automatische Erschließung und Musterkennung im Rahmen der [Digitalisierung](#digitalisierung).
* Systeme die Antworten, Texte und Medien erzeugen, umschreiben und zusammenfassen von einfachen Chatbots bis zu umfangreichen Sprachmodellen wie ChatGPT.

Die zunehmende Verfügbarkeit von leistungsfähigen KI-Systemen wird in absehbarer Zeit zu Änderungen in der Rezeption und Produktion von Medien führen und damit auch Auswirkungen auf die Arbeit von Bibliotheken haben.

## Zusammenfassung & Ausblick

Strukturierte Metadaten sind unverzichtbar für die Verwaltung und den Zugriff auf Ressourcen in Bibliotheken. Daher bilden sie und ihre Verarbeitung die Grundlage für praktisch alle von Bibliotheken angebotene IT-Dienste. Während Aufwand und Bedeutung von Datenverarbeitung auch in Zukunft hoch bleiben wird, ist davon auszugehen, dass der Einsatz semantischer Technologien (RDF) zur Zusammenführung heterogener Daten und von Verfahren der künstlichen Intelligenz zunehmen wird.

# Bibliotheksmanagementsysteme

|  |
| --- |
| Zusammenfassung |
| Ein Bibliotheksmanagementsystem ist ein IT-System, das die [Kernprozesse](#prozessabbildung) einer Bibliothek unterstützt. Das modular aufgebaute System verfügt über verschiedene [Komponenten](#komponenten) für die jeweiligen Kernprozesse Erwerbung, Katalogisierung/Erschließung und Ausleihe sowie für die Recherche. Die [Entwicklungsgeschichte](#geschichte) der Managementsysteme in Bibliotheken erstreckt sich mittlerweile über drei Generationen. Relevant für die Arbeit mit den Managementsystemen ist auch die [Anbindung an Verbundkataloge](#verbundkataloge) sowie die [Integration anderer Systeme](#integration). |

## Einleitung

Das BMS spielt eine zentrale Rolle für die meisten klassischen Geschäftsprozesse in Bibliotheken.

|  |
| --- |
| Definition |
| Ein **Bibliotheksmanagementsystem** ist ein Softwareprodukt, mit dem die Arbeitsprozesse rund um die Erwerbung, Erschließung, Ausleihe, den Zugriff und die Auffindbarmachung von Bibliotheksbeständen über Kataloge abgebildet und automatisiert werden können. |

Durch die Ausweitung der Aufgaben in den Bereichen Publikationsdienste, Open Science oder auch Lernort sind in neuerer Zeit jedoch noch weitere Aufgaben hinzugekommen, die durch die klassischen BMS nicht abgebildet werden. Darüber hinaus haben die frühen Systeme nur sehr unzureichende Möglichkeiten, die nötigen Informationen zu elektronischen Ressourcen und ihrer Zugänglichkeit abzubilden. Auch zur Unterstützung von neueren Aufgaben wie der Publikationsunterstützung oder der Verwaltung räumlicher Ressourcen werden separate Systeme genutzt. Daraus ergibt sich der Bedarf, das BMS an diese separaten Systeme anzubinden, was die Bedeutung von Schnittstellen und offenen Architekturen erhöht hat.

In diesem Text wird der Begriff Bibliotheksmanagementsystem verwendet. Teilweise wird im Deutschen auch der allgemeinere Begriff *Bibliothekssystem* verwendet. In der angloamerikanischen Literatur finden sich die Begriffe *Integrated Library System (ILS)* und *Library Management System (LMS)*, zuletzt aber auch *Library Services Platforms*.

## Geschichte der Bibliotheksmanagementsysteme

Ihren Ursprung haben Bibliotheksmanagementsysteme in den 1960er Jahren, als Bibliotheken damit begannen, Katalogdaten untereinander auszutauschen und auf diese Weise Prozesse zu optimieren. In dieser Zeit entwickelten sich auch die heute noch gebräuchlichen Austauschformate für Katalogdaten, z.B. MARC.

In den 1970er Jahren erlaubte die fortschreitende technische Entwicklung die Automatisierung weiterer Prozesse über den Datenaustausch hinaus. Zunächst war dies vor allem die Ausleihe mit der Verbuchung von Medien und der Erzeugung von Mahnschreiben. Auch die Verwaltung von Bestellungen im Rahmen der Erwerbung wurde möglich, so dass man in der Folge von *Integrated Library Systems* zu sprechen begann. Davon, dass Katalog-, Erwerbungs- und Nutzer\*innen-Daten an einem Ort gehalten und bearbeitet wurden, versprach man sich eine größere Effizienz der Arbeitsprozesse. Diese **erste Generation** von BMS beinhaltete teilweise auch schon digitale Funktionen für Bibliotheksnutzer\*innen wie über Telnet erreichbare Kataloge, die von Anfang an als integraler Bestandteil der BMS gesehen wurden (Borgman 1997).

Die Entstehung des World Wide Web in den 1990er Jahren hatte zunächst vor allem Einfluss auf die Benutzbarkeit der Kataloge, die Web-Oberflächen erhielten. Aber auch die anderen Komponenten der BMS wurden überarbeitet, und zwar zunehmend auch von kommerziellen Anbietern, während die ersten Systeme als Eigenentwicklungen von Bibliotheken entstanden. Die Landschaft an **Systemen der 2. Generation** war von den späten 1990er bis in die Nullerjahre sehr divers, ist zuletzt aber von vielen Übernahmen geprägt worden, so dass man von einem konsolidierten Markt sprechen kann (Breeding o. J.).

|  |
| --- |
| Info |
| Die **erste Generation** der Bibliotheksmanagementsysteme umfasste Grundfunktionen für die Ausleihe wie Verbuchung und Mahnung, für die Erwerbung die Verwaltung von Bestellungen und teilweise auch über Telnet erreichbare Kataloge für die Bibliotheksnutzer\*innen. Die **zweite BMS-Generation** verfügte über erweiterte Funktionalitäten zur Unterstützung der Kernprozesse sowie durch Weboberflächen der Kataloge aus. Die **dritte Generation** zeichnete sich durch stärkere Modularisierung und mehr Schnittstellen zur Anbindung weiterer Systeme aus. |

Seit den 2010er-Jahren vermarkten Anbieter eine neue Generation von BMS, die **Next-Generation Library Management Systems**, die auch **Library Services Platforms** genannt werden. Diese zeichnen sich durch verschiedene technische und funktionale Neuerungen aus. Die Datenhaltung erfolgt in der Regel cloudbasiert (auch wenn dies bei bei älteren Systemen im Grunde auch schon der Fall war), außerdem werden in der Regel mehr Schnittstellen zur Integration des Systems mit anderen Lösungen angeboten. Funktional wurden die Systeme vor allem um die Möglichkeit der Verwaltung von elektronischen Ressourcen erweitert sowie Statistik- und Reporting-Funktionalitäten verbessert.

Seit dem Ende der 1990er Jahre spielen auch wieder Lösungen eine Rolle, die nicht kommerziell sind. Diese Open Source-Lösungen haben in der Regel eine große Anwender-Community und lassen einen vielfältigen Markt für Support- und Wartungsdienstleistungen zu.

|  |
| --- |
| Abbildung 6.1: Evolution der Bibliotheksmanagementsysteme (nach Matthews und Block (2020), S. 7) |

|  |
| --- |
| Info |
| Nach Matthews und Block (2020) lässt sich die Geschichte der BMS in sechs überlappende Epochen einteilen (siehe [Abbildung 6.1](#fig-bmsevol)):   1. **System-Epoche**: Erste Schritte in den 1950er bis in die 1970er-Jahre hin zur Entwicklung von Software, z. T sehr experimentell, die die klassischen Geschäftsgänge von Bibliotheken in einem digitalen System abbilden sollen – dadurch prägt sich der Begriff „Bibliothekssystem“. Das Augenmerk bei der Entwicklung liegt besonders auf Nachbildungen des Leihverkehrs unter besonderer Beachtung der Identifikation überfälliger Medien. 2. **Epoche der Funktionalität**: Kommerzielle Bibliothekssoftware-Anbieter beginnen sich zu formieren, die erstmals eine integrierte Lösung der verschiedenen Automationsbereiche (Erwerbung, Katalogisierung, Zeitschriftenakzession, Verbuchung, Leihverkehr usw.) anbieten. Hierdurch entsteht die Bezeichnung „Integriertes Bibliothekssystem“ (IBS), der auf den aus dem US-amerikanischen Raum übernommenen Begriff „Integrated Library System“ (ILS) zurückgeht. In den 1980ern entstehen die ersten Online-Kataloge (OPAC), die die in Bibliotheken traditionellen Zettelkataloge nachbilden. 3. **Nutzer\*innen-Fokus-Epoche**: Durch die Erkenntnis, dass sich die Gewohnheiten von Bibliotheksnutzer\*innen im Zugang zu und Umgang mit Medien u.a. mit dem Aufkommen des WWW in ihrem Alltag zunehmend ändern (z.B. durch die Nutzung von Online-Shopping und Suchmaschinen), rücken die Bedürfnisse der Nutzer\*innen immer mehr in den Fokus bei der Entwicklung von Bibliothekssystemen. 4. **Epoche der Verbreiterung der Informationsressourcen**: Der Übergang in eine Phase, bei der Medien nicht mehr erworben, sondern digital lizenziert werden. Entsprechend entwickelt sich das Bedürfnis nach einem Electronic Resource Management (ERM) und neuartige BMS unterfüttern zum Ende der Epoche diesen Wandel mit einer von Medientypen unabhängigen Ressourcenverwaltung. 5. **Discovery-System-Epoche**: Systeme, die über den lokalen Medienbestand hinaus auch extern lizenzierte Inhalte über eine alleinige Suchplattform zugänglich machen, erfreuen sich zunehmender Beliebtheit bei den BMS-Betreiber\*innen. Sie sollen den Nutzer\*innen einen deutlichen Mehrwert bieten. Seit den 2010er Jahren sind es überwiegend kommerzielle Verlage, die umfangreiche e-Medien-Pakete oder Indizes von Volltext- und Bibliografie-Datenbank als lizenzierbare Resource Discovery Services anbieten. 6. **Wissensinnovation**: Bibliotheken realisieren überwiegend, dass Discovery-Systeme nicht ihre gewünschte Wirkung entfalten und sie sich deutlicher von Plattformen großer Tech-Unternehmen abgrenzen müssen. Wissen soll neu erschlossen werden mit innovativen Technologien wie 3D-Druck, Virtual Reality (VR), Open-Access-Repositorien etc. |

Tabelle 6.1: relevante Software-Produkte (Stand Mitte 2022)

|  | Organisation | Marktstatus | Open Source | Individuelle Entwicklung |
| --- | --- | --- | --- | --- |
| aDIS/BMS | aStec | ÖBs und WBs vor allem im BSZ | nein | durch aStec |
| Alma | ExLibris | WBs in Berlin, NRW, Bayerische Staatsbibliothek, Schweiz | nein | durch ExLibris, integrierte Apps in Eigenregie |
| FOLIO | Open Library Foundation | Einführung in WBs | ja | in Eigenregie oder durch Dienstleister |
| Koha | Koha Community | ÖBs und Spezialbibliotheken, in Planung im KOBV | ja | in Eigenregie oder durch Dienstleister |
| LBS | OCLC | WBs im GBV und Spezialbibliotheken | nein | durch VZG |
| LIBERO | LIBERO/Knosys | ÖBs und WBs | nein | durch LIBERO |

In [Tabelle 6.1](#tbl-aktuelle-sw-produkte) sind die aktuellen BMS mit der derzeit größten Marktreife und -durchdringung im deutschsprachigen Raum (Stand Mitte 2022, Sortierung nach Namen) angegeben. Weitere BMS wie ExLibris Aleph, SISIS Sunrise und allegro werden zwar auch noch an vielen Bibliotheken eingesetzt, aber nicht mehr wesentlich weiterentwickelt. Das Cloud-basierte System WMS von OCLC ist in Deutschland bislang nur vereinzelt im Einsatz. Für BibliothecaPlus ist von OCLC ein Nachfolger angekündigt.

Darüber hinaus gibt es mehrere kommerzielle Systeme, deren Funktionsumfang auf bestimmte Arten von Bibliotheken zugeschnitten ist, beispielsweise:

* [Perpustakaan](https://must.de/default.html?Lib.htm) ist in Schulbibliotheken verbreitet und wendet sich auch an nicht-bibliothekarisch vorgebildetes Personal,
* [NOS](https://de.wikipedia.org/wiki/NOS_(Bibliothekssoftware)) ist in internen Forschungs- und Behörden-Bibliotheken verbreitet,
* Quria von [Axiell](https://www.axiell.com/) ist in skandinavischen ÖBs verbreitet und löst im deutschsprachigen Raum das BMS BIBDIA ab.

Eine umfangreiche internationale Übersicht von BMS enthält der von Marshall Breeding gepflegte [Library Technology Guide](https://librarytechnology.org/products/). Für den Deutschsprachigen Raum gibt es Übersichten von Verbundzentralen oder Büchereifachstellen, z.B. Kluge (2022) für öffentliche Bibliotheken. Darüber hinaus sind Daten zu BMS systematisch in Wikidata erfasst und können beispielsweise unter <https://w.wiki/574K> abgefragt werden.

|  |
| --- |
| Info |
| Der IT-Lebenszyklus von BMS ist mit mehr als 20 Jahren im Vergleich zu anderen IT-Systemen eher lang. So wurde beispielsweise FOLIO im Rahmen des *Open Library Environment Project* bereits 2009 initiiert und wird wahrscheinlich erst im nächsten Jahrzehnt in die Wartungsphase übergehen. |

## Funktionalitäten von Bibliotheksmanagementsystemen

BMS sind in der Regel modular aufgebaut und verfügen mindestens über Module für folgende Funktionen:

* Erwerbung
* Katalogisierung / Erschließung
* Ausleihe
* ein Recherche-Modul, das sich vorwiegend an die Bibliotheksnutzer\*innen richtet

### Grundlegende Komponenten

Die Systeme der **1. und 2. Generation** können als sehr ausgereift bezeichnet werden und lassen vielfältige Möglichkeiten zu, bibliothekarische Geschäftsgänge in einem hohen Detaillierungsgrad abzubilden. Nachfolgend werden diese entsprechenden Aufgabenbereiche skizziert.

**Erwerbung** meint die Beschaffung benötigter Bestände bei Verlagen. Darunter fallen z.B. folgende Aufgabengebiete:

* Bestellungen
* Budgetverwaltung
* Rechnungsverwaltung
* Lieferantenverwaltung
* Zeitschriften- und Fortsetzungsabonnements
* Unterstützung [EDIFACT](https://unece.org/trade/uncefact/introducing-unedifact)-Standards
* Buchbinder

*Siehe auch* [*Prozessabbildung: Erwerbung*](#erwerbung-prozess)

**Katalogisierung** meint die Erschließung der verwalteten Medien und digitalen Quellen, z.B. anhand

* Übernahme von Fremddaten
* Anbindung an Verbünde
* Integration digitalisierter Medien

*Siehe auch* [*Prozessabbildung: Katalogisierung*](#katalogisierung-prozess)

**Ausleihe** meint vorwiegend die Verwaltung physischer Medien bzw. Objekte und regelt die Interaktionen mit Nutzer\*innen wie z.B.:

* Abbildung komplexer Reglements nach Benutzer- und Medientypen, Standorten usw (siehe auch [Benutzungsbedingungen](#benutzungsbedingungen))
* Ausleihfristen
* Verwaltung von Standorten
* Versand von Benachrichtigungen
* Anbindung an Selbstverbuchungslösungen
* Mahngebühren

*Siehe auch* [*Prozessabbildung: Ausleihe*](#ausleihe-prozess)

Das **Recherchemodul** stellt die Sicht für die Nutzer\*innen auf Bestände der Einrichtung zur Recherche und Kontofunktionen dar:

* Katalog (auch OPAC genannt)
* Benutzerkonto

*Siehe auch* [*Prozessabbildung: Katalog*](#katalog-prozess)

Die **Next Generation-Systeme** zeichnen sich gegenüber den Systemen der 1. und 2. Generation in der Regel durch andere Systemarchitekturen aus. Das heißt, sie verfügen über aktuellere technische Einzelkomponenten und Schnittstellen, auf deren Grundlage auch zahlreiche zusätzliche Funktionalitäten angeboten werden können. Im Einzelnen gibt es folgende Merkmale, die ein **Next Generation-System** kennzeichnen (Schweitzer 2016):

* Angebot als Software as a Service ([SaaS](#cloud))
* Mandantenfähigkeit
* Interoperabilität durch offene, standardisierte und dokumentierte Schnittstellen
* Verfügbarkeit von Datenbanken bzw. Knowledge Bases für bibliografische Daten und Lizenzinformationen
* oftmals kein fest integrierter Katalog, sondern Schnittstellen zu Discovery-Systemen
* Statistik-Werkzeuge
* Erzeugung von Semesterapparaten
* Anlegen von Favoriten-Listen
* Anzeige von Buchcovern

### Vergleich mit anderen Managementsystemen

Aufgrund der hohen Kosten für die Einführung oder die Migration eines BMS dürfte sich für viele Entscheider\*innen die Frage stellen, ob sich die Investition lohnt bzw. ob sich die Aufgaben auch mit anderen Lösungen erledigen lassen (siehe hierzu [Beschaffung und Marktanalyse](#beschaffung-und-marktanalyse)).

Systeme zur Automatisierung von Geschäftsprozessen gibt es in verschiedenen Branchen. Eine genauere Betrachtung der Aufgaben, die durch Automatisierung unterstützt werden sollen, kann aufzeigen, ob dafür ein Bibliotheksmanagementsystem oder eine andere Lösung besser geeignet ist.

Die folgenden Alternativen sind möglicherweise für kleine Einrichtungen relevant, die über sehr überschaubare Bestände verfügen und kaum oder wenig ausleihen:

Erfassung von Medien:

* Listen in einer Tabellenkalkulation (Excel, LibreOffice, …)

Erfassung und Web-Präsentation von Medien:

* [Library Thing for Libraries](https://www.librarything.com/forlibraries)
* [Zotero Groups](https://www.zotero.org/groups/)
* Stand-Alone-Lösungen für Electronic Resource Management wie [Coral](http://coral-erm.org)

Ausleihe

* Plugins für Wordpress wie [WebLibrarian](https://www.greengeeks.com/tutorials/create-a-library-management-system-in-wordpress/)

Erwerbung

* Finanzbuchhaltungssysteme wie SAP, [HIS-Hochschul-ERP](https://www.his.de/)

Nutzerdatenverwaltung

* IDM-Systeme

Bibliotheken mit einem jährlichen Zuwachs von über 500 Medien und verschiedenen Benutzertypen und Ausleihbedingungen ist die Nutzung eines BMS zu empfehlen, da hier eine gewisse Prozesseffizienz einerseits und eine Erschließungs- und Dienstleistungsqualität andererseits erreicht werden kann.

|  |
| --- |
| Info |
| Als gedankliches Experiment ist die Überlegung, auf ein BMS zu verzichten, jedoch gut geeignet, um sich über die Anforderungen klar zu werden. Insbesondere die Rolle des Bibliothekskataloges als Schnittstelle zu den Bibliotheksnutzer\*innen kann und sollte kritisch hinterfragt werden. Beispielsweise gab es [Überlegungen der Universitätsbibliothek in Utrecht](https://insights.uksg.org/articles/10.1629/2048-7754.174/), auf dieses klassische Instrument gänzlich zu verzichten. |

### Integration des BMS mit anderen IT-Systemen

Innerhalb der Bibliothek werden BMS meist zusammen mit anderen Softwaresystemen eingesetzt. Insbesondere sind dies:

* [Selbstbedienungsautomaten](#automatisierung-und-selbstbedienung) (Ausleihe, Rücknahme, Sortierung von Medien, Bezahlung von Gebühren)
* Dokumentenserver, Content Management Systeme und andere Repositorien
* Workflowsysteme (Digitalisierung von Altbestand; Publikationsunterstützung, …)

Weitere Systeme müssen für eine effektive Arbeit sinnvoll mit dem BMS verbunden werden:

* Haushaltssysteme wie SAP, HIS Haushalt-ERP
* [Identitätsmanagementsysteme](#identity-management) (Account-Verwaltung)
* Lieferantensysteme (bibliographische Daten, Bestell- und Rechnungsdaten)

Im bibliothekarischen Umfeld sind folgende Systeme relevant:

* der [Verbundkatalog](#verbundkataloge)
* die [Zeitschriftendatenbank](#Xef3cdf849515896b72e4da871885dcc60b5fbed)
* die [elektronische Zeitschriftendatenbank](#Xef3cdf849515896b72e4da871885dcc60b5fbed)

Für die regionale und überregionale Literaturversorgung (physische, Print-Medien, E-Medien) spielt die Anbindung an folgende Systeme eine wesentliche Rolle

* Fernleihe
* Dokumentenlieferdienste (wie [Subito](https://www.subito-doc.de/) und Fachinformations-Lieferdienste)

Die Anbindung an die entsprechenden Dienste (Zentraler Fernleih-Server, Fernleihdienst, Subito-Server etc.) ist für viele, aber durchaus nicht alle Bibliotheken relevant.

Im Zusammenhang mit dem Aufbau der Fachinformationsdienste für die spezialisierte Informationsversorgung in Deutschland werden in zunehmendem Maße Fachportale entwickelt. Relevante Katalog-Informationen werden aus möglichst vielen Bibliotheken regelmäßig abgerufen (Harvesting), in ein einheitliches Datenformat übertragen und anschließend als gemeinsamer Index für die übergreifende Recherche in Discovery-Systemen angeboten. Die BMS müssen entsprechend über [Standardschnittstellen](#datenformate-und-schnittstellen) die relevanten Katalogdaten in einem vereinbarten [Datenformat](#bibliographische-metadaten) bereitstellen.

### Verbundkataloge

In Deutschland haben sich Katalogverbünde in den 1970er und 1980er Jahren entwickelt. Zunächst haben sich die wissenschaftlichen Bibliotheken meistens auf Bundesland-Ebene für die Rationalisierung der Katalogisierung zu Verbünden zusammengeschlossen. Inzwischen sind in diesen Verbünden auch öffentliche Bibliotheken vertreten. Darüber hinaus gibt es mit WorldCat einen internationalen Verbundkatalog. Die Anbindung an WorldCat geschieht jedoch in der Regel nicht direkt über das lokale BMS sondern über den Bibliotheksverbund.

[Tabelle 6.2](#tbl-verbund) gibt eine Übersicht über die deutschsprachigen Bibliotheksverbünde.

Tabelle 6.2: Bibliotheksverbünde und -kataloge

| Verbund | Verbundkatalog | System |
| --- | --- | --- |
| BVB | B3Kat | ALEPH (Ex Libris) |
| BSZ | K10plus | CBS (OCLC) |
| GBV | K10plus | CBS (OCLC) |
| hebis | hebis | CBS (OCLC) |
| hbz | hbz | Aleph (Ex Libris)  Alma-Netzwerkzone (Ex Libris) |
| KOBV | B3Kat | Aleph (Ex Libris) |
| VÖBB (öffentliche Bibliotheken) | VÖBB | aDIS/BMS (aStec) |
| Österreichischer Bibliothekenverbund | OBV | Alma-Netzwerkzone (Ex Libris) |
| Swiss Library Service Platform (SLSP) | swisscovery | Alma |

### Anbindung an Verbundkataloge/Verbundkatalogisierung

Die Übernahme von bibliografische Daten oder - bei elektronischen Medien - Paket -bzw. Lizenzinformationen aus anderen Systemen ist für eine Bibliothek unabhängig davon, ob sie in einem Verbund organisiert ist, von Interesse. Eine Anbindung von bibliografischen Datenquellen, z.B. per Z39.50, für die Übernahme der entsprechenden Daten gilt daher als Mindeststandard. In Verbünden organisierte Bibliotheken katalogisieren in der Regel bereits in Verbunddatenbanken und wollen die Katalogisate dann verzögerungsfrei in die lokalen Systeme übernehmen.

Für Informationen zu elektronischen Medien gibt es neben den Verbunddatenbanken weitere Datenbanken bzw. Knowledge Bases, aus denen Paket- und Lizenzinformationen hervorgehen. Diese sind zum Beispiel

* die Zeitschriftendatenbank (ZDB) als zentrales Nachweissystem für Zeitschriften und Fortsetzungen in deutschen und österreichischen Bibliotheken
* die GoKB als kooperativ gepflegte Knowledge Base für elektronische Ressourcen

### Statistik und Reporting

Mitunter verfügen BMS über eigene Module für die Erstellung von Statistiken. Folgende Statistiken sind typischerweise erforderlich:

* Arbeitsstatistiken - Für die tägliche Arbeitsorganisation und die Abrechnung der Arbeitsleistungen gegenüber den Stakeholdern müssen in regel- und unregelmäßigen Abständen Statistiken und Bericht aus dem BMS erstellt werden. Die Inhalte werden von den Stakeholdern bestimmt.
* Deutsche Bibliotheksstatistik - Bibliotheken können sich entscheiden, Daten für die [Deutsche Bibliotheksstatistik](https://www.bibliotheksstatistik.de/) zu erfassen. Die notwendigen Daten sollten über das BMS ermittelt werden können. Durch die einheitliche Definition der statistischen Kennzahlen ist eine umfassende, vergleichende Auswertung aller Bibliothekssparten (wissenschaftliche, öffentliche, Spezialbibliotheken) möglich.
* Sonderstatistiken wie Statistiken der Fachinformationsdienste (FID)

Bei den Systemen der 1. und 2. Generation ist es bisweilen nötig zusätzliche Werkzeuge zum Einsatz zu bringen, um alle gewünschten Berichte zu erstellen (z.B. *BibControl* oder *Crystal Reports*). Während die integrierten Module vor allem auf die Daten des eigenen Systems fokussiert sind, können externe Werkzeuge auch Fremddaten aufnehmen, zum Beispiel Daten aus Besucherzählern.

### Bibliotheksorganisation

Bei der Implementierung oder Anpassung eines BMS ist die Organisation der Bibliothek, die Gestaltung der Prozesse sowie die räumliche Situation zu berücksichtigen. Handelt es sich zum Beispiel um einen öffentliche oder wissenschaftliche Bibliothek? Ist die Organisation als ein einschichtiges oder zweischichtiges System angelegt? Ist es eine einzelne Bibliothek oder eine Zentralbibliothek mit Zweigstellen?

Auch die Aufstellung der Medien innerhalb der Gebäude nimmt Einfluss auf die Ablauforganisation und damit die Konfiguration des Systems. Dies lässt sich anhand der folgenden Beispiele darstellen:

* Bücher können in einer anderen Zweigstellen ausgeliehen werden.
* Die Bibliothek verfügt über einen Magazinbestand, also physische Medien, die für die Nutzer\*innen nicht unmittelbar zur Verfügung stehen.

In beiden Fällen muss auch der Bestellprozess über das System abgebildet werden. Im Magazin bzw. der Zweigstelle sind der Anschluss und die Aufstellung von Druckern für die Erzeugung von Bestellzetteln zu berücksichten. Sind die Medien für die Bibliotheksnutzer\*innen direkt zugänglich, entfällt der Bestellschritt und der abzubildende Prozess beginnt mit der Ausleihverbuchung.

Auch das Rechtemanagement eines BMS ist abhängig von der Größe der Organisation. So sind ggf. verschiedene Berechtigungsstufen für die Bearbeitung von Daten im BMS für die Bibliotheksbeschäftigen einzuführen. Die Berechtigungen bilden die Arbeitsorganisation ab und berechtigen z.B. zum Lesen, Anlegen, Editieren oder Löschen von Ausleihbestellungen, Nutzer\*innen- oder Katalogdaten, Erwerbungsunterlagen, Gebühreninformationen u.ä.

### Benutzungsbedingungen

Die Benutzungsbedingungen werden durch die Ausleihpolitik der Bibliothek bestimmt. Die Gestaltung der Bedingungen erfolgt sowohl bezogen auf die Medien und die Bibliotheksnutzer\*innen. Dabei geht es um die Frage, was von wem ausgeliehen werden darf und, wenn eine Ausleihe möglich ist, wie und für welchen Zeitraum diese erfolgen kann.

Eine Grundlage zur Abbildung der Benutzungsbedingungen ist die Definition von Benutzungsgruppen. Die Benutzungsgruppen werden durch verschiedene Kriterien charakterisiert. Zur Illustration zwei Beispiele:

* Gruppenbildung Universitätsbibliothek:
  + intern: Studierende, Lehrende, weitere Universitätsangehörige
  + extern: externe Wissenschaftler\*innen, interessierte Öffentlichkeit
* Gruppenbildung öffentlichen Bibliothek: Kinder, Jugendliche, Erwachsene

Die Einteilung von Bibliotheksnutzer\*innen in Gruppen dient der einfacheren Zuweisung von Rechten und Ausleihbedingungen, wird aber auch für statistische Zwecke genutzt. Die Ausleihpolitik bestimmt, welche Rechte den verschiedenen Benutzungsgruppen zugewiesen werden. So erfolgt z.B. die Gruppeneinteilung in öffentlichen Bibliotheken i.d.R. nach dem Alter. Einerseits wird damit die Zugänglichkeit der Medien für Kinder und Jugendliche gesteuert. Andererseits dient diese Gruppierung der Einstufung der Gebühren (Kinder und Jugendliche zahlen oft weniger oder keine Gebühren).

Neben den Gruppen werden Ausleihbedingungen auch auf die Medien bezogen. Zur Illustration:

* Präsenzbestände vs. ausleihbare Medien,
* besonders wertvolle Medien oder
* elektronische Publikationen, die nur unter bestimmten Bedingungen und von bestimmten Benutzergruppen genutzt werden können.

Benutzungsbedingungen werden also sowohl durch die Zugehörigkeit zu einer Benutzungsgruppe als auch durch das Medium selbst bestimmt. Die Beschreibung der Benutzungsbedingungen ist somit eine wesentliche Voraussetzung für die Einrichtung des Ausleihmoduls eines BMS.

## Datenverwaltung in BMS

Ein BMS verwaltet zum einen Daten über die von der Bibliothek bereitgestellten oder vermittelten Ressourcen (vor allem physische und digitale Medien) und zum andern Daten über wesentlichen Arbeitsprozesse (beispielsweise Erwerbung und Ausleihe). Dabei lassen sich grob drei Arten von Daten unterscheiden:

* **Bibliographische Metadaten** zur Beschreibung von Ressourcen
* **Digitale Inhalte** wie Volltexte, Digitalisate und andere digitale Publikationen
* **Verwaltungsdaten** zur Unterstützung von Workflows

Die Datenhaltung erfolgt in der Regel in relationalen Datenbanken (MySQL, Oracle).

Zur sinnvollen Verarbeitung von Daten im BMS und in Integration mit anderen System müssen Daten bestimmten Datenformaten entsprechen, über Schnittstellen abruf- und ggf. änderbar sein und Mindestanforderungen an die Datenqualität genügen.

### Bibliographische Metadaten

Bibliographische Metadaten in Form von Titel-, Exemplar- und Normdaten bilden den Kern den klassischen Katalogs. An anderer Stelle tauchen diese Daten beispielsweise als Suchindizes für Discovery-Systeme auf. Das BMS verwaltet diese Daten um Ressourcen Auffindbar und Zugreifbar zu machen. Diese Daten können von verschiedenen Bibliotheken gemeinsam genutzt und in der Regel frei zur Verfügung gestellt werden (siehe Kapitel zu Open Data).

### Digitale Inhalte

Dies sind letztendlich die Daten die für die Nutzer\*innen der Bibliothek vor allem von Interesse sind. Im Falle von Open Access Publikationen bietet das BMS nur einen möglichen Weg zum Zugriff, für erworbene oder lizenzierte Inhalte muss das BMS dagegen unterschiedliche Zugriffsrechte unterstützen.

Digitale Inhalte werden in der Regel nicht direkt im BMS sondern in eigenen Content Management Systemen (CMS) und Repositorien verwaltet. Ein BMS muss mit diesen Systemen durch Verwendung gemeinsamer Datenformate, Import, Export und Verlinkung zusammenarbeiten können. Der Unterschied zwischen Metadaten und Inhalten ist dabei mitunter fließend und hängt vom Anwendungsfall ab. Reicht es oft Publikationen grob mit Metadaten zu beschreiben, so umfasst in anderen Fällen die Erschließung von Publikationen auch Dokumentstrukturen und inhaltliche Bestandteile wie z.B. einzelne Abbildungen.

*Das Kapitel* [*Digitalisierung*](#digitalisierung) *geht ausführlicher auf digitale Inhalte ein.*

### Verwaltungsdaten

Verwaltungsdaten dienen der Unterstützung von Arbeitsabläufen innerhalb der Bibliothek (siehe [Prozessabbildung](#prozessabbildung)). Diese Daten sind zum größten Teil nicht öffentlich und müssen insbesondere im Falle von Daten von Nutzer\*innen im Rahmen des [Datenschutz](#datenschutz) vertraulich behandelt werden.

Zur Interoperabilität mit anderen Informationssystemen innerhalb der eigenen oder übergeordneten Einrichtung gibt es in der Regel nur wenig übergreifend etablierte Standards und Schnittstellen, so dass hier oft zusätzliche Anpassungen an das BMS notwendig sind.

### Datenformate und Schnittstellen

Da Computer nicht selbständig mitdenken und interpretieren können, müssen Daten nach klar definierten Regeln aufgebaut sein. Diese Regeln sollten möglichst genau dokumentiert sein. Damit verschiedene Systeme Daten austauschen können, sollten möglichst etablierte Standardformate verwendet werden.

Trotz gemeinsamer Standards ist ein genaues Hinschauen immer erforderlich, da sich die Handhabung gleicher Formate in der Praxis zwischen verschiedenen Systemen und Einrichtungen unterscheidet.

Neben Standardformaten gibt es speziellere Anwendungsformate. Diese basieren allerdings in der Regel auf allgemeinen Strukturierungssprachen (*CSV*, *XML*, *JSON* oder *RDF*) die je nach BMS besser oder schlechter unterstützt werden.

* Beispiele für bibliographische Standardformate sind *MARC21*, *BIBFRAME* und als kleinster gemeinsamer Nenner Dublin Core. Das PICA-Format bzw. darauf aufbauende Formate ist vor allem als Internformat in den Bibliotheksverbünden GBV, BSZ und an der DNB verbeitet.
* Verbreitete Metadaten-Schnittstellen sind Z39.50, *SRU* und *OAI-PMH*.
* Beispiele für relevante Formate und Schnittstellen für digitale Inhalte sind PDF, *METS/MODS* und *IIIF*.
* Beispiele für relevante Schnittstellen für BMS-Verwaltungsdaten sind *LDAP* und *PAIA*.

Eine umfassende Übersicht von Datenformaten mit Schwerpunkt auf Formate, die für Bibliotheken relevant sind, bietet die Seite <https://format.gbv.de>.

### Datenqualität

Im Gegensatz zu physischen Dingen ist Daten von außen nicht anzusehen, ob sie unvollständig, veraltet oder aus anderen Gründen fehlerhaft sind. Ohne kontrolliertes Qualitätsmanagement muss davon ausgegangen werden, dass die Qualität von Daten kontinuierlich abnimmt. Zur Ermittlung und Verbesserung der Datenqualität tragen bei:

* **Richtlinien** legen einheitliche Regeln für Daten fest, beispielsweise durch Katalogisierungsregeln wie RDA (Soll-Stand)
* **Validierung** ermittelt die Übereinstimmung von Daten mit formal definierten Vorgaben (Ist-Stand)
* **Statistiken** geben quantitative Auskunft, zum Beispiel über die Anzahl erfolgreich importierter oder exportierter Datensätze

Nicht zuletzt beeinflussen auch die Möglichkeiten der Ein- und Ausgabe von Daten ihre Qualität, beispielsweise über die Usability der Katalogisierung.

## Marktanalyse und Beschaffung

|  |
| --- |
| Info |
| Der deutschsprachige [BMS-Markt 2022](#tbl-aktuelle-sw-produkte) ist überschaubar. Für den [Entscheidungsprozess](#entscheidungsprozess) sind daher vor allem auch der Umfang der gewünschten und gewichteten Funktionalitäten, [Varianten des Betriebs](#betriebsmodelle) (gehostet oder lokal) oder auch die Mitgliedschaft in einem Verbund als Kriterien heranzuzuziehen. |

Die Beschaffung eines BMS ist für eine Bibliothek eine große Herausforderung, nicht nur wegen der zu kalkulierenden Kosten sondern auch wegen des erheblichen Einflusses auf alle bibliothekarischen Arbeitsschritte. Der Aufwand für die Migration von Altdaten, die Revision von Geschäftsgängen und die Schulung von Personal muss bei der Beschaffung berücksichtigt werden. Nicht zuletzt ist die Wahl eines BMS auch eine strategische Entscheidung, da die Möglichkeiten auf zukünftige Anforderungen einzugehen je nach System und eigenen Ressourcen unterschiedlich ausfällt.

Es kann auch eine ethisch-moralische Entscheidung oder ein Commitment zu einer ökologisch-nachhaltigen Betriebsführung (öffentliche Einrichtungen als Vorzeigecharakter für einen ökologischen Wandel) sein, die Aspekte der nachhaltigen Beschaffung zu berücksichtigen, wie sie sich bei BMS als auch anderen IT-Anwendungen stellen, etwa die Konsequenzen des ökologischen Fußabdrucks der genutzten Infrastruktur (z. B. CO2-Ausstoß des Rechenzentrums).

Auch aus datenschutzrechtlicher Perspektive gibt es Voraussetzungen zu berücksichtigen, die gegen die Anschaffung bestimmter BMS-Lösungen sprechen (siehe Abschnitt [Datenschutz](#datenschutz)).

Vor diesem Hintergrund ist die Auswahlentscheidung für einzelne Bibliotheken oft ein langwieriger Prozess. Bei Teilnahme an einem Verbund können sich Bibliotheken durch diesen über die BMS, die vom Verbund unterstützt werden, informieren und beraten lassen (siehe [Übersicht deutscher Verbundsysteme](https://www.gbv.de/bibliotheken/geografische-uebersicht-der-deutschen-verbundsysteme)). Die Beschaffung und Einführung von BMS liegt immer in der Verantwortung der jeweiligen Bibliothek oder der Einrichtung, zu der die Bibliothek gehört.

Die Gründe für einen Systemwechsel sind primär technischer oder finanzieller Natur. Beispielhaft werden folgend einige Gründe aufgezählt:

* Das Altsysteme ist technisch überholt oder wird nicht mehr gewartet.
* Es fehlen Schnittstellen für die Integration des BMS in die lokale Informationsinfrastruktur.
* Die Kosten für den laufenden Betrieb sind zu hoch und sollen mit einem anderen System gesenkt werden.
* Eine Funktionserweiterung, z.B. für die Verwaltung von elektronischen Ressourcen, kann an dem bestehenden System nicht mehr vorgenommen werden.

### Entscheidungsprozess

Ein Entscheidungsprozess umfasst typischerweise folgende Schritte:

1. **Workflowanalyse:** Dokumentation bestehender und zukünftig gewünschter Prozesse, die mit dem BMS abgebildet werden sollen
2. **Anforderungsanalyse:** Zusammenstellung und Priorisierung der gewünschten Funktionalitäten und strategischen Zielen unter Einbeziehung aller Stakeholder
3. **Marktanalyse:** Auswahl der in Frage kommenden Lösungen und Betriebsmodelle
4. **Evaluation:** Vertiefte Beschäftigung mit einer Auswahl von Lösungen durch Ausprobieren von Test-Installationen und Kontakt mit Anwendungsbibliotheken
5. **Aufwandsabschätzung** von Migration, Einrichtung und Schulung
6. **Ausschreibung**, falls erforderlich
7. **Auswahlentscheidung**

### Auswahlkriterien eines BMS

Es kann davon ausgegangen werden, dass die aktuell am Markt verfügbaren Systeme die klassischen Geschäftsgänge (siehe Kapitel [Prozessabbildung](#prozessabbildung)) einer Bibliothek gut abbilden können. Die Anforderungen aus dem Kapitel Nutzer\*innenzentrierte Gestaltung gelten grundsätzlich natürlich auch hier.

Die Betrachtung einzelner Systeme einschließlich der Nutzungsszenarien und Use Cases kann sehr aufwändig werden. Daher empfiehlt es sich, die gewünschten Funktionalitäten zu bestimmen und durch die Stakeholder bewerten zu lassen. Die Bewertung kann beispielsweise in Form einer Matrix geschehen, in der die Funktionalitäten nach ihrer Bedeutung/Wichtigkeit einerseits und den zu erwarteten Aufwänden andererseits eingeordnet werden.

|  |
| --- |
| Abbildung 6.2: Beispiel für eine Matrix zur Einordnung von Funktionalitäten |

Zur Evaluierung der BMS können bestehende Anforderungskataloge für die Evaluierung von BMS herangezogen werden, zum Beispiel der gemeinsam [von HBZ und VZG entwickelte Kriterienkatalog](https://www.folio-bib.org/?page_id=247). Dieses umfangreiche Dokument zeigt die Anforderungen an alle Komponenten auf Grundlage der Analyse von sehr ausgereiften Prozessen in Altsystemen auf. Es empfiehlt sich, insbesondere diejenigen Funktionalitäten genau zu überprüfen, die strategisch von besonderer Bedeutung sind.

### Marktanalyse

Da es sich bei BMS um relativ spezialisierte Software handelt und in den letzten Jahren einige Produkte aufgekauft oder eingestellt wurden, ist der Markt sehr überschaubar (siehe Kapitel [aktuell relevanten BMS](#aktuell-relevante-software-produkte)).

Neben der Wahl konkreter Produkte gibt es grundsätzlich drei Möglichkeiten:

* Beitritt zu einem Bibliotheksverbund und Nutzung eines BMS, das von diesem Verbund unterstützt wird, zu den jeweils gültigen Konditionen
* Lizenzierung eines kommerziellen BMS und Einkauf einschlägiger Dienstleistungen für Hosting, Wartung und Support sowie Migration und individuelle Konfiguration
* Implementierung und individuelle Konfiguration eines Open Source-BMS, entweder in Eigenregie oder durch vollständige oder punktuelle Unterstützung von einschlägigen Dienstleistern für Hosting, Wartung und Support sowie Migration und individuelle Konfiguration

Die Vor- und Nachteile im Überblick:

|  | Verbund | kommerzielles BMS | Open Source BMS |
| --- | --- | --- | --- |
| Vorteile | regelmäßige Produktentwicklung  gewisser State-of-the-Art garantiert  klare Kosten- und Leistungsstruktur  große Anwendungscommunity | regelmäßige Produktentwicklung  einheitlicher Leistungsumfang  klare Verantwortlichkeiten | niedrige Anschaffungskosten  große Anwendungscommunities  oftmals regelmäßige Produktenwicklung  viele Dienstleister, die Services rund um Migration, Betrieb und individuelle Anpassung anbieten  offene Schnittstellen und Formate |
| Nachteile | begrenzte individuelle Anpassung  Wartezeiten bei individueller Anpassung | eher geringe individuelle Anpassbarkeit  relativ hohe und intransparente Preise  Abhängigkeit bei der Weiterentwicklung  evtl. Verlust der Datenhoheit  z.T. proprietäre (herstellerspezifische) Systeme und Schnittstellen | erfordert eigene IT-Kapazitäten oder Outsourcing  Risiko der Sicherung von Nachhaltigkeit und Kompatibilität |

Verbünde bieten in der Regel ein oder zwei Lösungen an, die entweder kommerziell oder Open Source sind. Die Mitgliedschaft in Verbünden kann ein kostengünstiger Weg sein, um mit einem BMS und dazugehörigen Dienstleistungen versorgt zu werden. Allerdings steht möglicherweise nicht allen Bibliotheken die Mitgliedschaft in einem Verbund offen oder bedingt andere Nachteile (z.B. den Zwang, an der Fernleihe teilzunehmen und begrenzte Möglichkeiten zur individuellen Anpassung der Software).

Bei kommerzieller Software fallen typischerweise Lizenzkosten an, die sich nach der Größe der Bibliothek oder der übergeordneten Einrichtung richten (z.B. an der Anzahl von Mitarbeitenden, Studierenden oder Einwohner). Dabei werden einmalige Beschaffungs- und jährlichen Wartungskosten unterschieden. Es muß klar vereinbart werden, welche Dienste mit den Wartungskosten (Support, Update auf neue Versionen, …) abgegolten sind.

Der Betrieb der Lösungen kann von den Anbietern oder anderen Dienstleistern (Verbund, andere kommerzielle Anbieter) übernommen werden (Cloud/Software as a Service), d.h. die Bibliotheken brauchen keine eigenen Server zur Verfügung stellen und administrieren. Eine Installation auf eigenen Servern (On-Premise-Lösung) erfordert hingegen eigenes, ausgebildetes Personal.

Bei Open Source-Lösungen gibt es keine initialen Anschaffungskosten. Bei Verfügbarkeit entsprechender Server-Infrastruktur und erfahrenem Personal kann eine Bibliothek die Software selbst installieren und in Betrieb nehmen oder diese Leistungen von Dienstleistern einkaufen.

Die initiale Konfiguration sowie die Migration von Daten aus einem Alt-System können ebenso von den Bibliotheken selbst durchgeführt werden oder sind Teil des Kauf-/Wartungsvertrages.

Die laufende Betreuung des Betriebs von BMS erfordert speziell geschultes und berechtigtes Personal - sogenannte System-Bibliothekar\*innen. In wenigen Fällen wird die Systembetreuung an Dienstleister (beim Hoster) übergeben.

## Prozessabbildung

Für den Einsatz eines BMS bilden Prozessbeschreibungen bzw. Workflows eine wesentliche Grundlage. Auf der Basis der Abbildung der Kernprozesse wie Erwerbung, Katalogisierung, Ausleihe sowie der Rolle des Systems und anderer Akteure können Anpassungen (leichter) vorgenommen werden.

Zur Modellierung, Dokumentation und Visualisierung von Workflows bietet sich klassischerweise eine Modellierungssprache wie BPMN (Business Process Model and Notation) an. Für diese und verwandte Sprachen existieren umfangreiche Werkzeuge und Toolchains, mit denen einerseits Prozesse erstellt werden können, gleichzeitig aber auch – sollte das nötig sein – die modellierten Prozesse automatisiert werden können. Im Endeffekt bedeutet dies, dass aus dem Prozessmodell Programmcode erzeugt wird.

|  |
| --- |
| Abbildung 6.3: Beispiel einer BPMN-Prozessabbildung |

Lässt man die Aspekte der Prozessautomatisierung oder Codegenerierung außer acht, so lässt sich auch eine abgespeckte BPMN-ähnliche Semantik nutzen, um Prozesse zu dokumentieren und zu visualisieren. Andere Alternativen zur Modellierung finden sich in den verschiedenen Diagrammformen der UML (Unified Modelling Language).

Aus heutiger Sicht sollten für die im Folgenden genannten Bereich Prozessbeschreibungen erstellt werden, damit potentielle BMS an Hand dieser geprüft werden können. Hierbei könnten sich Notwendigkeiten für Änderungen in den Prozessabläufen der Bibliothek ergeben, die auf Basis der Beschreibungen genauer adressiert werden können.

### Nutzer\*innen

Als Nutzer\*innen werden in diesem Kapitel diejenigen Menschen bezeichnet, die mit einem BMS interagieren. Man unterscheidet zwischen den Bibliotheksbeschäftigten, die mit dem Modulen Ausleihe, Erwerbung, Katalogisierung, ERM etc. interagieren, und den Bibliotheksnutzer\*innen (oft auch als Leser\*innen bezeichnet), die mit dem BMS über das Modul OPAC oder nur indirekt über ein Discovery-System oder ein anderes Drittsystem mit dem BMS in Kontakt kommen.

### User-Interfaces für verschiedene BMS-Anwender\*innen

Die Bibliotheksbeschäftigten und die Bibliotheksnuter\*innen haben verschiedene Sichtweisen auf ein BMS. Bibliotheksbeschäftigte müssen über das User-Interface bei ihrer Arbeit spezifisch durch die Workflows geführt werden. Dabei ist auf eine einheitliche Benutzungsführung und Gestaltung der Oberfläche zu achten.

Für die Bibliotheksnutzer\*innen steht die Information über die Dienste der Bibliothek, deren Bestand und die Nutzung des Bestandes im Vordergrund. Bibliotheksnutzer\*innen kommen dabei häufig mit mehreren IT-Systemen in Kontakt (BMS- OPAC-Modul, Web-Server, Discovery-System, …). Daher sollte auch hier auf eine einheitliche Oberfläche der eingesetzten IT-Systeme geachtet werden, auch bezüglich Accounts und Login, zumindest aber auf ein einheitliches Design und eine einheitliche Benutzerführung.

Es ergeben sich daraus die folgenden Anforderungen

1. Intuitive Benutzbarkeit
2. Barrierearme Gestaltung
3. Responsivität

Diese Themen werden auch in den Abschnitten zu [rechtlichen Rahmenbedingungen](#rechtliche-rahmenbedingungen) und im Kapitel zu den [Anforderungen an Bibliotheks-IT](#anforderungen) angesprochen.

### Erwerbung

Ein BMS sollte das Bibliothekspersonal bei den folgenden Aufgaben unterstützen:

1. Überprüfen von vorhandenen Beständen (Vorakzession)
2. Aufgabe von Bestellungen bei definierten Lieferanten auf verschiedenen Wegen
3. Verwaltung von Lieferantendaten
4. Anlegen und Verwalten von Bestellungen von Zeitschriften und Fortsetzungswerken
5. Überwachung von Bestellungen
6. Anlegen und Verwalten von Budgets
7. Akzessionierung von Medien
8. Rechnungsverwaltung inkl. Schnittstellen für haushalterische Systeme
9. Verwaltung von Bindeaufträgen
10. Verwaltung von Nicht-Kauf-Bestellungen

Diese Aufgaben lassen sich mit den am Markt befindlichen Systemen in der Regel gut abbilden. Allerdings werden die meisten Bibliotheken für die Verwaltung von notwendigen Bestellungen von Materialien jenseits des Bibliotheksbestandes (Büromaterial, IT-Ausstattung etc.) zusätzliche haushalterische Systeme einsetzen. Das Erwerbungsmodul ist insofern meist nur eine Komponente im Haushaltswesen.

### Verwaltung von elektronischen Ressourcen

Für die Verwaltung elektronischer Ressourcen sollten folgende Aufgaben unterstützt werden:

1. Erfassung von Lizenzinformationen nach unterschiedlichen Erwerbungsmodellen wie Pakete, Allianz- oder Nationallizenzen
2. Zuordnung von digitalen Inhalten zu Paketen
3. Verwaltung von Paketen
4. Bezug von bibliografischen Daten von Aggregatoren und Verlagen
5. Unterstützung der direkten Verlinkung auf Volltexte aus Katalogen und Discovery-Systemen
6. Auslieferung von aussagekräftigen Zugangsinformationen in Kataloge und Discovery-Systeme
7. Unterstützung bei der Bereitstellung von digitalen Inhalten jenseits von proprietären Apps

Die BMS der 1. und 2. Generation haben erhebliche Defizite bei der Verwaltung von elektronischen Ressourcen. Die Bereitstellung von entsprechenden Funktionalitäten ist daher ein Alleinstellungsmerkmal von BMS der neuen Generation.

Alternativ können aber auch separate, sogenannte Electronic Resource Management-Tools eingesetzt werden (Coral, GoKB und LAS:eR).

### Katalogisierung

Bei der Katalogisierung müssen folgende Tätigkeiten unterstützt werden:

Erfassung von unterschiedlichen Medientypen gemäß aktueller Metadaten-Standards

1. Übernahme von Katalogdaten aus Bibliotheksverbünden
2. Möglichkeit der Integration von Normdaten
3. Erfassung von lokalen Daten
4. Konfigurierbarkeit von Erfassungsmasken

### Katalog

Der Katalog ist die Sicht für die Bibliotheksnutzer\*innen auf die Bestände der Bibliothek. An das Katalogmodul werden folgende Anforderungen gestellt:

1. Web-Interface nach aktuellen Standards bezüglich Barrierefreiheit, Responsivität etc.
2. Angebot von Möglichkeiten der Suche nach bekannten Titeln
3. Angebot von Möglichkeiten der Suche nach Themen
4. Filterung von Trefferlisten nach formalen oder inhaltlichen Kriterien bzw. Standorten
5. Anzeige von Verfügbarkeitsinformationen
6. Anzeige von Neuerwerbungslisten
7. Kontobezogene Funktionalitäten (Einsicht, Verlängerung, Vormerkung, Bestellungen)
8. Anzeige von Neuigkeiten und wichtigen Links auf der Startseite des Katalogs
9. Anpassbarkeit der Katalogoberfläche an das Corporate Design (wenigstens Logo und Farbschema)

In der Geschichte der BMS war das Katalogmodul eher ein Nebenprodukt der Katalogisierungsarbeit. Durch die Veränderungen im Informationsverhalten seit Entwicklung des WWW ist insbesondere auf das Katalogmodul ein besonderer Innovationsdruck entstanden. Auf diesen Druck haben Bibliotheken mit dem Angebot von Discovery-Systemen reagiert, die als alternative Benutzungsschnittstelle zu den klassischen OPACs aufgebaut wurden und neben einem modernen Design auch Suchmaschinen-typische Funktionen wie Facettierung oder Unterstützung bei der Formulierung von Suchbegriffen bieten. Diese Funktionen sind in den BMS der neuen Generation standardmäßig enthalten.

Es entstanden durch den erwähnten Innovationsdruck verschiedene Konstrukte, die Daten der Bibliothek den Nutzer\*innen zur Verfügung zu stellen:

1. Klassischer Katalog (OPAC) als Bestandteil des BMS
2. Katalog als separates Modul (nicht Bestandteil des BMS), selbst entwickelt, zugekauft oder als Open Source
3. Discoverysystem als Bestandteil des BMS: Daten aus dem eigenen Bestand sowie Fremddaten, die als Metadaten zur Verfügung stehen.
4. Discoverysystem als zugekauftes Modul eines anderen Herstellers oder als Eigenbau mit zugekauften Metadaten oder als Open Source mit offenen Daten oder zugekauften Metadaten.

Bei Punkt 4 entsteht die Herausforderung, die im BMS gehosteten Informationen, zum Beispiel über den Ausleihstatus/Verfügbarkeit, auch in der Oberfläche des Discovery-Systems aktuell darzustellen.

### Ausleihe

Ein BMS sollte die folgenden Aufgaben der Ausleihe unterstützen:

1. Anlegen von Benutzergruppen, Standorten, Medienarten
2. Abbildung der in den Benutzungsordnungen festgelegten Ausleihbedingungen, z.B. Leihfristen nach Benutzergruppen, Standorten, Medienarten
3. Verbuchung von Medien (Ausleihe, Rücknahme)
4. Konfiguration von Ausdrucken für Bestellzettel und Vormerkungen
5. Ermöglichen von Bestellungen und Vormerkungen
6. Mahnwesen (Fristen, Mahnstufen)
7. Benachrichtigungen für Bestellungen, Vormerkungen, Mahnungen, Leihfristerinnerungen
8. Gebührenverwaltung
9. Erzeugung von Listen (überfällige Medien, nicht abgeholte Vormerkungen)
10. Anbindung an Bezahlsysteme (Kassenautomaten, Online-Bezahlsysteme)
11. Anbindung von Verfügbarkeits- und Kontoinformationen an Discovery-Systeme
12. Anbindung an Automatisierungslösungen und externe Verbuchungssysteme (etwa mittels RFID)

Die Parametrisierung der Ausleihe ist ein besonders komplexer Bereich der BMS-Installation aufgrund der Vielzahl von zu beachtenden Benutzungsregeln, der Sensibilität der Daten und der besonderen Relevanz eines reibungslosen Betriebs beim Versand von Benachrichtigungen. Ein Beispiel für eine solche Komplexität ist die das Verhalten bei Feiertagen: Hier muss ein Schließtagekalender regelmäßig gepflegt werden, um zu vermeiden, dass Leihfristenden auf Feiertage oder Wochenenden fallen.

### Automatisierung und Selbstbedienung

Als Automatisierung wird die Möglichkeit bezeichnet, die Geschäftsgänge einer Bibliothek mit digitalen Werkzeugen abzubilden und durchführen zu können. Dazu sind Maschinen notwendig, die die entsprechenden Funktionen anbieten. Das schon recht betagte Standardprotokoll für die Kommunikation zwischen BMS und Automat ist SIP2. Dieses Protokoll hat den Nachteil, dass es ohne Verschlüsselung entwickelt wurde und daher - sofern es sich um ein BMS in der Cloud handelt zumindest - über [stunnel](https://en.wikipedia.org/wiki/Stunnel) verschlüsselt getunnelt wird. Moderne BMS unterstützen mittlerweile zusätzlich auch allgemeine Kommunikationsprotokolle, etwa über *REST*, sodass das Tunneln von Verbindungen nicht mehr nötig ist. Außerdem ist man nicht mehr daran gebunden, dass anzubindende Geräte SIP2 unterstützen, was deutlich mehr Marktalternativen öffnet.

Nachfolgend werden Automaten für die Selbstbedienung im Bereich der Ausleihe dargestellt.

#### Selbstverbucher / Ausleihautomaten

Selbstverbucher / Ausleihautomaten bestehen meist aus einer Auflagefläche für die auszuleihenden Medien, einer Schnittstelle für Bibliotheksausweise sowie einem PC, der die Endgeräte verwaltet und mit dem BMS kommuniziert. Bei einer funkgestützten Medienerkennung (RFID) gibt es die Möglichkeit der Stapelverbuchung, es werden also vom Automaten mehrere gestapelte Bücher erkannt und zur Verbuchung angeboten. Bei einer Barcodegestützten Medienerkennung wird jedes Medium einzeln verbucht.

Bibliotheksausweise gibt es in verschiedensten Ausprägungen: Barcode (1D-Code), Funkchip (u.U. proprietär, Bsp.: Intercard), QR-Code (2D-Code). Die 1D- oder 2D-Codes können entweder auf Papier oder in einer App auf dem Smartphone beigebracht werden. Die Schnittstelle im Automaten muss auf die vorhandenen Ausweistypen vorbereitet sein.

Bei Nichtvorhandensein einer separaten Rückgabeanlage kann der Selbstverbucher / Ausleihautomat auch eine Rückgabefunktion anbieten. Zumeist werden die zurückgegebenen Medien unsortiert gesammelt; im Anschluss erfolgt die Sortierung durch das Bibliothekspersonal.

Nach der Rückgabe- oder Ausleihverbuchung muss der Selbstverbucher / Ausleihautomat auch die Buchsicherung (sofern vorhanden) bedienen. Bei der in vielen Bibliotheken auslaufenden EM-Sicherung (elektromagnetisch über einen im Medium eingeklebten magnetisierbaren Metallstreifen) geschieht dies über die Ansteuerung eines Elektromagneten mit hörbarem Feedback an die Nutzer\*innen („klack“). Bei RFID-Sicherung wird bei erfolgter Verbuchung ein Sicherungsbit auf dem RFID-Chip verändert. Aufgrund der größeren Geschwindigkeit dieses Vorganges geschieht dies ohne Feedback an die Nutzer\*innen.

#### Rückgabeautomat / -sortierung

Ein separater Rückgabeautomat hat zum einen den Vorteil, dass die Prozesse Ausleihe und Rückgabe bei starker Nutzung entzerrt werden und zum anderen, dass eine Sortierung der zurückgegebenen Medien möglich ist. Die Medien werden von den Nutzer\*innen auf ein Förderband gelegt und eingezogen (außer Reichweite des Nutzer\*innen. In dieser Position wird der Barcode auf dem Medium oder der RFID-Chip gelesen. Wird keines der beiden erkannt, wird das Medium wieder zurückgegeben. Bei erfolgreicher Erkennung und Verbuchung im BMS (und anschließender Aktivierung der Buchsicherung) wird im BMS mithilfe der Signatur oder Mediennummer erfragt, wie das Medium sortiert werden soll. In den meisten BMS gibt es dazu Tabellen, die z.B. über die Anfänge von Signaturen oder anderen Kriterien (Bsp: „SN …“ in Wagen 3, „ist vorgemerkt“ in den Wagen x) arbeiten. Steht das Sortierziel fest, wird das Medium über Förderbänder zu dieser Stelle transportiert und abgeworfen. Das Ziel kann ein sog. Tray sein, ein oben offener Korb oder Wagen, oft mit einem gewichtgesteuerten Boden, damit die Medien nicht allzu tief fallen. Alternativ bieten immer mehr Hersteller sog. Ergocarts an, auf die die Medien so geschichtet werden, dass sie am Regal Rückenschonend aus einem Stapel entnommen und einsortiert werden können.

Üblicherweise gibt es am Rückgabeautomaten keine Authentifizierung.

Es gibt auch Rückgabeautomaten, die eine erneute Ausleihe des gerade zurückgegebenen Werkes an den/ie gleiche Bibliotheksnutzer/in ermöglichen. Dies ist in den Fällen sinnvoll, wenn die maximale Leihfrist / maximal mögliche Verlängerungen der Leihfrist erreicht ist und den/die Bibliotheksnutzer/in das Buch weiter nutzen möchte und das Medium nicht anderweitig bestellt ist.

#### Kassenautomat

Ein Kassenautomat ermöglicht die personalfreie Bezahlung der offenen Gebühren. Auch hier wird erst der Nutzungsausweis eingelesen und nach einer optionalen Passworteingabe die offenen Gebühren angezeigt. Die Gebühren können dann mit Bargeld oder Bargeldlos gezahlt werden. Auf eine Bargeldzahlung wird zunehmend verzichtet, da das Handling von Bargeld aufwändig und teuer ist.

#### Fernleihautomat

Aus einem Fernleihautomat können Fernleihen personalfrei an Nutzer\*innen ausgegeben werden. Da diese Bücher weder mit dem eigenen System der Bibliothek gesichert noch verbuchbar sind, muss eine separate Verbuchung durchgeführt werden. Die Nutzer\*innen bekommen eine Nachricht, dass ihr bestelltes Medium in einem Fach mit der Nummer xy bereit liegt sowie eine PIN zur Öffnung dieses Faches. Sobald das Fach geöffnet wird, wird das Medium auf das Konto des Nutzers/der Nutzerin verbucht. Auch eine Öffnung des Faches mit einem funkgesteuerten Nutzungsausweis statt der PIN ist möglich.

#### Sicherungsgates

Sicherungsgates erkennen unverbuchte Medien, die die Bibliothek verlassen. Die dafür übliche Technik war in den letzten Jahrzehnten die EM-Sicherung, also die Erkennung der Magnetisierung von metallischen Streifen, die in die Medien geklebt waren. Mit der Umstellung auf RFID geschieht die Buchsicherung über Funk, ein Sicherungsbit im Speicher der RFID-Chips wird untersucht. Bei EM-Sicherung ist der maximale Abstand zwischen zwei Gates zur halbwegs zuverlässigen Erkennung ca 90 cm und stellen somit eine Einschränkung des Zugangs, z.B. bei der Nutzung mit Rollstühlen, dar. Etwa der gleiche Abstand ist notwendig bei RFID-HF, bei RFID-UHF (Reichweite bis zu 10m) ist ein sehr großer Abstand möglich und somit der Verzicht auf eine Einengung des Ausgangs.

Bei Erkennung eines gesicherten (und nicht entliehenen) Mediums ertönt ein Warnton. Bei manchen Systemen wird das entsprechende Medium mit Titel und Cover auf einem Monitor angezeigt.

Sicherungsgates verhindern nicht Diebstahl, Diebe wählen andere Wege. Sicherungsgates verhindern das versehentliche Verlassen der Bibliothek mit unverbuchten Medien.

### Anbindung von Systemen über Schnittstellen

Ein BMS muss in der Lage sein, mit anderen Systemen automatisiert Daten auszutauschen. Diese Austauschprozesse betreffen folgende Szenarien

1. Bereitstellung von Konto- und Verfügbarkeitsinformationen, z.B. über [PAIA](https://verbundwiki.gbv.de/display/VZG/PAIA) und [DAIA](https://verbundwiki.gbv.de/display/VZG/DAIA)
2. Anbindung an Buchhaltungssysteme wie SAP oder HIS Haushalt-ERO
3. Anbindung an Tools für statistische Auswertungen (s.a. [Kapitel Statistik](#statistik))
4. Bereitstellung von bibliografischen Daten
5. Recherche in Fremddatenbeständen, z.B. über Z39.50
6. Schnittstellen zu Kataloganreicherungsdiensten (Buchcover)
7. Schnittstellen zu IDM-Systemen (s.a. [Kapitel IDM](#identity-management))
8. Schnittstellen zu einschlägigen Plattformen der jeweiligen Zielgruppen, zum Beispiel Lernmanagementsysteme

Die Systeme der neuen Generation verfügen in der Regel über Schnittstellen, über die sie in die bestehenden Informationsinfrastrukturen, d.h. die umgebenden Systeme, eingebunden werden können.

Eine Schnittstelle (engl. Interface oder manchmal auch API - application programming interface) bildet einen definierten Kommunikationsweg zwischen Verschiedenen Systemen als „Gesprächspartner“. Im bibliothekarischen Universum gibt es für diese Fälle auch schon viele etablierte Austauschformate, etwa SIP2. Ein BMS „von der Stange“ kann im Regelfall die üblichen Austauschformate unterstützen, sodass ein Austausch zwischen den gängigen Systemen einfach möglich ist. Hierzu zählen insbesondere der jeweilige Bibliotheksverbund, etwa zum Austausch von Metadaten oder für das verteilte Lizenzmanagement, aber auch nutzer\*innen-nahe Dienstleistungen, wie die Fernleihe.

### Nicht-bibliothekarische Schnittstellen

Ein BMS existiert im Regelfall nicht nur für sich oder nur im Kosmos der eigenen und anderer Bibliotheken, sondern ist auch in die lokalen IT-Strukturen eingebunden.

Ein gutes Beispiel ist der Einsatz eines BMS an einer Hochschule: Im Regelfall sind alle Mitglieder einer Hochschule auch gleichzeitig (potenzielle) Nutzer\*innen der Bibliothek. Die Daten der Mitglieder dieser Einrichtung werden an einer zentralen Stelle verwaltet und sollen durch andere Systeme, z.B. im Bibliothekssystem, durch Verknüpfung nachgenutzt werden. Dies ist die Rolle des Identity Managements (IDM).

#### Identity Management

Ein IDM (Identity Management System) ist ein System, mit dem die Basisdaten von Personen und Gruppen an zentraler Stelle verwaltet werden könne. Dies sind etwa persönlich Daten, Kontaktdaten und Organisationsstrukturen. Der Gedanke daher ist, dass alle relevanten Informationen nur an einer zentralen Stelle vorgehalten werden, und in anderen Systemen keine Dubletten erzeugt werden, die dann auch noch irgendwie synchronisiert werden müssten. Damit Personen in einem System eindeutig identifiziert werden können, existiert zumeist eine oder mehrere eindeutige IDs, etwa die Matrikelnummer eines Studierenden.

Das IDM hält im Regelfall mehr Daten über eine/n Nutzer/in bereit, als von den jeweiligen verbundenen Systemen benötigt werden. Beispielsweise könnte in einem IDM vorgehalten werden, dass eine Person Mitarbeiterin einer Hochschule ist, dass sie zu einer gewissen Fakultät der Hochschule gehört und dass sie zu einer bestimmten Arbeitsgruppe gehört. In der Kommunikation des BMS mit dem IDM ist jedoch nur die erste der Informationen relevant, etwa um die Ausleihkonditionen der Person festlegen zu können. Daher wird in der Kommunikation mit einem IDM im Regelfall auch auch gewisser Scope mitgegeben, damit nur die wirklich für das konsumierende System relevanten Informationen mitgegeben werden; Prinzip „Datensparsamkeit“.

Ein IDM kann als Identity Provider zu einem Authentifizierungsdienst werden. Über diesen Dienst kann man dann unter Umständen ein Single Sign On realisieren, bei dem die Daten der Nutzer\*innen nicht mehr an den Service oder Content Provider weitergegeben werden sondern nur noch eine Art Ticket, dass eine Erlaubnis regelt. Im Idealfall gilt diese einmalige Anmeldung dann für einige/viele Service-/Contentprovider, so dass der/die Nutzer/in sich nur einmal anmelden muss um viele Dienste zu nutzen.

Authentifizierungsprotokolle sind bspw.: [Shibboleth / SAML2](https://en.wikipedia.org/wiki/Shibboleth_(software)), [OpenID](https://de.wikipedia.org/wiki/OpenID)

Softwareprodukte für IDM sind: SAP (mit Plugins), Microsoft Active Directory, uvm

##### Speicherung von Nutzer\*innendenaccounts

Ein Account besteht aus den Kontaktdaten des Menschen sowie Authentifizierungsinformationen. Hier ist Datensparsamkeit nach DSGVO geboten. Für die Speicherung aller personenbezogenen Daten müssen die Notwendigkeiten oder rechtlichen Gründe nachgewiesen werden. Als Beispiel kann die Speicherung des Geburtsdatums angesehen werden. Wird für die Begründung für die Speicherung des Geburtsdatums die Prüfung der Volljährigkeit oder die Befähigung eines Seniorentarifes herangezogen, ist davon auzugehen, dass die Speicherung des Geburtsdatums nicht notwendig ist. Wird zur Begründung eine als notwendig erachtete Adressermittlung bei Behörden angegeben, ist die Speicherung der Geburtsdatum möglich, das eine Adressermittlung (zur Wiederbeschaffung vermisster Exemplare) möglich. Die Speicherung nutzungsbezogener Daten wie Verweise auf die ausgeliehenen Medien, angefallene Gebühren, offene Bestellungen und bestellte Digitalisate muss in der Regel nicht explizit begründet werden.

Sofern die übergreifende Institution über eine Datenbank zur Speicherung der Accounts verfügt ([IDM](#identity-management), Identity Management) ist eine Anbindung an diese sinnvoll. Dieses IDM enthält dann allerdings nicht notwendigerweise die externen Nutzer\*innen.

Die technisch einfachste Lösung für Accounts der externen Bibliotheksnutzer\*innen ist die Speicherung im IDM der übergeordneten Einrichtung, sofern vorhanden. Komplexer ist die Speicherung in einem separaten System, da dann bei Autorisierung u.U. mehrere Systeme abgefragt werden müssen.

Datenschutzbezogene Vorgehensweisen auch in Bezug auf personenbezogene und personenbeziehbare Daten von Nutzer\*innen finden sich in Abschnitt [Datenschutz](#datenschutz) im Kapitel zum technischen Betrieb eines BMS.

#### Bezahlsysteme

Online-Payment, Kassensysteme/-automaten (siehe auch Kapitel [Kassenautomat](#kassenautomat))

#### E-Rechnung

E-Rechnungen müssen seit 2020 von Einrichtungen des Bundes, der Länder und Kommunen verarbeitet werden können. Der Umgang mit E-Rechnungen ist sehr unterschiedlich geregelt. Zum Teil nehmen Einrichtungen nur noch an einer zentralen Stelle E-Rechnungen entgegen. In anderen Einrichtungen werden E-Rechnungen dort entgegen genommen, wo die Bestellungen ausgelöst wurden. Es gibt verschiedene Formate in der eine E-Rechnung übermittelt werden kann (PDF, XML oder direkt per [EDIFACT](https://de.wikipedia.org/wiki/EDIFACT)).

Elektronische Rechnungen kommen immer dann ins Spiel, wenn Bestellvorgänge von neuen Medien über das BMS abgewickelt werden. In diesem Zusammenhang entstehen Rechnungen von Lieferanten, die von der Bibliothek oder ihrer Organisation zu begleichen sind.

Ohne eine „E-Rechnungs-Workflow“ würde dies bedeuten, dass Rechnungen der Lieferanten bei der Bibliothek eingehen, einem Bestellvorgang zugeordnet werden müssen, von der jeweiligen Rechnungsstelle beglichen und schließlich wieder im BMS „abgehakt“ werden müssen. Diese repetitiven Workflows lassen sich mittlerweile weitgehend automatisieren. Das BMS ist in der Lage elektronisch übermittelte Rechnungsdaten automatisiert den jeweiligen Bestellprozessen zuzuordnen Bei einer gleichzeitigen Anbindung eines elektronischen Rechnungswesens z.B. über SAP können auch die Zahlungsinformationen automatisiert zugeordnet werden und somit ein Bestellvorgang komplett automatisiert abgeschlossen werden.

#### Statistik

Mit dem Begriff „Statistik“ können verschiedene Dinge im Rahmen eines BMS gemeint sein, etwa Betriebsstatistiken, wie die Rechnerauslastung eines Servers, auf dem das BMS betrieben wird. In diesem konkreten Fall ist jedoch mit „Statistik“ gemeint, dass die Nutzungsdaten des BMS in eine Form gebracht werden können, in der Mitarbeiter\*innen der Bibliothek Informationen ziehen können, die zur Dokumentation, zum Reporting oder zur weiteren Arbeit benutzt werden können.

Beispiele für Statistiken sind ganz klassische die Ausleihzahlen einer Bibliothek, ggf. aufgeteilt nach verschiedenen Themen oder Fächern, die den Bedarf der an bestimmten Medien preisgeben. Dies könnte für die Leitung einer Bibliothek relevant sein, oder für das Budgetmanagement der Einrichtung. Ein anderes Beispiel wären detaillierte Ausleih- und Benutzungsstatistiken, sowie konkrete Informationen zum Bestand. Mit einer solchen Datenlage können Fachbibliothekar\*innen etwa gezielt Bestandsmanagement betreiben.

Manche LMS haben Statistikfunktionen schon mit dabei. Andere halten ihre Daten in einer Datenbank und diese müssen aktiv exportiert werden. Wieder andere bieten entsprechende Schnittstellen, über die statistische Daten exportiert werden können.

Je nach Anforderung an den Umfang und an die Arbeit, die mit statistischen Auswertungen erfolgen soll, kann die Entscheidung fallen, die aus dem BMS kommenden Daten einfach nur in eine Tabellenkalkulation zu exportieren, oder eine speziell auf die statistische Datenanalyse zugeschnittene Statistik-Software einzuspielen. Ergo kommen hier Software wie Excel, BibControl oder gar komplexe Statistik-Plattformen wie SPSS in Frage. BMS wie ExLibris Alma bringen hierbei schon eigene Statistik-Module mit, die eine externe Lösung überflüssig machen.

Alma (und andere BMS) können automatisiert oder manuell [COUNTER](https://www.projectcounter.org/counter-sushi/)-Reports für statistische Daten der Nutzung digitaler Medien importieren.

## Technischer Betrieb

Der technische Betrieb eines BMS variiert je nach [Betriebsmodell](#betriebsmodelle) (lokale Installation, gehostete Variante oder Cloud-Dienst). [Kosten](#kosten) entstehen dabei für Lizenz- und Wartungsverträge sowie für Betriebsressourcen. Für den Betrieb sind weiter das [Monitoring](#monitoring) sowie die Aspekte der [IT-Sicherheit](#it-sicherheit), [Backup](#backup-und-rollback) und [Datenschutz](#datenschutz) zu berücksichtigen.

### Kosten

Die Anschaffungskosten eines BMS machen nur einen kleinen Teil aus. Wichtiger ist, sich über folgende Kosten klar zu werden:

* Personalkosten für den laufenden Betrieb
* Lizenzkosten und Wartungsverträge der Software
* Betriebsressourcen, wie z.B. Serverraum, Energieverbrauch, Wartung, Backuplösungen

Personalkosten und Ressourcen richten sich hauptsächlich nach Art der Installation (Lokal, Hosting oder Cloud). Lizenzkosten sind teilweise nach Größe der Einrichtung gestaffelt, d.h. sie richten sich nach Anzahl der verwalteten Medien und/oder Endnutzer\*innen.

Insbesondere der Punkt Personalkosten kann zu einem Engpass bzw. Risiko werden, denn in vielen Fällen zeigt sich, dass einige wenige Personen durch ein BMS gebunden werden und gleichzeitig auch die einzigen sind, die das System in der Tiefe bedienen können. Wirklich kritisch wird es, wenn nur eine einzige Person diese Rolle erfüllt. Je mehr Verantwortung beim Betrieb auf das Personal vor Ort fällt (lokaler Betrieb), desto wichtiger wird dieser Aspekt. Selbst bei der Nutzung eines Cloud-BMS ist davon auszugehen, dass für die fachliche Administration der Software Personal dauerhaft gebunden ist. Bei dieser Betriebsmethode gibt der Anbieter meistens den Updatezeitpunkt vor, insofern müssen unter Umständen Workflows in der Bibliothek aufgrund von Änderungen in der Software durchgeführt werden, ohne dass man die zeitliche Planung dafür in der Hand hat.

Um Personalengpässe zu vermeiden, ist es sinnvoll, Einführungsprozesse nur in einer Expertengruppe durchzuführen und Verantwortlichkeiten auf mehrere Schultern zu verteilen (Ausfallsicherheit, Urlaubsvertretung usw). Auch die gute Dokumentation teils komplexer Zusammenhänge sollte bedacht werden, damit Fachwissen nicht nur in den Köpfen einiger weniger Mitarbeiter\*innen schlummert.

### Installation & Updates

Zur Einrichtung eines BMS gehört:

* **Installation auf einem Server:** erfordert i.d.R. Kenntnisse in Systemtechnik (Hardware, Server, Kommandozeile...). Wenn Hosting durch Drittanbieter geleistet wird (Cloud, Dienstleister wie Verbundzentrale o.A.), verändert sich diese Aufgabe. Sie entfällt, wenn der Hoster spezialisiert auf das Hosting von BMS ist (bspw. Verbundzentrale), sie wird geringer, wenn der Hoster eher allgemein aufgestellt ist.
* **Konfiguration/Parametrisierung:** Teilweise über Administrator-Oberfläche möglich, teilweise nur über Konfigurationsdateien. Erfordert vor allem Kenntnisse der eigenen IT-Infrastruktur und der verwendeten Schnittstellen und Formate. Die Grenzen zwischen Konfiguration und Programmierung eigener Erweiterungen sind fließend. Zu beachten ist auch die Migration bestehender Daten in das neue System.

Nach Einrichtung werden BMS laufend erweitert. Fehler werden behoben und neue Funktionen kommen hinzu. Die Aktualisierung kann je nach Produkt agil in kleinen, häufigen Schritten erfolgen (monatlich, wöchentlich oder häufigere Updates) oder in längeren Zeitabschnitten (oft quartals- oder halbjahres-weise).

### Open Source

Wird ein System auf Open-Source-Basis eingesetzt, sollte eine Verständigung darüber erfolgen, ob und unter welchen Bedingungen lokale Anpassungen am System auch der Community zur Verfügung gestellt werden. Hierzu müssen die Lizenzbedingungen des Systems geprüft werden.

### Laufender Betrieb

Während des laufenden Betriebs ist es wichtig, sich über den aktuellen Betriebszustand des Systems ein klares Bild machen zu können. Dieser „Statusbericht“ kann sich über alle Ebenen des Systems ziehen: Wie viel Speicherplatz ist noch frei? Ist das System für alle Nutzer\*innen erreichbar? Sind verbundene Systeme verfügbar und betriebsbereit? Je nach Betriebsmodell werden diese Fragestellungen durch klassisches IT-Monitoring abgedeckt, benötigen teilweise aber auch bibliotheksspezifische Lösungen.

#### Monitoring

Monitoring-Lösungen für den Betrieb von IT-Infrastrukturen sind beispielsweise [Check\_MK](https://checkmk.com/) oder [Prometheus](https://prometheus.io/). Diese Anwendungen bieten eine kontinuierliche Überwachung von Systemen anhand definierter Metriken und warnen die Administratoren aktiv, wenn definierte Werte bestimmte Grenzen überschreiben.

Die Nutzung einer Monitoring-Lösung wird umso relevanter, je mehr Betriebsverantwortung für das BMS bei der Einrichtung liegt. Beim Cloud-BMS liegen zwar viele der Verantwortungen beim Betreiber der Software, trotzdem sollte zumindest die reine Verfügbarkeit des Systems auch von der nutzenden Einrichtung überwacht werden.

#### Notfallbetrieb

Bzgl. der Themen Support, Wartung & IT-Sicherheit, als auch Fehlersuche und -vorbeugung, unterscheiden sich die Aufwände für die Einrichtung je nach gewähltem Betriebsmodell erheblich. Jedoch weichen die zu nutzenden Prinzipien bei diesen Themen nicht grundlegend ab zu anderen zu wartenden Systemen in der IT-Welt.

Dazu gehören Maßnahmen zur Aufrechterhaltung des Bibliotheksbetriebs im Notfallbetrieb. Dies kann ein temporärer Offlinebetrieb des Systems sein. In diesem Fall werden die Prozesse mit den Daten abgewickelt, die zum Zeitpunkt des Offline-Gangs im System vorhanden waren. Wenn das System wieder online geht, muss gewährleistet werden, dass Änderungen an den Daten aus der Offlinezeit nachvollzogen werden (Beispiele: Ausleihen, Erwerbungen, Rechnungsbearbeitung, Benutzerdatenänderungen). Im Idealfall erledigt das die genutzte Komponente oder das BMS selbst.

Bei lokalen Installationen sollte man je nach Größe der Einrichtung ebenfalls über ein Spiegelsystem des BMS nachdenken. Dieses wird parallel auf dem aktuellen Stand gehalten und kann einspringen, wenn das laufende BMS ausfällt.

Es empfiehlt sich in jedem Fall neben dem Einsatz eines Produktivsystems mindestens eine Test-Instanz und ggf. eine oder mehrere Entwicklungs-Instanzen des BMS zu betreiben. So können neue Funktionen schneller umgesetzt werden ohne den laufenden Betrieb durch unerwartete Fehler zu gefährden.

Zu beachten ist weiterhin, die Nutzer\*innen des BMS (intern als auch extern) bei Problemen zu informieren. Dabei sind vor allem von Bedeutung, welche Interaktionen nicht mehr möglich sind, ob es alternative Möglichkeiten für die Nutzer\*innen gibt und wann das System voraussichtlich wieder zur Verfügung steht.

### IT-Sicherheit

Um ein BMS vor den zunehmenden Angriffen durch böswillige Akteure (Hacking, Malware, Ransomware) abzusichern, können die folgenden Empfehlungen als Grundlage dienen (Breeding, Marshall 2022):

* Die Infrastruktur um das BMS herum sollte durch starke Sicherheitsvorkehrungen getragen werden.
* Die Gefahr kurzfristig entstehender Sicherheitslücken sollte nicht unterschätzt werden.
* Cloud-basierte Systeme sollten aktiv überwacht und der Überblick behalten werden.
* Anbieter sollten aufgefordert werden, die Konzepte ihrer Sicherheitsvorkehrungen offenzulegen.
* Gerade Administrator\*innen sollten ihre Zugänge gesondert absichern.
* Es sollte sichergestellt werden, dass jede Software stets auf dem aktuellen Stand ist, sowohl auf den Arbeitsplatz-PCs als auch den Servern.

Allgemein gilt auch immer der Grundsatz: „Bleiben Sie wachsam, in Bezug auf ungewöhnliche Ereignisse auf Ihren IT-Systemen“.

### Backup und Rollback

Für den Fall, dass der Betrieb eines BMS lokal erfolgt, ist es wichtig, dass sich die Einrichtung über Backup und Rollback der Software Gedanken macht. Da dies ein generelles Thema des Betriebs von IT-Systemen ist, wird im Folgenden auf die Spezifika für BMS eingegangen, und Themen wie das Backup von Servern lediglich angerissen.

Folgende Aspekte sollten im Rahmen von BMS besondere Beachtung finden:

* Definition von Backup-Zyklen: wie oft werden welche Daten in welchem Umfang auf welche Art gesichert? Es können hier durchaus verschiedene „Sicherungsaspekte“ mit unterschiedlichen Zyklen definiert werden.
* Definition des Umfangs der Sicherung. Sollen die Daten komplett gesichert werden, sollen nur Veränderungen gesichert werden? Wichtig für diese Entscheidung ist die Frage, wie schnell ein System wiederhergestellt werden kann/soll.
* Ein off-site-Backup sollte in die Überlegungen einbezogen werden, also ein kompletter Satz einer Sicherung, der außerhalb der Institution gelagert wird. Dabei ist der Datenschutz zu berücksichtigen, u.U. müssen die Sicherungsdaten daher verschlüsselt werden.
* Sicherung von Daten, die bei rechtlichen Fragen von Relevanz sein können: Bsp. Rechnungen, Ausleihen, Mahnungen.

Ganz allgemein ist die Frage zu klären, wer Verantwortung für die Einrichtung, Durchführung und die regelmäßige Kontrolle der Sicherungen hat. Der letzte Punkt meint hierbei einerseits das Monitoring der erfolgreichen regelmäßigen Ausführung von Sicherungen, aber auch der Test der erstellten Sicherungen, etwa durch periodisches Einspielen auf einer Testinstanz des BMS.

### Zusammenspiel Hard- und Software

Das BMS steht in der IT-Landschaft einer Bibliothek im Regelfall nicht alleine, sondern kommuniziert mit anderen Hard- und Softwaresystemen. Hierzu gehören beispielhaft (s.a. Kapitel [Automatisierung)](#automatisierung-und-selbstbedienung):

* Lesegeräte: (Barcode-)Scanner und Chip-Lesegeräte für Benutzungsausweise und/oder Medien
* Selbstverbucher für die Ausleihe und/oder Rückgabe
* Rückgabeautomaten, die ggf. auch eine automatische Vorsortierung von Medien übernehmen
* Sicherungsgates zur Detektion nicht entliehener Medien an Ein- und Ausgängen
* Drucker zur Erstellung von Quittungen, Ausweisen, Labeln, usw.

Für den Fall, dass die externen Systeme nicht lokal an einem Computer angeschlossen sind, sondern über das Netzwerk der Einrichtung angebunden sind, gibt es vielfach etablierte bibliothekarische Schnittstellen (APIs), etwa [SIP](https://en.wikipedia.org/wiki/Standard_Interchange_Protocol), oder man setzt auf moderne, allgemeine API-Standards wie [REST](https://de.wikipedia.org/wiki/Representational_State_Transfer).

### Datenschutz, User-Tracking, Analytics

Innerhalb der EU gilt seit 2018 die Datenschutz-Grundverordnung (DSGVO), nach der personenbezogene Daten grundsätzlich zu schützen sind.

Im Kontext eines BMS sind die anfallenden personenbezogenen Daten etwa:

* durch Nutzer\*innen bei der Anmeldung angegebenen Daten für den Bibliothekszugang
* mit dem Benutzerkonto verbundene Ausleihvorgänge und Mahnhistorien
* die Protokolle (Logs) über Online-Zugriffe auf das BMS (z.B. IP-Adresse, Seitenaufrufe)

Um den Schutz personenbezogener Daten gewährleisten zu können, gibt es verschiedene Ansätze:

* Verschlüsselung: Daten werden auf verschlüsselten Servern gespeichert, ebenso ist die Übertragung Ende-zu-Ende verschlüsselt
* Separierung: Personendaten werden getrennt von nicht-sensiblen Daten gehalten (siehe [IDM](#identity-management))
* Pseudonymisierung: Nutzer\*innen-Daten werden mit Pseudonymen präpariert, sodass sie nicht mehr oder nur unter großem Aufwand den einzelnen Personen zuzuordnen sind
* Anonymisierung: Daten werden derart verändert, dass sie nicht rückverfolgbar sind (z.B. Maskierung IP-Adressen)

Die Entscheidung zur Verschlüsselung und Separierung von Daten sollte bereits im Vorfeld des Betriebs eines BMS getroffen werden.

Die Pseudonymisierung und Anonymisierung kann auch im Laufe der Erhebung der personenbezogenen Daten zur Anwendung kommen, sofern bestimmte Daten nicht mehr für einen konkreten Zweck erforderlich sind.

Leider sind personenbezogene Daten für Bibliotheks-Statistiken oft notwendig (siehe [Kosten](#kosten)). In diesem Fall sollten ebenfalls pseudonymisierte oder anonymisierte Datensätze zur Grundlage genommen werden.

Wenn ein BMS durch einen externen Anbieter gehostet wird (siehe [Betriebsmodelle für serverbasierte Software](#X3a5ca0018f511a6e73c5bcce2da503f97a39d48)), muss Folgendes sichergestellt sein:

* Die Verschlüsselung der Datenübertragung (Ende-zu-Ende-Verschlüsselung)
* Betrieb und Steuerung der Server innerhalb der EU ([DSGVO](https://de.wikipedia.org/wiki/Datenschutz-Grundverordnung))
* Der Ausschluss von User-Tracking durch Ad-Tech (Werbe-Netzwerke)
* Der Abschluss eines [Datenverarbeitungsvertrags im Auftrag](https://de.wikipedia.org/wiki/Datenverarbeitung_im_Auftrag)

In der Kombination eines IDM mit einem cloudbasierten BMS außerhalb der EU wäre denkbar, die personenbezogene Daten dort in pseudonymisierter Form speichern zu lassen oder Personendaten von nicht sensiblen Daten zu trennen.

Für alle personenbezogenen und personenbeziehbaren Daten sind Lösch- oder Anonymisierungsfristen festzulegen. Die Anonymisierungsfristen ergeben sich aus den Vorgaben der DSGVO und müssen betrieblichen und rechtlichen Aspekten genügen. So ergeben sich Fristen für die Speicherung von Daten über Gebühren (Entstehung, Bezahlung, ...) aus den Landeshaushaltsordnungen oder anderen für die Einrichtung maßgeblichen Regelungen. Betriebliche Gründe für die Länge von Speicherfristen von personenbezogenen und personenbeziehbaren Daten können sich aus Fristen für Einsprüche ergeben.

Die über die vergangenen Jahrzehnte geschehenen sukzessiven Aufkäufe kleinerer BMS-Service-Provider durch einige wenige große kommerziellen Bibliotheksdienstleister hat ganze Firmenkonglomerate entstehen lassen, die inzwischen den Bibliotheksmarkt dominieren. Einige von ihnen, die Dienste für wissenschaftliche Bibliotheken anbieten, wandeln sich in den letzten Jahren zu Data-Analytics-Konzernen. In diesem Zuge präparieren sie ihre cloud-basierten BMS-Lösungen mit Trackern, die Verhaltensprofile über die Nutzer\*innen erstellen. Durch die ebenfalls seitens der Anbieter gestellten Zugangsauthentifizierungssysteme wird versucht, zusätzlich eine möglichst hohe Personalisierung bei der Erstellung einzelner Profile zu erreichen. Die dabei entstehenden Datenflüsse werden für gewöhnlich nicht transparent gemacht (Siems 2022). Der Einsatz solcher Analytics-Technologien unterminiert die Integrität konventioneller IDM-Systeme und tangiert somit nicht nur datenschutzrechtliche Belange, sondern auch die IT-Sicherheit. Idealerweise sollte bereits vor der Anschaffung einer BMS-Lösung abgeklärt werden, ob solche Analytics-Technologien eingesetzt werden. Im Zweifelsfall sollte immer der\*die lokale Datenschutzbeauftragte oder IT-Sicherheitsbeauftragte hinzugezogen werden.

## Zusammenfassung und Ausblick

Ein BMS ist im Normalfall kein statisches System - vielmehr muss es aufgrund der sich verändernden Bedürfnisse einer Bibliothek und deren Nutzer\*innen stetig angepasst werden.

Der Import, Export oder auch die Zusammenführung von Daten erfordert klar definierte Metadaten und Schnittstellen für den freien Austausch aus gut nachnutzbaren Quellsystemen. Dies ist vor allem erforderlich bei der aktuell stärkeren Entwicklung hin zu Open Data und öffentlicher Datennutzung. Die Integration und Interaktion mit anderen Informationssystemen nimmt also zu. Vor allem herkömmliche BMS der zweiten Generation kommen hier schnell an ihre Grenzen.

# Discovery & Retrieval

|  |
| --- |
| Zusammenfassung |
| Im Gegensatz zum klassischen Online Public Access Catalogue (OPAC) bieten Discovery-Systeme mehr Inhalte, Funktionen und in der Regel auch modernere Nutzungsoberflächen. In diesem Kapitel werden typische [Komponenten](#komponenten) und [Funktionen](#funktionen) von Discovery-Systemen vorgestellt sowie Hinweise zum Aufbau und [Betrieb](#betrieb) gegeben. |

## Einleitung

Als Discovery-Systeme werden Rechercheplattformen bezeichnet, die möglichst alle Dienste einer Bibliothek über einen einheitlichen Zugang nutzbar machen. Insbesondere beschränken sich die recherchierbaren Medien nicht nur auf den lokalen Bestand. Die Benutzung und der Funktionsumfang orientieren sich dabei an gängigen Suchmaschinen und Verzeichnissen im Web.

## Ursprung und Motivation von Discovery-Systemen

Die Entstehung von Discovery-Systemen zu Beginn der 2000er Jahre hatte mehrere Gründe: Bibliothekarische Recherchesysteme spielten im Informationsverhalten insbesondere von studentischen Nutzer\*innen nur noch eine untergeordnete Rolle. Parallel zeichnete sich ab, dass die dritte Generation der Bibliotheksmanagementsysteme bezüglich ihrer OPAC-Module stagnierte, vornehmlich in Bezug auf das Design, aber auch hinsichtlich ihrer Funktionalitäten. Außerdem wurde Suchmaschinen-Technologie als Open Source-Software verfügbar, so dass technisch aufgeschlossene Einrichtungen eigene Experimente mit der Indexierung bibliographischer Daten begannen.

Zum gegenwärtigen Zeitpunkt sind Discovery-Systeme in wissenschaftlichen und zunehmend auch in öffentlichen Bibliotheken verbreitet. Es gibt eine Reihe von Produkten kommerzieller Anbieter\*innen und einige Open Source-Projekte. Discovery-Systeme können von Bibliotheken selbst oder durch Hosting-Anbieter wie Verbundzentralen, Hersteller\*innen und kommerzielle Dienstleister\*innen betrieben werden. Die Hersteller kommerzieller Bibliotheksmanagementsysteme der neueren Generation bieten Discovery-Systeme an, die besonders gut mit dem [BMS](#bibliotheksmanagementsysteme) der gleichen Hersteller\*innen zusammenarbeiten.

Wenn Bibliotheken neben dem Bestandskatalog andere Repositorien betreiben (Dokumenten-Server, Digitalisate-Server, Forschungsdaten-Server, …) ist die Einführung eines Discovery-Systems eine Möglichkeit, diese Datenbestände gemeinsam zugänglich zu machen.

## Bestandteile von Discovery-Systemen

### Komponenten

Ein Discovery-System umfasst verschiedene Komponenten. Dazu gehören

* eine Benutzungs- oder Rechercheoberfläche ([Frontend](#frontend)),
* der [Suchindex](#suchindex) (ein oder mehrere Quell-Indizes) sowie die operativen Aspekte
* [ETL-Prozesse](#etl-prozess) und
* die Konfiguration der Rechercheoberfläche.

Angebunden ist häufig auch eine Komponente zur [Authentifizierung und Autorisierung](#authentifizierung-und-autorisierung).

#### Frontend

Die Rechercheoberfläche (User Interface) umfasst typischerweise eine Startseite, eine einfache und eine erweiterte Suche, eine Trefferliste mit Facetten sowie eine Detailseite. Mitunter sind auf der Startseite auch thematische Sucheinstiege verfügbar, z.B. ein Browsing über eine Klassifikation oder Sammlungen. Außerdem gibt es meistens einen persönlichen Bereich, in dem auf das eigene Bibliothekskonto im BMS zugegriffen und gespeicherte Suchanfragen und Literaturlisten verwaltet werden können.

Die Gestaltungsmöglichkeiten für Design und Layout des User Interface reichen von einer einfachen optischen Anpassung bei Schriften, Farben und Logos bis hin zu größeren Veränderungen im Seitenaufbau, je nachdem, welcher Art das eingesetzte System ist (Eigenentwicklung auf Open Source-Basis, gehostetes kommerzielles System o.Ä.).

#### Suchindex

Zentraler Bestandteil eines Discovery-Systems sind auf Grundlage etablierter Suchmaschinentechnologie wie *Apache Solr* und *Elasticsearch* entwickelte **Suchindizes**.

Der Index eines Discovery-Systems enthält Metadaten und ggf. damit verknüpfte Daten wie Volltexte, Inhaltsverzeichnisse und Übersetzungen. Im Suchindex kommen also Daten unterschiedlicher Art und Herkunft zusammen. Die Daten können in einzelnen Kollektionen aufbereitet sein, z.B. nach Bestandsdaten von (Teil-)Bibliotheken oder Verbünden, Daten einzelner Verlage, Metadaten aus Repositorien etc. Es besteht die Möglichkeit, den Suchraum des Discovery-Systems individuell zu konfigurieren. Der Aufbau eines Index in Eigenregie ist bei entsprechenden Prozesskenntnissen und personellen Kapazitäten möglich und schafft Freiheiten zur Berücksichtigung eigener Datenkollektionen.

Einige Discovery-Systeme können Suchanfragen gleichzeitig an mehrere Suchindizes senden und die Treffer aus den unterschiedlichen Suchindizes in einer Gesamtliste zusammenführen. Dies setzt allerdings eine Koordination der genutzten Suchindizes voraus. Dieser Aufbau ermöglicht es den Bibliotheken auch, verschiedene Datenquellen in ihrem Discovery-System gemeinsam zugänglich zu machen. Teils werden die Quellen selbst ausgewertet (z.B. Harvesting der Daten des eigenen Katalogs, relevanter Repositorien, …), teils werden dafür andere freie oder kommerzielle Suchindizes (K10plus-Zentral, Gemeinsamer Verbünde Index, EBSCO-Discovery-Index, ExLibris Central Discovery Index,…) genutzt.

Die Daten, die in Suchindizes aufgenommen werden sollen, werden im Rahmen eines [ETL-Prozess](#etl-prozess) aus verschiedenen Datenquellen (Kataloge/BMS, Repositorien, bibliographische Fachdatenbanken, …) gesammelt, konvertiert und dann in den Suchindex geladen. Für jede Quelle muss dieser Prozess entsprechend eingerichtet und für Aktualisierungen regelmäßig ausgeführt werden.

Bei den ETL-Prozessen werden die Daten aus den verschiedenen Datenquellen transformiert. Dabei werden z.B. die MARC-Struktur mit Feldern, Indikatoren und Unterfeldern in eine einfachere Feldstruktur überführt. Die Daten werden entsprechend den verschiedenen Such- und Navigationsbedürfnissen in unterschiedliche Indexfelder überführt. Ein Datum (z.B. Name des\*der Autor\*in) kann für verschiedene Suchtypen unterschiedlich aufbereitet und mehrfach im Index gespeichert werden. Bei der Aufbereitung werden die Daten analysiert und etablierte Verfahren zur Relevanzberechnung für die Sortierung innerhalb der Trefferliste eingesetzt.

Für die Bildung von Facetten aus den Einträgen der Trefferliste werden spezielle Daten ermittelt. Diese Facetten-Daten sind technisch gesehen Suchbegriffe und dienen der nachträglichen Verfeinerung der Trefferliste. Beispiele für Facetten, die zur Einschränkung genutzt werden, sind Namen von Autor\*innen, Schlagwörter, Medienarten, Standorte physischer Medien oder auch Kennzeichnung von Open Access-Material und Vieles mehr).

### Funktionen

Ein Discovery-System ist mehr als ein reines Nachweissystem. Der Funktionsumfang umfasst daher auch mehr als die reine Recherche. Der Anspruch an ein Discovery-System, alle Informationen zu Medien an einer Stelle zu bündeln, sollte prinzipiell auch alle Dienstleistungen zu diesen Medien umfassen. Daher sollten auch Informationen zur Bereitstellung von Literatur enthalten, weitere Dienste integriert und eine Personalisierung möglich sein.

#### Recherche

Hauptfunktionen des Discovery-Systems sind die Recherche, die Anzeige von Metadaten und die Hinführung zur Nutzung der Medien. Im Einzelnen geht es um folgende Punkte:

* einfache Suche ohne Spezifizierung eines Suchfeldes
* Suche in Feldern der Metadaten (Titel, ISBN, Schlagworte)
* erweiterte Suche mit Möglichkeiten der Verknüpfung von Suchen in verschiedenen Feldern
* Navigation in Trefferlisten über Facetten und Sortierung
* Detailanzeige einzelner Treffer
* Export von Literaturangaben

Die Suche in Discovery-Systemen nutzt in der Regel verschiedene Funktionen der Suchmaschinentechnologie, um einen eingegebenen Suchbegriff gegen den Index abzuprüfen. Daher liefern Discovery-Systeme mit dem Suchparadigma „beste Treffer“ statt „exakte Treffer“ mehr Treffer als Bibliothekskataloge (Steilen 2012). Sie nutzen außerdem Algorithmen für die Relevanzsortierung (*Ranking*), um die Trefferlisten möglichst nutzungsorientiert aufzubereiten. Die Sortierungsalgorithmen sorgen bei Übereinstimmungen von Suchbegriff und Indexeintrag in definierten Feldern (Titel, Schlagwort, …) für eine Bevorzugung. Anders als bei Web-Suchmaschinen gehen Popularitätsdaten wie die Anzahl von Ausleihen, Aufrufen und Zitationen bislang nicht in das Ranking ein.

Zu den Funktionalitäten für die Suchunterstützung gehören auch die Autovervollständigung sowie die Vorschlagsfunktion von Suchbegriffen. In beiden Fällen wird der Suchindex in Echtzeit geprüft. Es gehört zu den zentralen Zielen von Discovery-Systemen, Null-Treffer-Meldungen zu vermeiden.

Facetten sind ebenfalls eine für Suchmaschinen typische Funktion und dienen der Eingrenzung von Treffermengen. Hierfür werden einzelne Metadatenfelder wie Schlagwörter, Namen von Verfasser\*innen oder Dokumenttypen in Bezug auf eine Suchanfrage ausgewertet und nach Vorkommenshäufigkeit sortiert. Den Facetten wird eine wichtige Rolle beim entdeckenden Suchen zugesprochen. Zur Präsentation der Facetten in der Rechercheoberfläche gibt es verschiedene Möglichkeiten (siehe [Abbildung 7.1](#fig-vufind)). Die Auswahl der angebotenen Facetten muss jedoch gut vorbereitet werden. Fehlen die entsprechenden Metadaten bei bestimmten Titeln, können durch Facettierung auch Treffer verloren gehen.

Die Weiterverwendung von Literaturangaben wird durch verschiedene Exportmöglichkeiten unterstützt. In der Regel lassen sich Angaben per Mail verschicken, ausdrucken oder in unterschiedlichen Formaten und Zitierstilen herunterladen.

|  |
| --- |
| Abbildung 7.1: Beispiel eines Rechercheergebnis in einem Discovery-Interface ([Quelle](https://core.coll.mpg.de/)) |

#### Bereitstellungsdienste

Die Evaluationen früher Discovery-Systeme haben bereits gezeigt, dass Informationen darüber, ob und wie ein gefundenes Medium zugänglich ist, von zentraler Bedeutung sind. Diese Bereitstellungsdienste, auch Delivery-Funktionen genannt, umfassen für physische und digitale Medien jeweils unterschiedliche Aspekte.

Bereitstellungsdienste für physische Medien:

* Nachweise von Standorten, Ausleihbarkeit und aktuellem Ausleihstatus
* Verlinkung zu Verbundkatalogen mit Fernleihmöglichkeiten
* Verlinkung zu Fernleihe und Dokumentlieferdiensten
* Möglichkeit zur Anfrage nach einer Digitalisierung oder Bereitstellung in einem Semesterapparat
* Möglichkeit zur Abgabe eines Anschaffungsvorschlags

Bereitstellungsdienste für digitale Medien:

* idealerweise eine auf das jeweilige Nutzungsszenario angepasste Zugangs-URL
* weitere Zugangs-URLs
* Hinweise zur Nutzung elektronischer Medien, z.B. zur Zugänglichkeit über VPN, notwendigen Readern, DRM etc.

Die Verfügbarkeit und Entleihbarkeit von physischen Medien, die der Bibliothek gehören, werden über eine sogenannte Verfügbarkeitsrecherche, die das Discovery-System im Hintergrund ausführt, ermittelt und angezeigt. Diese Abfragen werden mittels [Schnittstellen](#schnittstellen) zu den Ausleihmodulen der Bibliotheksmanagementsysteme durchgeführt. Diese Schnittstellen können proprietär oder offen sein. Beispiele für Hersteller-unabhängige Schnittstellen sind die *Patrons Account Information API* (PAIA) als offene Schnittstelle und das *Session Initiation Protocol* (SIP2) als intern genutzter Standard oder das NISO Circulation Interchange Protocol (*NCIP*). Verschiedene Discovery-Systeme unterstützen diese oder andere Schnittstellen zum Ausleihsystem in Form von sogenannten Treibern – beispielsweise unterstützt VuFind die Anbindung an FOLIO durch einen eigenen FOLIO-Treiber.

Bei den digitalen Medien ist die größte Herausforderung, den jeweils besten von in der Regel mehreren Zugangslinks für ein Medium zu identifizieren und zur Anzeige zu bringen. Zur Ermittlung des besten Zugangslinks sind in der Regel mehrere Prüfschritte erforderlich. Idealerweise sind solche Prüfschritte konfigurierbar, allerdings ist diese Funktion oftmals kein integraler Bestandteil von Discovery-Systemen, sondern ein eigener Dienst. Ein Beispiel für einen solchen separaten Dienst ist der Webdienst DAIA+ (Keßler 2018). Eine andere Möglichkeit ist der Einsatz sogenannter Link Resolver. Beim Link Resolving wird über die Metadaten ein Hyperlink zu Diensten der Bibliothek ermittelt. Es wird vorrangig bei der Ermittlung von Diensten für Metadaten zu solchen Medien genutzt, die nicht aus dem BMS der Bibliothek und E-Ressourcen stammen. Ein Verfahren für das Link-Resolving ist die *Open-URL* (NISO-Standard [*Z39.88*](https://www.niso.org/publications/z3988-2004-r2010)).

### Anreicherungsdienste

Die Ergänzung von bibliotheksseitig erstellten Metadaten mit weiteren Informationen gibt es bereits in den klassischen OPACs. Beispiele sind gescannte Inhaltsverzeichnisse, Links auf Wikipedia-Artikel oder die Integration von Buchcovern.

Zu den am häufigsten genutzten Anreicherungsdiensten gehören:

* Cover-Anzeigen
* kontextabhängige Infoboxen mit Informationen aus Nachschlagewerken, z.B. Autor\*innenenportraits z.B. via *Wikidata* und *GND,* Informationen aus Nachschlagewerken wie Munzinger
* Empfehlungsdienste mit Hinweisen auf Literatur zum selben Thema (z.B. *BibTip*, *bX*)
* Visualisierungen von Buchstandorten über Gebäudeinformationssysteme (z.B. *Mapongo, V:Scout*)
* Integration mit weiteren Diensten, z.B. der Leseförderungs-App *Antolin*

Grundsätzlich erlaubt die Systemarchitektur von Discovery-Systemen die Integration von diesen und anderen Diensten über einschlägige [Schnittstellen](#schnittstellen), so dass sich über die gelisteten Dienste noch zahlreiche weitere Möglichkeiten ergeben.

#### Personalisierung

Discovery-Systeme erlauben in der Regel eine Anmeldung in einem persönlichen Bereich, der folgende Funktionalitäten umfassen kann:

* Einsicht in das Bibliothekskonto einschließlich der Möglichkeit zum Vormerken und Verlängern
* Speicherung von Suchanfragen
* Speicherung von Literaturlisten
* Alerting-Dienste

Literaturlisten können alternativ dazu auch sitzungsbasiert gespeichert werden. Dauerhaft gespeicherte Listen lassen sich auch veröffentlichen und damit allgemein zugänglich machen, was auch die Präsentation von Auswahllisten oder Semesterapparaten erlaubt.

Häufig können auch Suchanfragen gespeichert werden. Die Einrichtung von Alerting-Diensten hilft den Nutzer\*innen, sich mit wenig Aufwand über neue Titel informieren zu lassen. Alerting-Dienste beinhalten das regelmäßige (automatisierte) Absetzen einer Suchanfrage und das Versenden von Information, wenn die Suchanfrage veränderte Trefferlisten (in der Regel: neue Titel) liefert.

#### Thematische Sucheinstiege

Wie beschrieben bieten Trefferlisten mit Facetten und Empfehlungen zwar durchaus auch die Möglichkeit, sich eine Treffermenge zu erschließen. Allerdings fehlt Discovery-Systemen genau wie OPACs häufig die Möglichkeit, eine systematische Suche durchzuführen. Teilweise wird ein Browsing durch die klassifikatorische Inhaltserschließung angeboten, jedoch fehlen vielen Datensätzen entsprechende Daten und das Browsing bezieht sich jeweils nur auf Teilmenge des Suchraums.

Aus diesem Grund werden derzeit verschiedene Ansätze erprobt, um eine thematische Suche zu ermöglichen. Hierzu zählen u.a. folgende Projekte und Dienste:

* ein Nachbau der Browsing-Funktion an physischen Bücherregalen, z. B. bei dem kommerziellen Dienst [Blended Shelf](https://www.eikekleiner.de/blended-shelf/)
* die Nutzung von Normdaten zur Erstellung von Übersichtsseiten, z. B. im [Katalog des Deutschen Literaturarchivs Marbach](https://www.dla-marbach.de/katalog-beta/)
* die Visualisierung von Treffermengen und den darin enthaltenen Zusammenhängen, wie zum Beispiel bei [Open Knowledge Maps](https://openknowledgemaps.org/), in einer [prototypischen Installation der SLUB Dresden](https://data.slub-dresden.de/) oder mit dem kommerziellen Dienst *Yewno*.

Diese Projekte und Dienste sind jedoch entweder noch relativ neu oder wenig verbreitet und nicht oder nur mit Aufwänden nachnutzbar. Im Rahmen einer strategischen Planung für den Einsatz eines Discovery-Systems muss daher abgewogen werden, ob und wie ein thematischer Sucheinstieg umgesetzt werden soll, zumal für eine Darstellung im Sinne einer optimalen User Experience jeweils auch erhebliche Design-Aufwände entstehen.

## Aufbau und Betrieb eines Discovery-Systems

### Betriebsmodelle

Der Betrieb eines Discovery-Systems stellt vergleichbare Anforderungen und unterliegt ähnlichen Rahmenbedingungen wie beim [Betrieb eines Bibliotheksmanagementsystems](#technischer-betrieb).

Im Inhouse-Betrieb werden alle Komponenten selbst durch die Bibliothek betrieben und damit sind hier die weitestgehenden Anpassungen möglich. Dies wird meist nur bei kleinen oder sehr speziellen Datenbeständen (z. B. durch die Fachinformationsdienste) oder durch sehr große Einrichtungen gemacht. Oft trifft man auch hybride Lösungen, in denen neben einem vergleichsweise kleinen eigenen Index ein kommerzieller oder nicht-kommerzieller Index genutzt wird.

In einem Hosting-Betrieb wird die gesamte Infrastruktur durch eine\*n Dienstleisterbereitgestellt. Dabei erfolgt die Indexierung in der Regel durch einheitliche Indexierungsverfahren, die von allen teilnehmenden Bibliotheken gemeinsam genutzt werden. Bei diesen Lösungen werden alle Daten in einen einheitlich aufgebauten, in einer Cloud gehosteten Index eingespielt, die Frontends sind nur eingeschränkt individualisierbar und lassen sich ausschließlich durch Konfigurationen parametrisieren. Zusatzfunktionen lassen sich über [Schnittstellen](#schnittstellen) anbinden. Wesentlicher Vorteil dieser Systeme ist ein vergleichsweise geringer Wartungsaufwand, ihre gute Skalierbarkeit und durch standardisierte Workflows und ihre hohe Betriebssicherheit. Als Hoster\*innen von Discovery-Systemen treten Bibliotheken, Verbünde und kommerzielle Anbieter auf.

Ein Spezialfall des Hostings ist die Nutzung von Cloud-Services externer Anbieter\*innen für den Betrieb von BMS und Discovery-Systemen. Mehrere Hersteller\*innen von BMS und Discovery-Systemen sind gleichzeitig Betreiber\*innen von solchen Cloud-Lösungen. In diesen Fällen wird die Software (BMS, Discovery-System) nicht mehr lizenziert, sondern über eine jährliche Pauschale Nutzung, Update und Betrieb des jeweiligen Software-Systems abgegolten.

Beim Hosting oder bei der Nutzung von Software, die in der Cloud betrieben wird, spricht man von einer „[Datenverarbeitung im Auftrag](https://de.wikipedia.org/wiki/Datenverarbeitung_im_Auftrag)“. Die Verantwortung für Datenschutz und Datensicherheit bleibt bei der Bibliothek als Auftraggeberin.

### Marktsituation

Die ersten Discovery-Systeme haben Bibliotheken selbst entwickelt, im deutschsprachigen Raum z.B. die *E-LIB* an der Staats- und Universitätsbibliothek Bremen oder das *beluga*-System an der Staats- und Universitätsbibliothek Hamburg. Seit Ende der 00er Jahre gibt es auch kommerzielle Systeme am Markt, entweder als Teil von Bibliotheksmanagementsystemen der neuesten Generation oder auch als individuell lizenzierbare Systeme. Die Open Source-Lösung *VuFind* ermöglicht es, verschiedene Suchindizes unter einer Oberfläche nutzbar zu machen, so dass es eine relativ große Vielfalt von Nutzungsszenarien gibt.

#### Kommerzielle Komplettsysteme

Im Wesentlichen gibt es zwei vergleichbare Anbieter\*innen von Komplettsystemen für Discovery-Systeme

* ExLibris mit *Primo* und *Summon*
* EBSCO mit *Ebsco Discovery-Service*

Diese Systeme bieten eine fertige Lösung, in die lokale Bestandsdaten und weitere lokale Metadaten integriert werden können. Es fallen jährliche Lizenzgebühren sowie einmalige Implementierungskosten an. Beide genannten Systeme sind weit verbreitet. Diese Systeme sind ausschließlich über die Cloud der jeweiligen Hersteller\*innen nutzbar. Diese sorgen für eine hohe Verfügbarkeit und regelmäßige Softwarepflege. Individuell zu prüfen sind vor einem Einsatz folgende Fragen:

* Einbindung von Verfügbarkeitsinformationen
* Datenschutzrechtliche Fragen (Ort des Hostings, Verfahrensbeschreibungen)
* Datenhoheit

Die Indizes dieser Systeme können separat lizenziert und beispielsweise an VuFind-Systeme angebunden werden.

Ein weiteres kommerzielles Discovery-System ist *WorldCat Discovery*, das allerdings die Nutzung von WorldCat als Suchindex voraussetzt.

#### Open Source-Systeme

Unter den von Bibliotheken selbst entwickelten Discovery-Systemen sind international VuFind und Blacklight am weitesten verbreitet.

[VuFind](https://vufind.org/) lässt sich an verschiedene kommerzielle und frei verfügbare Komponenten wie Indizes und Bibliotheksmanagementsysteme anbinden. In den deutschsprachigen Ländern besteht eine lebendige Anwender\*innengemeinschaft, die sich regelmäßig trifft. Mit [Qcovery](https://www.qcovery.de/) und [finc](https://finc.info/) gibt es zwei Sub-Communities für wissenschaftliche Bibliotheken, die sich die Aufgaben der Pflege und Weiterentwicklung der Software unter sich aufteilen. Die Software basiert auf PHP.

[Blacklight](http://projectblacklight.org/) ist hauptsächlich im angloamerikanischen Raum verbreitet, aber auch bei [Europeana](https://www.europeana.eu/) Einsatz. Die Software basiert auf Ruby on Rails.

Das von der VZG entwickelte System [Lukida](https://www.lukida.org/) spielt vor allem im Rahmen des Index K10plus-Zentral eine Rolle und wird primär als SaaS angeboten.

#### Indizes

Neben den kommerziellen Anbieter\*innen bieten im Bereich wissenschaftlicher Bibliotheken einige Verbundzentralen auf Suchmaschinen-Technologie basierende Indizes an, teilweise für die teilnehmenden Bibliotheken, teilweise auch darüber hinaus für die nicht-kommerzielle Nutzung. Diese frei verfügbaren Indizes sind für Bibliotheken, die ihre Bestandsdaten an einen Verbund liefern, eine hervorragende Möglichkeit, um relativ kostengünstig an ein Discovery-System zu kommen, da die Erstellung eines eigenen Index mit hohen Investitionen verbunden ist. Metadaten-Kollektionen enthalten der ALBERT-Index des Kooperativen Bibliotheksverbundes Berlin-Brandenburg sowie der Gemeinsame Verbündeindex für Bestandsdaten aus allen wissenschaftlichen sowie vielen Spezial- und öffentlichen Bibliotheken.

### Auswahl- und Entscheidungsprozesse

Sofern ein Discovery-System nicht Teil des BMS ist, ist die Einführung immer mit beträchtlichen Aufwänden verbunden, die aus initialen Kosten für die Implementierung und laufenden Kosten für die Pflege bestehen. Diese Kosten fallen unabhängig davon an, ob es sich um ein kommerzielles oder ein Open Source-System handelt. Sie richten sich nach unterschiedlichen Kriterien und dürften im Bereich der initialen Kosten im höheren vierstelligen Bereich liegen. Grundsätzlich sind die Entscheidungsprozesse bei Auswahlentscheidungen mit denen für ein Bibliotheksmanagement-System vergleichbar (vgl. Abschnitt [Marktanalyse und Beschaffung](#marktanalyse-und-beschaffung)).

Allerdings müssen die strategischen Vorteile eines Discovery-Systems sehr deutlich und auf den lokalen Bedarf hin herausgearbeitet werden. Es hat sich als hilfreich erwiesen, dass Bibliotheken klar definieren, an welche Zielgruppen sich ein Discovery-System richtet und welche Aufgaben es erfüllen soll. So könnte man beispielsweise argumentieren, dass eine Suche nach Signaturen, die aufgrund der häufigen Komplexität der entsprechenden Systeme in der Regel recht aufwändig zu implementieren ist, in einem Discovery-System nicht nötig ist, weil diese Suche ohnehin hauptsächlich vom Bibliothekspersonal durchgeführt wird. Das Bibliothekspersonal kann diese Suchen im BMS durchführen. Es sollte auch geklärt werden, ob der klassische OPAC nach Einführung eines Discovery-Systems überhaupt weiter angeboten werden soll.

Auch der Zuschnitt der Suchräume sollte genau bedacht werden, vor allem, wenn über lokale Bestandsdaten hinaus eigene Metadatenkollektionen (z.B. aus institutionellen Repositorien) integriert und durch eigene Suchfilter angesprochen werden sollen. Generell kann davon ausgegangen werden, dass auf die initiale Implementierung eines Discovery-Systems eine längere, oft mehrjährige Phase der Optimierung folgt, die idealerweise konsequent auf die Usability und User Experience der Hauptzielgruppen ausgerichtet ist (vgl. Kapitel [Anforderungen an die IT-Entwicklung](#anforderungen)).

Die grundsätzliche Entscheidung für ein Discovery-System beinhaltet auch einen Wechsel der Suchparadigmen. Die Einführung eines Discovery-Systems kann nur dann sinnvoll erfolgen, wenn die Abkehr der Dualität von Bestandsverzeichnis und Bibliographie sowie den traditionellen Suchparadigmen strategisch erwünscht ist und von entsprechenden Schulungen für das Bibliothekspersonal begleitet wird.

Wenn ein Discovery-System im Hosting genutzt werden soll, relativieren sich die oben gemachten Aussagen zur Flexibilität, die Open Source-Systeme bieten, da die Hoster\*innen in diesem Fall die Möglichkeiten festlegen, die durch die Bibliotheken genutzt werden können. Ebenso verschieben sich die gemachten Aussagen zur Verantwortung für Betriebssicherheit und Verfügbarkeit.

### Monitoring und Weiterentwicklung

Wie jedes IT-System brauchen auch Discovery-Systeme kontinuierliches technisches Monitoring (vgl. Kapitel Management von IT-Diensten), aber auch konzeptionelle Betreuung. Anders als der klassische Bibliothekskatalog sind Discovery-Systeme angetreten, um sich konsequent nach dem Informationsverhalten der Nutzer\*innen zu richten. Daraus ergibt sich, dass sowohl die Implementierung als auch die weitere Entwicklung möglichst kleinschrittig und unter Einbeziehung von Analysen der Nutzung erfolgen sollten. Neben den klassischen Methoden der Usability-Forschung (siehe Kapitel [Wie beziehen wir unsere Nutzer\*innen ein?](#einbeziehung)) bietet sich als niedrigschwellige Methode vor allem die Analyse von Logfiles an. Mit der Software [Matomo](https://matomo.org/) kann, auch unter Berücksichtigung von datenschutzrechtlichen Vorschriften, ermittelt werden, welche Anfragen an ein System gestellt werden.

## Vergleich mit klassischen Bibliothekskatalogen

Da Discovery-Systeme die Metadaten und Volltexte anders als die klassischen OPACs aufbereiten, sind Suchstrategien und -ergebnisse in beiden Systemen unterschiedlich.

Discovery-Systeme richten sich in der Regel an Benutzer\*innen, die den Umgang mit bibliographischen Recherchesystemen wie Katalogen und Fachbibliographien nicht gewohnt sind und die mit den Nutzungsmustern bedient werden sollen, die sie auch aus dem Web gewohnt sind.

Neben der Recherche nach bibliographischen Informationen sollen Discovery-Systeme auch den Zugriff bzw. die Bereitstellung von Medien unterstützen. Dieser auch als **Delivery** bezeichnete Prozess hat sich bereits in der frühen Phase der Discovery-Systeme als zentrales Element aus Sicht der Nutzer\*innen herausgestellt. Die Anbindung an Ausleihsysteme und Link Resolver ist daher ein wichtiges Qualitätskriterium.

Neuere BMS wie FOLIO und *Alma* enthalten zum Teil gar keinen klassischen OPAC mehr. Mit diesen Systemen muss daher immer ein zusätzliches Discovery-System eingesetzt werden.

Viele Bibliotheken erschließen **E-Ressourcen** nicht in vollem Umfang in ihrem BMS. Daher sind im OPAC, der zu diesem BMS gehört, die E-Ressourcen nicht oder nur eingeschränkt auffindbar. Wenn die Bibliothek ein Discovery-System betreibt, können Metadaten zu E-Ressourcen über einen [ETL-Prozess](#etl-prozess) in den Index des Discovery-Systems geladen werden. Voraussetzung dafür ist, dass den Metadaten mittels Electronic Ressource Management (ERM) entsprechende Nutzungslizenzen zugeordnet sind.

Tabelle 7.1: Vergleich typischer Eigenschaften von OPAC/Katalog und Discovery-System

|  | OPAC/Katalog | Discovery-System |
| --- | --- | --- |
| Suchraum | nur lokaler Bestand, nur selbständige Werke | lokaler Bestand, aber auch Verbunddaten, bibliographische Daten, Volltexte… |
| Suchprinzip | exakte Suche, feldbasierte Suche mit Boolescher Logik | best match/natürlichsprachige Suche |
| Suchunterstützung | eher wenig | Autovervollständigung, Suchvorschläge, Facetten |
| Sortierung | standardmäßig nach Aktualität | standardmäßig nach Relevanz |
| Mehrwertdienste | Buchcover, Listen, Exportformate | Buchcover, Listen, Stöbern/Entdecken |
| Metadatenmodell | bibliothekarisches Schema mit Hierarchien und Verweisen | „flache Version“ eines bibliothekarischen Schemas |

## Grenzen und Alternativen

Discovery-Systeme sind in der Regel nur einer von vielen Bausteinen in der Prozesskette der Recherche, Bewertung und Beschaffung von Literatur und spielen an unterschiedlichen Stellen eine Rolle. Sie helfen dabei, Literatur zu entdecken und Zugangswege zu ermitteln und brechen die traditionelle Grenze zwischen Katalog und Bibliografien durch einen zentralen Sucheinstieg auf. Trotz dieser Stärken können die Systeme nachgewiesene Medien nur begrenzt kontextualisieren und bewerten und bleiben in der Praxis oft hinter den Erwartungen zurück (Christensen 2022). Je nach Anwendung spielen daher alternative Systeme weiterhin eine Rolle:

* Komplexe bibliographische Angaben, zum Beispiel zum Erscheinungsverlauf von Zeitschriften oder mehrbändigen Werken, oder die Suche nach Signaturen lassen sich möglicherweise schneller über **herkömmliche bibliothekarische Instrumente** beziehungsweise Spezialdatenbanken wie die des BMS ermitteln.
* Zum Entdecken von Literatur eignen sich auch allgemeine Suchmaschinen oder spezielle **Academic Search Engines** wie Google Scholar sowie gänzlich andere Wege wie bestehende Literaturverzeichnisse, Empfehlungslisten auf Lernplattformen und Webshops.

Insbesondere Webshops haben im Vergleich zu Discovery-Systemen sehr personalisierte Such- und Empfehlungsdienste, die jedoch auf einer intensiven Auswertung des jeweiligen Nutzungsverhaltens basieren. Die Verwendung dieser Daten zur Personalisierung ist auch in Discovery-Systemen denkbar, wird aber aus Datenschutz- und Neutralitätsgründen grundsätzlich eher abgelehnt.

Eine vergleichsweise neue Herangehensweise insbesondere an das entdeckende Suchen bieten **Wissensgraphen** (knowledge graphs), die die vielfältigen Beziehungen zwischen Dokumenten und damit verknüpften Elementen darstellen und visualisieren. Die Anforderungen an die Qualität der so aufbereiteten Daten sind jedoch ungleich höher. Entsprechende Systeme existieren bereits in ausgewählten Bereichen, zum Beispiel die Plattform [SoNAR](https://sonar.fh-potsdam.de/) zur historischen Netzwerkanalyse. Ein ernstzunehmendes Beispiel für einen allgemeinen Wissensgraphen ist die Datenbank Wikidata mit ihren bibliographischen Inhalten *WikiCite* und dem dazu gehörigen Browsing-Interface [Scholia](https://scholia.toolforge.org/) (siehe Abbildung [Abbildung 7.2](#fig-scholia)).

|  |
| --- |
| Abbildung 7.2: Thematisches Netzwerk von [Publikationen in und über Scholia](https://scholia.toolforge.org/topic/Q45340488) |

Grundsätzlich gilt, dass die Grenzen zwischen Discovery-Systemen und Alternativen in der Praxis fließend sind und dass Discovery-Systeme perspektivisch um Funktionen anderer Systeme erweitert werden können und sollten.

# Digitalisierung

|  |
| --- |
| Wichtig |
| Dieses Kapitel befindet sich in einer ersten öffentliche Entwurfsversion. Feedback siehe Links im rechten Menü und [Hinweise zur Mitarbeit](#mitarbeit). |

|  |
| --- |
| Zusammenfassung |
| Neben der reinen Bewahrung von Kulturgütern stellt sich Bibliotheken und anderen Kultureinrichtungen die Herausforderung, diese weltweit im Netz zugänglich zu machen. Die Digitalisierung von Kulturgut erfordert verschiedene Prozesse und Tools, die in diesem Kapitel vorgestellt werden. |

## Grundlagen

### Begriffsbestimmung

**Digitalisierung** steht in diesem Kapitel für die Überführung von materiellen Kulturgütern in digitale Formate. Bibliotheken und andere kulturbewahrende Institutionen nehmen sich für die Digitalisierung sowohl Schriftgut (Bücher, Handschriften, Nachlässe) als auch audiovisuelle Medien und andere Objekttypen vor. Dabei wird oft auch von **Retrodigitalisierung** oder **Bestandsdigitalisierung** gesprochen, wenn die Objekte nicht originär digital sind („born-digital“), sondern bereits physische vorlagen und erst später digitalisiert werden.

### Geschichte

Als Vorläufer der Digitalisierung kann die Mikroverfilmung betrachtet werden, bei der bereits seit dem 19. Jahrhundert analoge Dokumente mit dem Ziel der Langzeiterhaltung, der Vervielfältigung oder als platzsparender Ersatz für Originale in ein anderes analoges Format transformiert wurden. Die Überführung in digitale Formate begann in den 1990er Jahren. Das Ziel war hier zunächst vor allem die weltweite Zugänglichmachung von besonderen Werken oder Beständen für die wissenschaftliche Nutzung. Waren es zunächst vor allem bedeutende Druckwerke des 16., 17. und 18. Jahrhunderts, die digitalisiert wurden, hat sich der Fokus der Digitalisierung auch zeitlich, thematisch und formal sehr erweitert.

In Deutschland war und ist die Deutsche Forschungsgemeinschaft (DFG) ein wesentlicher Akteur, der durch Projektförderung erhebliche Impulse für die großflächige Retrodigitalisierung gegeben hat. Gleichzeitig sorgten die in geförderten Projekten verbindlichen [DFG-Praxisregeln](https://zenodo.org/records/7435724), die auch in anderen Projekten oft angewendet werden, für ein hohes Maß an Einheitlichkeit und Interoperabilität zwischen verschiedenen Beständen. Neben technischen Parametern für den [Scan-Prozess](#scanner-und-scanverfahren) (Auflösung, Farbtiefe u.a.) und Speicherformaten (TIFF unkomprimierte Master-Images) werden dort Standards zur [Speicherung von Metadaten](#metadaten-und-schnittstellen) festgelegt, die von der Library of Congress entwickelt wurden und weiter gepflegt werden. Die DFG hat darüber hinaus mit dem [DFG-Viewer](https://dfg-viewer.de/) einen Minimalstandard in der [Objektpräsentation](#präsentieren) entwickeln lassen. Zu den Standards der DFG-Förderung gehört auch, dass jedes Werk nur einmal digitalisiert werden soll.

Eine weitere große Digitalisierungs-Initiative ging und geht von Google aus, die weltweit und in Deutschland insbesondere bei der Bayerischen Staatsbibliothek in großem Umfang alte Druckwerke digitalisieren (Google Books).

Ergänzend digitalisieren, veröffentlichen, erschließen und editieren private Akteure digitale Kulturgüter – Akteure in Wikimedia-, Open Data- und Open Science-Gemeinschaften, oder sie arbeiten ehrenamtlich in einschlägigen bibliothekarischen Projekten mit.

### Status quo

Zahlreiche Bibliotheken mit umfangreichen unikalen Beständen haben sich auf den Weg gemacht, Retrodigitalisierung als festes Arbeitsfeld in die eigene Organisation aufzunehmen. Für die Unterstützung der Prozesse und für die Präsentation der Ergebnisse im Web haben sich unterschiedliche Softwarelösungen etabliert, zu denen aktive Anwendergemeinschaften gehören. Zusammen mit den geklärten Standards ist ein Rahmen geschaffen, in dem diese Bibliotheken kontinuierlich neue Digitalisierungsprojekte starten. Größere Häuser öffnen zum Teil ihre Infrastruktur anderen kleineren Einrichtungen in Landesprogrammen wie z.B. in [Sachsen](https://www.slub-dresden.de/ueber-uns/projekte/landesdigitalisierungsprogramm) oder [Hamburg](https://www.sub.uni-hamburg.de/bibliotheken/projekte-der-stabi/hamburger-kulturgut-im-netz-hakin.html). In Berlin unterstützt die Arbeitsstelle [digiS](https://www.digis-berlin.de/digis/) mit einem jährlichen [Förderprogramm](https://www.digis-berlin.de/foerderprogramm/info/) seit 2012 Digitalisierungsprojekte in unterschiedlichen technischen Umgebungen - um nur einige Beispiele aus der heterogenen föderalen Struktur herauszugreifen.

### Ausblick

Neben der reinen bildhaften Digitalisierung wird von Forschenden zunehmend auch eine Erstellung und Bereitstellung von Volltexten erwartet. Diese kann entweder über einfache Text-Bild-Zuordnung oder auch mit aufwändiger semantischer Auszeichnung erfolgen. Die von der DFG geförderte Initiative OCR-D erstellt Werkzeuge und Infrastrukturen, um für die Werke des 16., 17. und 18. Jahrhunderts, aber auch für viele weitere Digitalisate eine effiziente und alltagstaugliche Volltext-Erstellung zu ermöglichen.

Außer Druckwerken und Handschriften werden zunehmend weitere Medien und Objekte digitalisiert. Das sind nicht nur einschlägige Textdokumente aus Bibliotheken und Archiven, sondern auch Objekte aus Sammlungen und Museen, die teilweise auch in 3D digitalisiert werden, und audiovisuelle Medien. Außerdem wird versucht, zusätzliche Attribute der materiellen Objekte zu erfassen (z.B. Strukturerkennung, Kontextinformationen).

Zudem werden sich GLAM-Einrichtungen auch immer stärker des Bedarfes bewusst, Digitalisate auch jenseits eigener Portale auffindbar und nutzbar zu machen (siehe [Digitalisierung als Beitrag zur Open GLAM-Bewegung](#Xabc6d7253650c5e5284c03076d215a642b91408)).

## Prozesse

Die Digitalisierung von Kulturgütern gliedert sich in mehrere Prozessschritte. Es haben sich dabei einrichtungsübergreifend ähnliche Prozessketten etabliert, zu deren Unterstützung [Workflowmanagementsysteme](#workflowmanagementsysteme) (WMS) entwickelt wurden und die Projektsteuerung erleichtern.

|  |
| --- |
| BILDUNTERSCHRIFT FEHLT |

Als Ausgangspunkt jedes Digitalisierungs-Vorhabens ist eine gründliche Planung der Arbeitsschritte unverzichtbar. Dabei sind die lokalen Rahmenbedingungen und die Infrastruktur der digitalisierenden Einrichtung ebenso zu berücksichtigen wie die konkreten Ziele, die zu erzeugenden Formate und Ergebnisse sowie die Verknüpfungen zu externen datenverarbeitenden Systemen. Die [DFG-Praxisregeln](https://doi.org/10.5281/zenodo.7435724) geben wichtige Hinweise, was bei den Prozessen beachtet werden sollte. WMS können diese Schritte steuern und teilweise auch automatisieren.

Für die Planung von Zeitungsdigitalisierung sollte im Vorwege besonders genau auf die Einzelschritte im Prozess geschaut werden oder kollegiale Beratung aus anderen Häusern eingeholt werden, um angesichts der Massen an Material den effizientesten Weg bei der Erschließung zu finden, evtl. Dienstleister optimal einzubinden und alle Möglichkeiten von Automatisierung auszuschöpfen.

### Vorbereiten: Auswahl von Objekten

Eine sinnvolle Auswahl zu digitalisierender Objekte hängt von Faktoren ab, deren Wichtigkeit gegeneinander abzuwägen ist. Dazu zählen beispielsweise

* die bestehenden Kapazitäten: verfügbare Scanverfahren, Arbeitsplätze, personelle Verfügbarkeiten),
* die technischen Ausstattung: Scanner, Speicherkapazität, Anbindung an eine Langzeitarchivierung
* zeitliche Faktoren: Projektlaufzeit, Förderbedingungen, Anlässe,
* Digitalisierungswürdigkeit: Einzigartigkeit, Nachfrage
* Erhaltungszustand: Dringlichkeit, Machbarkeit
* existierende Vorarbeiten: vorhandene Metadaten, Vervollständigung bereits angefangener Vorhaben

Die Prioritäten für die eigene Bestandsdigitalisierung sollten immer wieder überprüft werden, wenn sich die genannten Faktoren ändern.

Nach der Auswahl von zu digitalisierenden Objekten folgen weitere vorbereitende Schritte. Dazu gehören z.B.

* das Ausheben der Werke aus dem Bestand
* das Überprüfen und ggf. das Herstellen der Digitalisierungs-Tauglichkeit
* das Anlegen von Vorgängen im Workflowmanagementsystem
* das Erzeugen von Laufzetteln
* die Vorbereitung von Übergaben physischer Objekte an etwaige Scan-Dienstleister

### Digitalisieren

Die Entscheidung für das oder die am besten geeigneten Digitalisierungsverfahren ist individuell am Objekt zu treffen und hängt von verschiedenen Faktoren ab, wie der Art des Objekts, der gewünschten Genauigkeit und den verfügbaren Ressourcen.

#### 2D-Objekte

Für das Scannen von 2D-Objekten („Flachware“) hat sich als Zielformat der Bilddigitalisierung unkomprimiertes [TIFF](https://de.wikipedia.org/wiki/Tagged_Image_File_Format) etabliert. Es werden mit Aufsichtscannern oder fotografischen Reproständen Einzelseiten aufgenommen und vom Bildausschnitt her wird mit schwarzem Rand digitalisiert. Verschiedene Auflösungen („Derivate“) für die Webpräsentation werden nachgelagert erzeugt oder dynamisch von einem Bildserver zur Verfügung gestellt.

Für die verschiedenen Digitalisierungsvorhaben bilden die [DFG-Praxisregeln](https://doi.org/10.5281/zenodo.7435724) eine gute Orientierungshilfe - für die Auflösung, den Farbraum oder andere technische Parameter. Für die Beurteilung und Sicherstellung der objektiven Bildqualität haben sich zudem Digitalisierungsstandards etabliert, die mittlerweile auch von Scannerherstellern und Scandienstleistern berücksichtigt werden. Erwähnenswert sind hier vor allem [Metamorfoze](https://www.metamorfoze.nl/english/digitization) und der daraus hervorgegangene [ISO-19263 Standard](https://www.iso.org/standard/64220.html) .

Bei der Zeitungsdigitalisierung ist im Interesse der Forschung die Erzeugung von Volltexten nach der Digitalisierung ein wichtiges Ziel. Diese Perspektive bedingt im Hinblick auf die Qualität der OCR eine Präferenz für eine Digitalisierung vom Original gegenüber einer Digitalisierung vom Mikrofilm.

#### 3D-Objekte

Mittlerweile haben einige Bibliotheken angefangen, über flache, zweidimensionale Informationen hinaus auch die Tiefe und Mehrdimensionalität des kulturellen Erbes und Wissens innerhalb ihrer Bestände zu erfassen. Bibliotheksbestände sind häufig heterogen und ergeben sich aus einer fortgesetzten Sammlungsgeschichte, so dass auch spezielle dreidimensionale Objekte, wie Globen, Spiel- oder PopUp-Bücher oder historische Objekte des Buchdrucks wie Lettern, Teil eines Bestandes sein können. In diesen Fällen hat sich eine Digitalisierung in allen drei Dimensionen als ein entscheidendes Werkzeug erwiesen, um die Mehrdimensionalität in digitale Formate zu übertragen und um über Inhalte hinaus den Zustand und Details als Kulturerbe-Objekte wiederzugeben. Dies betrifft ebenso seltene Bücher und Manuskripte mit einzigartigen oder kunstvollen Bindungen. Um die drei Dimensionen angemessen abzubilden, sind [spezielle Erfassungsmethoden](#scanner-und-scanverfahren) notwendig.

#### AV-Medien

Bibliotheken und andere GLAM-Einrichtungen archivieren häufig auch analoge audiovisuelle Medien. Für die Digitalisierung der unterschiedlichen Quellen (z.B. Magnettonbänder, Film von 8mm bis 35mm, Selbstschnittplatten, Videobänder) werden im Falle einer Massendigitalisierung zumeist externe Dienstleister beauftragt, so dass die Herausforderungen hier in erster Linie in der inhaltlichen Erschließung und der Erhaltung des originalen Materials liegen.

**Für alle Materialarten gilt:** Eigene Scanner sind nur erforderlich, wenn der Scanprozess nicht an einen Dienstleister ausgelagert werden soll oder kann. Nach der Auswahl geeigneter Geräte (am besten nach Beratung durch erfahrene Einrichtungen) empfiehlt sich eine Teststellung, bei der diese unter Realbedingungen erprobt werden können.

### Erschließen

#### Erschließen von Metadaten

Erschließen bedeutet das Zusammentragen von allen verfügbaren Informationen zu einem Objekt und die Codierung in Form von [Metadaten](#daten-metadaten). Im Kontext der Digitalisierung sind drei [Arten von Metadaten](#arten-von-metadaten) besonders relevant:

* **Administrative Metadaten** mit Informationen zu Herkunft, Erhaltungszustand, technische Merkmale, Rechteinformationen.
* **Deskriptive Metadaten** zur bibliografischen Beschreibung des Objektes.
* **Strukturelle Metadaten** für Gliederungselemente wie z.B. Kapitel in Texten, Segmente in 3D-Objekten, Frames in Filmen.

Bibliografische Basisdaten wie auch strukturelle Metadaten können in einem WMS entweder importiert oder manuell erfasst werden. Viele der Systeme sind dafür individuell konfigurierbar und können je nach Art des Projektes angepasst werden. Ein wichtiges Metadatum sind [Identifikatoren](#identifikatoren) (URN, DOI, PURL…) und deren Generierung. Für die Strukturierung von Textdokumenten ist das [Stukturdatenset für den DFG-Viewer](https://dfg-viewer.de/strukturdatenset) ein praxistauglicher Referenzrahmen.

Um die Nachnutzung der digitalisierten Objekte zweifelsfrei zu kennzeichnen, gehören Informationen zur Lizenzierung in die administrativen Metadaten. Dafür hat eine Expertengruppe innerhalb von DINI detaillierte [Empfehlungen für Rechteinformationen in Metadaten](https://wiki.dnb.de/pages/viewpage.action?pageId=217533652) aufgestellt.

Strategien für das Crowdsourcing von Erschließungsarbeiten in offenen Informationsinfrastrukturen wie dem Wikiversum ermöglichen potenziell zusätzliche Mehrwerte, auch im Nachhinein und durch Dritte.

#### Zusätzliche inhaltliche Erschließung durch Volltexte und Annotationen

Erschließung endet nicht bei der Zuordnung von Metadaten. Ziel der Erschließung ist auch eine Abbildung des Werkes in nachnutzbarer Form. Bei Textdokumenten ist das die Generierung von Volltexten. Für die Auswahl eines Verfahrens und deren Parametrisierung sind Eigenschaften der Vorlage (Schriftart, Layout, Sprache, Papier- und Druckqualität u.a.) sowie die benötigte Qualität der Volltexte entscheidend.

Waren anfangs vor allem kommerzielle OCR-Werkzeuge dominierend (z.B. [ABBYY](https://de.wikipedia.org/wiki/ABBYY)), haben sich mittlerweile einige Open Source-Projekte (z.B. [Tesseract](https://de.wikipedia.org/wiki/Tesseract_(Software))) entwickelt, die qualitativ vergleichbare oder bessere Ergebnisse erzeugen können. Dank neuer technologischer Ansätze sind auch Lösungen für handschriftliche Dokumente verfügbar (z.B. [OCR4all](https://www.ocr4all.org/about/ocr4all)). Mit Förderung durch die DFG arbeitet die Initiative OCR-D an einem frei nutzbaren Workflow, der neben der eigentlichen Volltexterkennung weitere Schritte (z.B. automatische Layouterkennung) einbindet, die die Qualität der Erschließung verbessern.

Die Generierung von Volltexten kann als (automatischer) Prozessschritt im [Workflowmanagementsystem](#workflowmanagementsysteme) eingebunden werden. Die Ergebnisse werden in der Regel in [standardisierten Datenformaten](#metadaten-und-schnittstellen) geliefert und können dann in Präsentationssystemen z.B. für die Volltextsuche genutzt oder zum Download angeboten werden.

Neben Volltexten können je nach Kontext auch Annotationen als weiteres Element der Erschließung eine Rolle spielen, die aber in der Regel nicht im Digitalisierungsprozess, sondern erst im Nachhinein entstehen.

### Präsentieren

Die Präsentation von digitalisierten Objekten und den dazugehörigen Metadaten im Web ist ein wesentliches Ziel der Digitalisierungs-Aktivitäten.

Die Präsentation der digitalisierten Objekte hängt von der Art von Objekten und beabsichtigten Nutzungsszenarien ab und erfordert eigene Werkzeuge, die unabhängig von dem genutzten Workflowmanagementsystem ausgewählt und konfiguriert werden. Basis für Präsentationslösungen sind die im Workflow erzeugten standardisierten XML-Files (METS/MODS).

Funktional wird von einer Präsentationslösung eine konfigurierbare Recherche-Umgebung für Metadaten und Volltexte erwartet. Bei der Präsentation des einzelnen Objektes werden die Strukturdaten zusammen mit den dazu erfassten Metadaten als Inhaltsverzeichnis angeboten und es werden zur Arbeit mit den Scans Bildwerkzeuge erwartet (z.B. Zoomen, Rotieren, Thumbnail-Übersicht). Zunehmend wichtige Funktionen sind Download-Angebote oder die Bereitstellung von iiif-Manifesten zur Weiternutzung in anderen Viewern.

Mit dem [DFG-Viewer](http://dfg-viewer.de/) existiert eine Minimallösung für die Anzeige von Digitalisaten, die aus dem Katalog verlinkt werden kann. Für eine umfassendere Präsentation von Digitalisaten entscheiden sich Einrichtungen für eine Integration der Digitalisate in ein existierendes Repositorium oder für den Aufbau von Web-Angeboten mit speziellen Werkzeugen für Digitalisate wie z. B. *Kitodo.Presentation*. Ein wichtiges Element für mehr Sichtbarkeit der eigenen Digitalisierungs-Aktivitäten ist eine Datenlieferung an die Deutsche Digitale Bibliothek.

Zunehmend relevanter wird die Suche nach Lösungen für eine Präsentation von rechtebehafteten Materialien für eingeschränkte Personenkreise. Repositorien TODO: Link auf Kapitel in Text zu forschungsnahen Diensten haben hier in der Regel schon anpassbare Lösungen. Bei der Präsentation von Digitalisaten existieren prototypische oder an Institutionen individuell angepasste technische Lösungen. Standards für die Beschreibung von Zugriffsrechten in den Metadaten oder generische Präsentationslösungen auf Basis dieser Metadaten befinden sich noch im Abstimmungsprozess oder der Entwicklung.

Wikimedia Commons kann darüber hinaus als zentraler Medienspeicher für die Portalfamilie der Wikimedia-Bewegungn fungieren. Von dort werden digitalisierte Objekte und ‚born digital‘-Dokumente dezentral in den anderen Portalen eingebettet. Die Erschließung mit strukturierten Daten erfolgt mittels Wikidata (siehe [Digitalisierung als Beitrag zur Open GLAM-Bewegung](#Xabc6d7253650c5e5284c03076d215a642b91408)).

## Werkzeuge

### Scanner und Scanverfahren

#### 2D

Scanner und ihre Verfahren lassen sich grundlegend nach der Eigenschaft ihrer zu scannenden Objekte und der Zielsetzung des Scan-Ergebnisses unterscheiden.

Im Bereich des 2D-Scans unterstützen z.B. spezielle Buchscanner den Scanprozess im Rahmen der Digitalisierung effizient und für das zu digitalisierende Material auf schonende Weise. Das sind in der Regel Auflichtscanner oder fotografische Systeme, bei denen das Material von oben abgelichtet wird. Je nach Art und Zustand des Materials können dabei Buchwippen bzw. -wiegen mit verschiedenen Winkeln oder Glasplatten eingesetzt werden. Roboter können das automatische Umblättern der Seiten erledigen, sind dabei aber für empfindliche Bestände weniger schonend. Berührungslose Scanner sind vor allem für Handschriften vorgesehen.

|  |
| --- |
| BILDUNTERSCHRIFT FEHLT! |

Die 2D-Scanner können sich in verschiedenen Eigenschaften wie Größe der Formate, Farbtiefe, Auflösung, Art und Qualität der Optik und Mechanik sowie in der Verarbeitungsgeschwindigkeit unterscheiden. Um eine hohe und gleichbleibende Bildqualität zu gewährleisten, ist eine regelmäßige Kalibrierung der Systeme und Auswertung der Aufnahmen anhand von Test-Targets (z.B. UTT-Target oder GoldenThread) erforderlich. Neben der Hardware unterscheiden sich Scanner in der Software, durch die Rohdaten des Scanners für die weitere Nutzung bearbeitet (z.B. beschneidet oder glättet), optimiert und in gewünschten Zielformaten speichert.

#### 3D

Bei der 3D-Erfassung eines Objektes kann zwischen photo- und scanbasierten Verfahren unterschieden werden.

So werden bei der Photogrammetrie Fotos aus verschiedenen Blickwinkeln aufgenommen. Dabei wird die Anzahl der Fotos derart erhöht, dass sie mit Hilfe von Software zu einem 3D-Modell verarbeitet werden können. Nach diesem Prinzip funktionieren auch 3D-Scanning-Apps für Smartphones, die 3D-Scans von Objekten mithilfe der Kamera eines Smartphones ermöglichen.

Demgegenüber stehen verschiedene Scanverfahren. Beim Laser-Scanning werden z.B. Laserstrahlen auf das Objekt gerichtet, und die reflektierten Daten werden verwendet, um ein präzises 3D-Modell zu erstellen. Neben dem Laser lassen sich solche 3D-Modelle auch mit strukturiertem Licht (Streifenlichtscans) oder Computertomographie erzeugen.

Bei der 3D-Modellierung wird ein 3D-Modell nach Vorlagen manuell und mit Hilfe von spezifischer Software erstellt. Dies kommt insbesondere für die digitale Rekonstruktion von physisch nicht erhaltenen (Teilen von) Objekten zur Anwendung.

Alle Verfahren können miteinander kombiniert werden.

Wird beispielsweise ein Pop-up Buch oder ein Globus 3D-digitalisiert, sollen zusätzlich die kinetischen Aspekte (ein Globus dreht sich um seine Achse, bei einem Pop-up Buch lässt sich eine Lasche herausziehen usw.) unter Umständen auch im Digitalisat nachvollziehbar bleiben. Besteht das Originalobjekt aus Einzelteilen und werden diese entsprechend in Segmenten digitalisiert, können diese Segmente interaktiv nutzbar gemacht werden oder/und Animationen eingesetzt werden. Ein einfacher Scan reicht in diesem Fall nicht aus, da hier lediglich "die Hülle im Ganzen" erfasst wird. Eine manuelle Nachbearbeitung ist nötig.

### Workflowmanagementsysteme

Bei diesen Systemen handelt es sich im Grunde um klassische Workflowmanagementsysteme ([WMS](https://de.wikipedia.org/wiki/Workflow-Management-System)), die auf die Abbildung und Modellierung der [oben beschriebenen Prozesse](#prozesse) spezialisiert sind. Das Ziel der Systeme ist es, die Prozesse technisch abzubilden und zu organisieren bzw. die Zustände der im Prozess befindlichen Objekte zu speichern und darzustellen.

|  |
| --- |
| Beispiel für einen im WMS eingerichteten Workflow |

Aufgrund der hohen Spezialisierung ist die Auswahl an WMS für bibliothekarische Digitalisierungsprozesse überschaubar. Dazu gehören sowohl kommerzielle Produkte, wie z.B. [Visual Library](https://www.semantics.de/visual_library/) aber auch frei verfügbare Open Source Systeme wie [*Kitodo*](https://www.kitodo.org/).

Der wesentliche Unterschied zwischen den Tools – neben der Frage der Lizenz – ist deren Modularität. Bei einigen Lösungen sind z.B. die Erstellung, Anreicherung und Präsentation fest in einem System zusammengefasst, andere wählen z.B. den Ansatz, die Präsentation von der Erstellung zu trennen, sodass die Darstellung z.B. auch mit spezialisierten Präsentationswerkzeugen, wie Content-Management-Systemen erfolgen kann.

### Metadaten und Schnittstellen

An vielen Stellen des Digitalisierungsprozesses sind Schnittstellen (APIs) zu Drittsystemen notwendig. Das gilt sowohl für die Einbindung externer Metadaten (beispielsweise der Import aus Nachweissystemen via *SRU* aus Verbundkatalogen oder Kalliope) als auch die Weitergabe von Daten an externe Portale wie die Deutsche Digitale Bibliothek (*DDB*).

Als Austauschformat verwenden diese Schnittstellen meist [XML-Datenformate](#xml-basierte-datenformate) wie *PICA/XML*, *METS/MODS*, *EAD* und *LIDO*, die auf Struktur- oder Metadaten digitalisierter Objekte spezialisiert sind. [Relevante Schnittstellen](#schnittstellen) sind vor allem *SRU*, *OAI-PMH* und *IIIF*. Die Nutzung gemeinsamer Datenformate ermöglicht unter anderem die Auffindbarkeit, Verknüpfung, Präsentation und Verwendung von Objekten in unterschiedlichen Zusammenhängen.

Für einzelne Ressourcentypen und Anwendungen existieren oft genauere Anwendungsprofile, die die Verwendung von Metadaten-Standards erweitern oder einschränken.

**Beispiel:** Das METS/MODS-[Anwendungsprofile](https://wiki.deutsche-digitale-bibliothek.de/display/DFD/5.+Anwendungsprofile+und+Best+Practice+Guides) der Deutschen Digitalen Bibliothek legt Anforderungen an Metadaten fest, wenn diese in das Portal eingebunden werden sollen. Zur [Validierung der Daten](https://wiki.deutsche-digitale-bibliothek.de/display/DFD/Schematron-Validierungen+der+Fachstelle+Bibliothek#SchematronValidierungenderFachstelleBibliothek-Schematron-Dateien) stellt die DDB passende Werkzeuge zur Verfügung.

### Betriebsmodelle

Bei dem Einsatz der Werkzeuge im Digitalisierungsprozess stellt sich immer die Frage, ob sich die Anschaffung der Hardware oder die hausinterne Installation und Anpassung von Software lohnt oder aber Bestandteile ausgelagert werden können. Wie bei allen IT-Systemen sind grundsätzlich [drei Betriebsmodelle](#betriebsmodelle) denkbar:

* **Lokale Installation**: Hard- und Software wird durch betriebseigenes Personal gehostet, betreut und weiterentwickelt. Dieses Modell ist vor allem dann empfehlenswert, wenn bereits Digitalisierungskompetenz in der Institution vorhanden ist und die Digitalisierung einen dauerhaften Schwerpunkt in den Alltagsprozessen einnimmt bzw. einnehmen soll.
* **Externes (Cloud-)Hosting**: Sofern nicht z.B. rechtliche Gründe dagegen sprechen, ist dieses Modell empfehlenswert, wenn nur punktuell oder in geringem Umfang digitalisiert werden soll und dabei keine eigene Expertise notwendig ist oder aufgebaut werden soll.
* **Hybrid-Lösungen:** Es werden nur Teile des Prozesses ausgelagert, z.B. das Scannen der Objekte, während z. B. die Inhaltsanreicherung durch Metadaten im Haus verbleibt.

Die Hybrid-Variante wird im Alltag am häufigsten umgesetzt. Die Herausforderung besteht darin, die Schnittstellen zwischen externen und internen Prozessen genau zu definieren und festzulegen. So muss zum Beispiel externen Scan-Dienstleistern Zugriff auf das intern betreute Workflowmanagementsystem gewährt werden, damit digitalisierte Objekte direkt im System abgelegt werden können.

## Digitalisierung als Beitrag zur Open GLAM-Bewegung

Mit Digitalisierung leisten Bibliotheken und andere GLAM-Einrichtungen wichtige Beiträge dazu, dass Kulturgüter weltweit zugänglich, gut auffindbar und nachnutzbar sind. Neben der einrichtungseigenen Präsentation und aggregierenden Portalen wie der „Deutschen Digitalen Bibliothek“ oder der „Europeana“ spielen auch offene Wissensplattformen im Web eine wichtige Rolle, um Daten zugänglich zu machen, zu teilen und Nutzende an der Erschließung zu beteiligen. Entsprechende Möglichkeiten sollen im Folgenden skizziert werden, da es sich um einen wichtigen Trend bei der Digitalisierung und Nutzung handelt.

Offener Zugang und offene Lizenzen für freie Nutzungsszenarien, standardisierte Schnittstellen und offene Metadaten vereinfachen das Teilen von Daten bzw. Informationen aus digitalisierten Quellen – Texten, Illustrationen und Sammlungen – als freies Wissen.

### Veranstaltungsformate und Handlungsfelder

Hackathons und Editathons sind Veranstaltungsformate, um kurzfristig die Bekanntheit digitaler Sammlungen zu steigern, Datenexperimente zu ermöglichen und Datenkollaborationen in Pilotprojekten anzustoßen. Dabei können Impulse entstehen für weiterentwickelte Datenanwendungen. Lerneffekte entstehen dabei nicht nur bei den beteiligten Menschen, sondern auch institutionell bei den beteiligten Bibliotheken und anderen GLAM, die ggf. Bedürfnisse, Ideen und Defizite von Nutzenden kennenlernen. Ein Beispiel Hackathons ist die Initiative [„Coding da Vinci“](https://codingdavinci.de/).

Crowdsourcing und Citizen Science sind Handlungsfelder, die in Bibliotheken zunehmend Aufmerksamkeit wecken. Hierbei gewinnen Bibliotheken über externe engagierte Dritte zusätzliche Ressourcen für die Erschließung, Auswertung oder Korrektur von Daten. Ein Beispiel dafür ist das Projekt „[Die Datenlaube](https://de.wikiversity.org/wiki/DieDatenlaube)“. In der deutschsprachigen Wikisource korrigiert eine Gemeinschaft Ehrenamtlicher seit 2008 die von verschiedenen Bibliotheken bereitgestellten Scans der Illustrierten „Die Gartenlaube“. Diese Wikisource-Volltexte und -Illustrationen werden mit offenen Metadaten in Wikidata strukturiert erschlossen.

|  |
| --- |
| BILDUNTERSCHRIFT FEHLT |

### Offene Infrastrukturen

Offene digitale Infrastrukturen spielen für die Speicherung, Erschließung und Vernetzung von Digitalisaten eine wichtige Rolle, nicht zuletzt auch, um die beschriebene Beteiligung zu ermöglichen. Das prominenteste Beispiel dafür sind die Projekte der Wikimedia-Foundation, deren Grundprinzipien von Offenheit und Referenzierbarkeit die bibliothekarischen Erschließungs- und Präsentationskomponenten ergänzen:

* **Wikimedia Commons** fungiert als zentraler Medienspeicher. Detailerschließungen z.B. mittels relativer Bildpositionen in den Metadaten von Commons-Dokumenten ist möglich (Structured Data on Commons) (siehe [Wikidata Image Positions](https://wd-image-positions.toolforge.org/)).
* Mit der offenen Wissensdatenbank (Knowledge Graph) **Wikidata** können insbesondere Metadaten strukturiert und damit maschinenlesbar gespeichert, gemeinsam editiert, visualisiert, analysiert, geteilt und verknüpft werden.
* Die vielsprachigen Versionen der **Wikipedia** speichern Wissen in enzyklopädischen Artikeln und ermöglichen so die Zusammenführung von Digitalisaten mit Texten, Illustrationen aus den Wikimedia Commons, Referenzen, Links und damit Kontext.
* **Wikisource** ist eine Quellensammlung gemeinfreier Werke und besteht sowohl aus transkribierten Volltexten und deren Illustrationen als auch thematischen Linksammlungen zu relevanten Digitalisaten in Digitalen Sammlungen öffentlicher und wissenschaftlicher Bibliotheken, Archive, Museen und Galerien (GLAM).

Herausforderungen aus IT-bibliothekarischer Sicht bestehen darin, Wikimedia-Portale und bibliothekarische Werkzeuge durch Schnittstellen und Datentransformation ([ETL-Tools](#etl-prozess)) miteinander zu verknüpfen. GLAM-Labore (Open GLAM Labs) können dafür als Experimentier- und Ausbildungsgsort fungieren – in Bibliotheken und institutionenübergreifend z.B. im Rahmen von Landesdigitalisierungsprogrammen.

## Checkliste für Digitalisierungsprojekte

Diese **Checkliste** dient als Leitfaden für die Planung und Umsetzung von Digitalisierungsprojekten in Bibliotheken und hilft sicherzustellen, dass die Ziele erreicht und die digitalen Sammlungen effektiv genutzt werden können.

**Vorbereitungsphase:**

1. *Festlegung der Ziele*: Klären Sie die Hauptziele und den Zweck der Digitalisierung (z.B. Zugänglichkeit, Erhaltung, Forschung).
2. *Ressourcenplanung*: Ermitteln Sie die erforderlichen Ressourcen wie Verfügbarkeit der Objekte, Personal, Technologie, Budget und Zeitrahmen.
3. *Rechtliche Überlegungen*: Klären Sie urheberrechtliche und rechtliche Fragen bezüglich der digitalisierten Materialien.
4. *Auswahl der Objekte*: Wählen Sie die Objekte aus, die digitalisiert werden sollen, u. a. basierend auf ihrer Bedeutung und ihrem Zustand.
5. *Metadatenplanung*: Entwickeln Sie einen Plan zur Erfassung und Verwaltung von Metadaten für die digitalisierten Materialien.

**Technische Umsetzung:**

1. *Scantechnologie*: Wählen Sie die geeignete Scantechnologie und -auflösung entsprechend der Art der Objekte und deren Verwendungszweck aus.
2. *Qualitätskontrolle*: Implementieren Sie Prozesse zur Sicherstellung der Qualität der Digitalisate.
3. *Dateiformate*: Entscheiden Sie über die geeigneten Dateiformate für die digitalisierten Materialien.
4. *Speicher und Sicherung*: Planen Sie die Speicherung und regelmäßige Sicherung der digitalen Sammlung.

**Metadaten und Beschreibung:**

1. *Metadatenstandards*: Verwenden Sie etablierte Metadatenstandards, um die Beschreibung der Objekte zu standardisieren.
2. *Dokumentation*: Dokumentieren Sie den Digitalisierungsprozess und die verwendeten Metadaten.
3. *Zugänglichkeitsprüfung*: Überprüfen Sie, ob die digitalen Sammlungen barrierefrei und benutzerfreundlich sind.

**Langzeitpflege**:

1. *Erhaltungsstrategien*: Planen Sie langfristige Erhaltungsstrategien für digitale Materialien.
2. *Aktualisierung und Migration*: Berücksichtigen Sie die Notwendigkeit von Aktualisierungen und Migrationen von Dateiformaten und Systemen.

**Zugang und Bereitstellung:**

1. *Plattformwahl*: Wählen Sie die geeignete Plattform oder das geeignete System zur Präsentation und Verwaltung der digitalen Sammlungen.
2. *Nutzungsrechte*: Klären Sie die Nutzungsrechte und -bedingungen für die digitalen Inhalte.
3. *Nutzerfeedback*: Implementieren Sie Mechanismen zur Erfassung von Nutzerfeedback zur Verbesserung der digitalen Sammlungen.

**Veröffentlichung und Dissemination:**

1. *Verbreitung*: Planen Sie die Verbreitung der digitalisierten Sammlungen über Online-Kanäle, soziale Medien und Partnerschaften.
2. *Marketing*: Entwickeln Sie Marketingstrategien, um die digitalen Sammlungen einem breiten Publikum bekannt zu machen.
3. *Evaluation*: Führen Sie regelmäßige Bewertungen durch, um den Erfolg Ihrer Digitalisierungsprojekte zu messen und zu verbessern.

## Zusammenfassung und Ausblick

Viele Bibliotheken mit historischen Beständen sind im Feld der Kulturgutdigitalisierung engagiert und kooperativ unterwegs. Werkzeuge, Standards und Dienstleister stehen zur Verfügung, um die unikalen Bestände ins Netz zu bringen. Um die Bandbreite der Objekte zu erweitern, finden gerade wichtige Entwicklungen statt, um auch AV-Medien und Objekte jenseits von 2D in die Routine aufzunehmen. Ebenso werden die Bibliotheken mit den Ergebnissen aus OCR-D immer mehr in der Lage sein, auch ohne Expertenwissen gute Volltexte zu erzeugen bzw. ihre gesamten Digitalisate inkl. handschriftlicher Dokumente mit Volltexten anzureichern. Hier sind ggf. auch kooperative Infrastrukturen aufzubauen, um die Fortschritte in den technischen Möglichkeiten schnell zur Verfügung zu stellen und eine regelmäßige Erneuerung der Volltexte zum Standard werden zu lassen.

Erste Bibliotheken machen sich mit der schnell wachsenden Menge an Volltexten auf den Weg, mit Hilfe ausgewählter Tools aus dem Feld des Maschinenlernens, ihre Suche semantisch anzureichern auf Basis der vorgeschalteten automatisierten Analyseverfahren.

Mit dem großen Schatz der Volltexte wird es darüber hinaus wichtiger werden, diesen auch den Forscher:innen zusätzlich über eine API anzubieten, um individuelle Korpora für den Download in eigene Umgebungen zusammenstellen zu können. Hier landet man dann auch schnell bei der Suche nach weitergehenden nachhaltigen Lösungen für urheberrechtlich geschützte Materialien.

Aber auch die kollaborativen Aktivitäten an einzelnen Digitalisaten im Wikiversum oder in Citizen Science Projekten sind getragen von dem Leitgedanken, dass die digitalisierten Kulturgüter in neue Zusammenhänge gestellt werden und Interessierte zu einem Thema zusammen bringen.

# Forschungsnahe Dienste

|  |
| --- |
| Wichtig |
| Dieses Kapitel befindet sich in einer ersten öffentliche Entwurfsversion. Feedback siehe Links im rechten Menü und [Hinweise zur Mitarbeit](#mitarbeit). |

|  |
| --- |
| Zusammenfassung |
| Bibliotheken bieten [Dienste für die Wissenschaft](#einleitung) im Rahmen des digitalen Wandels von Publikationsprozessen und Open Science an. Dazu gehören [Dienste zur Publikation](#publikationsdienste), zur [Verwaltung von Forschungsdaten](#forschungsdatenmanagement) und [Forschungssoftware](#forschungssoftware) sowie [Informationen über Forschungsprozesse](#forschungsinformationssysteme).  [*Übergreifende Themen*](#übergreifende-themen) *wie diese* [*Zertifikate und Standards*](#zertifikate-und-standards) *und* [*zusammen mit externen Dienstleistern*](#zusammenarbeit-mit-dienstleistern) *werden in diesem Kapitel auch behandelt, der Text muss aber noch überarbeitet werden!* |

## Einleitung

Unter forschungsnahen Diensten werden verschiedene Bibliotheksservices zusammengefasst, die Wissenschaftler\*innen im gesamten Forschungsprozess unterstützen sollen und die überwiegend im Kontext von digitalem Wandel und Open Science angesiedelt sind (<https://www.o-bib.de/bib/article/view/5718>). Dazu zählen z. B. Services in den Bereichen [Forschungsdatenmanagement](#forschungsdatenmanagement), Bibliometrie sowie verschiedene [Publikationsdienste](#publikationsdienste).

Einige dieser Dienste, z. B. Repositorien für Zeitschriftenartikel, die unter Open-Access-Bedingungen zweitveröffentlicht werden dürfen, gehören schon seit Jahrzehnten zum Dienstleistungsrepertoire wissenschaftlicher Bibliotheken. Inzwischen betreiben größere Einrichtungen zudem oft spezialisierte Repositorien für ein Spektrum verschiedener Objekttypen: Publikationen, Zeitschriften, Bücher, Open Educational Resources, Forschungsdaten, Forschungssoftware und mehr. Dazu gehören auch diverse Dienste, die übergreifend den Forschungsoutput bestimmter lokaler oder fachlicher Forschungscommunitys besser auffindbar oder messbar machen sollen, darunter Forschungsinformationssysteme und Dienstleistungen im Bereich von Metriken oder zur Verwaltung von Artikelveröffentlichungsgebühren in Open-Access-Journals (APCs).

|  |
| --- |
| Definition |
| **Open Science** bezeichnet den Ansatz, wissenschaftliche Forschungsergebnisse, Daten und Methoden frei zugänglich und transparent zu teilen. Teilaspekte davon sind Open Access, Open Data, Open Educational Resources und Open Source). Ziel ist es, die Zusammenarbeit sowie die Reproduzierbarkeit von Forschung zu fördern, Innovationen zu beschleunigen und den gesellschaftlichen Nutzen wissenschaftlicher Erkenntnisse zu maximieren. Dieser offene Ansatz erleichtert es Forschenden weltweit, Informationen frei zu nutzen, zu teilen und weiterzuentwickeln. |

Forschungsförderer und Ministerien erwarten in Deutschland heute von den Forschungseinrichtungen eine Transformation des Publizierens hin zu Open Access und Open Data. Dazu gehören themenbezogene institutionelle Policys, Beratungs- und Schulungsangebote sowie technische Dienste. Unter anderem auf der Ebene der europäischen Forschungsförderung wird versucht, durch Initiativen wie [*CoARA*](https://coara.eu/) die Maßstäbe der Forschungsbewertung weiterzuentwickeln - weg von klassisch bibliometrischen Indikatoren wie h-Index und Journal-Impact-Faktor hin zu einer Würdigung vielfältiger verschiedenartiger Forschungsergebnisse.

In Abgrenzung zur Dominanz von Großverlagen ist im Bereich der wissenschaftlichen Informationsinfrastrukturen darüber hinaus der Anspruch entstanden, Dienste - bis hin zur Ebene von Entwicklung und Betrieb zugehöriger IT-Infrastrukturen - durch die Communitys der Forschenden selbst zu betreiben. In der englischsprachigen Fachdiskussion sind dazu Begriffe wie „scholar-led publishing“ oder „scholar-led conferences“ geprägt worden, vgl. exemplarisch hierzu das [Scholar-Led.Network Manifesto](https://graphite.page/scholar-led-manifesto/). Bibliotheken sehen sich hier als vertrauenswürdige, öffentlich finanzierte Dienstleister für die Wissenschaftscommunitys.

Der Einsatz von und die aktive Mitarbeit an Open Source-Software ist bei den forschungsnahen Diensten stärker als in anderen Bereichen der Bibliotheks-IT eine Selbstverständlichkeit. Auch die Bedeutung der gemeinschaftlichen Pflege offener Standards, Datenmodelle und Daten-Gemeingüter ist hier besonders stark ausgeprägt.

## Publikationsdienste

Die Veröffentlichung wissenschaftlicher Erkenntnisse ist ein zentraler Bestandteil des wissenschaftlichen Arbeitens. Solange dieser Prozess noch in den Geschäftsprozessen klassischer Subskriptionszeitschriften abgebildet war, war auch die Rolle der wissenschaftlichen Bibliotheken eher auf der Seite der Bereitstellung der Literatur. Mit der zunehmenden Stärkung des Open-Access-Gedankens und dem Aufkommen neuer Geschäftsmodelle rückten auch die wissenschaftlichen Bibliotheken stärker in die Rolle des Unterstützers und Ermöglichers. Die Digitalität der gesamten Prozesskette vom Schreiben bis hin zur Veröffentlichung der Artikel in teilweise von den Bibliotheken getragenen Infrastrukturen erfordert die Einbindung unterschiedlicher IT-Werkzeuge, die im Folgenden näher beschrieben werden.

### Journal Publishing-Dienste

Open Access bedeutet, dass wissenschaftliche Literatur kostenfrei und öffentlich im Internet zugänglich ist, sodass Interessierte die Volltexte lesen, herunterladen, kopieren, verteilen, drucken, in ihnen suchen, auf sie verweisen und sie auch sonst auf jede denkbare legale Weise benutzen können, ohne finanzielle, gesetzliche oder technische Barrieren jenseits von denen, die mit dem Internetzugang selbst verbunden sind. Zum ersten Mal wurde dieser Gedanke in der [Grundsatzerklärung der Budapester Open-Access-Bewegung](https://www.budapestopenaccessinitiative.org/translations/german-translation) formuliert.

Durch die Transformationsprozesse im wissenschaftlichen Publikationswesen weg von den traditionellen Abonnement-Modellen hin zu Open Access sehen sich Bibliotheken zunehmend auch in der Rolle eines Publikationsdienstleisters. Dies kann einerseits die Gründung eines Universitätsverlages bedeuten, andererseits aber auch die Schaffung der benötigten Infrastruktur für wissenschaftliche Zeitschriften. Befördert durch die Verlagerung der Kosten weg vom Lesen hin zum Publizieren wird der *scholar-led*-Ansatz immer gefragter und Bibliotheken geraten in die Situation, Expertise in diesem Bereich aufbauen zu müssen.

Zur Schaffung einer technischen Infrastruktur lässt sich z. B. mit der Software [*Open Journal Systems (OJS)*](https://ojs-de.net/ueber-ojs) eine Plattform zur Verfügung stellen, die die strukturierte Veröffentlichung von Zeitschriften(-artikeln) ermöglicht. Parallel dazu müssen auch die erforderlichen Abläufe und Organisationsstrukturen angepasst werden. Personelle Ressourcen müssen hier ebenso bedacht werden. In erster Linie gilt es, die Herausgeber\*innen-Teams der Zeitschriften zu unterstützen. Gleichzeitig sollte die Bibliothek auch technischen Support für einreichende Autor\*innen bieten. Der Funktionsumfang von OJS ermöglicht es auch, einen Workflow für den Peer-Review-Prozess abzubilden. Auch hier liegt Potenzial für die Unterstützung durch Bibliotheken. Wichtig ist somit ein Überblick über den Gesamtprozess des wissenschaftlichen Publizierens und nicht nur die Software-Aspekte.

Parallel dazu entwickeln sich derzeit "alternative Publikations-Plattformen” wie Preprint-Dienste, (Micro-)Blogs, Data Journals und ähnliche Dienste, die traditionelle Publikationswege wie peer-reviewte Journals ergänzen, siehe dazu auch die entsprechende [Publikation der Initiative Knowledge Exchange](https://doi.org/10.21428/996e2e37.3ebdc864). Hier können Bibliotheken ebenso in die Rolle des Dienstleisters treten, Instanzen geeigneter Open-Source-Software hosten und diese Entwicklung unterstützen. Dies kann auch im Zusammenhang mit der externen Kommunikation von Bibliotheken (Link zu entsprechendem Kapitel) betrachtet werden. Ergänzend gibt es zahlreiche weitere Dienste, die das Open-Access-Publizieren vor allem administrativ unterstützen, von denen im Folgenden einige kurz erläutert werden.

### Open-Access-Dienste

Neben der Unterstützung bei Open-Access-Veröffentlichungen haben sich im Zusammenhang mit der Open-Access-Transformation weitere IT-Dienste in Bibliotheken entwickelt.

#### Verlags-Software

Da im Sinne der „scholar-led infrastructure“ zunehmend Universitätsverlage gegründet werden, existieren hier auch eigene Software-Lösungen für die Verwaltung der Titel, Lagerbestände, Kund\*innen-, Adress- und Versanddaten. Häufig ist ein eigener Webshop mit entsprechenden Shopfunktionalitäten integriert und es gibt Schnittstellen zur Buchhaltung und Auslieferung. Für kleinere Verlage reicht vermutlich ein normaler Webshop aus; Spezialsoftware kann verlagsspezifische Anforderungen allerdings besser abbilden. Betrieb und Administration der Programme gehören mit in das zugehörige Diensteportfolio.

#### Publikationsfonds-Verwaltung

Publikationsfonds zur Finanzierung von Open-Access-Publikationen sind an vielen wissenschaftlichen Bibliotheken fest verankerte Hilfsmittel, um die Transformation des wissenschaftlichen Publikationswesens hin zu Open Access zu unterstützen. Durch sie werden anfallende Article Processing Charges (APC) nach durch die Einrichtung vorgegebenen Kriterien (mindestens anteilig) zentral gezahlt. Zum Monitoring der entstehenden Kosten ist häufig eine Tabellenkalkulationssoftware ausreichend. Für einen Vergleich verschiedener Einrichtungen und ein einheitliches Reporting existieren jedoch Services wie [OpenAPC](https://openapc.net/). Zur einheitlichen Darstellung aller Publikationskosten hat sich mit [OpenCost](https://www.opencost.de/metadatenschema/) ein xml-Metadatenschema etabliert.

#### Open Refine

Sollen Daten aus unterschiedlichen Quellen zusammengeführt werden oder ein Abgleich gegen externe Datenbanken stattfinden, ist [*OpenRefine*](https://openrefine.org/) ein geeignetes Werkzeug: Über offene Schnittstellen können Datensätze sowohl erweitert als auch beispielsweise in Wikidata exportiert werden. Hierbei ist keine zentrale Installation notwendig. Die Software wird lokal auf den Endgeräten ausgeführt. Lediglich eine korrekte Einstellung der Firewall, um die Kommunikation mit externen Datenbanken zu ermöglichen, gilt es zu beachten.

#### Journal-Finder-Werkzeuge

Neben den üblichen Fragen zu Lizenzen bei Open-Access-Publikationen spielt auch die Wahl eines geeigneten Journals eine zentrale Rolle. Derzeit verbreitete Journal-Finder-Werkzeuge wie [B!SON](https://service.tib.eu/bison/) oder der [oa.finder](https://finder.open-access.network/) greifen für die Journal-Daten zwar beide auf das Directory of Open Access Journals ([DOAJ](https://doaj.org/)) zu, verfolgen jedoch unterschiedliche Ansätze. Der oa.finder zeigt eine Liste von Zeitschriften mit Filteroptionen und den jeweiligen Förderbedingungen an. Diese werden aus bestehenden Transformationsverträgen abgeleitet und enthalten keine spezifischen Förderbedingungen. B!SON bietet teilnehmenden Einrichtungen die Möglichkeit, eine eigene Liste mit Förderbedingungen zu hinterlegen. Treffer werden direkt mit den korrekten Förderbedingungen angezeigt. Zu diesem Zweck muss ein Web-Backend entsprechend der Vertragssituation gepflegt werden.

### Repositorien für Forschungsergebnisse

Zentral für die Veröffentlichung jedweder Art von wissenschaftlichem Output ist ein geeigneter Ort für deren Veröffentlichung, ganz besonders im Hinblick auf die zunehmende Datengetriebenheit der Wissenschaften.

Heute haben sich Open-Access-Repositorien als verlässliche Speicherdienste für wissenschaftliche Ergebnisse etabliert, seien es [Forschungsdaten](#forschungsdatenmanagement), textuelle Publikationen, Softwarecode oder andere Materialien (siehe **?@fig-informationstypen**). Für die wissenschaftliche Kommunikation ist dabei die Zitierbarkeit eine wichtige Voraussetzung. Dazu muss sichergestellt werden, dass publizierte Ergebnisse in der veröffentlichten Form erhalten werden, das heißt nicht verändert oder gelöscht werden. Da Internetadressen als flüchtig gelten, werden zur Identifikation [Persistent Identifier-Systeme](#persistent-identifier) eingesetzt. Die Zitierfähigkeit und damit der langfristige, möglichst originalgetreue Erhalt der einmal eingestellten Informationen grenzen Repositorien gegenüber anderen Arbeitsplattformen wie Sync-and-Share-Plattformen (z. B. Nextcloud), Content-Management-Systemen (zur Erstellung von Blogs und Internetseiten) sowie virtuellen Forschungsumgebungen (mit integrierten Funktionen z. B. für die Datenanalyse) ab.

Es lassen sich grundlegend zwei Typen von Repositorien unterscheiden: Disziplinspezifische Repositorien sammeln die Inhalte einer bestimmten Forschungsdisziplin (Suche nach Disziplin möglich über [re3data](https://www.re3data.org/)), während generische Repositorien Inhalte unterschiedlicher Disziplinen aufnehmen (z. B. [Zenodo](https://zenodo.org)). Einen Sonderfall fachübergreifender Repositorien bilden institutionelle Repositorien, die speziell für die Mitglieder der jeweiligen Institution zur Verfügung stehen. Einige Repositorien widmen sich zudem speziell der Sammlung bestimmter Datentypen (z. B. Preprint- oder Publikationsserver, Forschungsdatenrepositorien, Bilddatenbanken, …). Insbesondere im Bereich Forschungsdaten und Open Educational Resources gibt es seitens der Nutzer\*innen häufig den Wunsch, dass in Repositorien auch eine interaktive Arbeit mit den Materialien möglich ist. Derzeit sind Repositorien aber in der Regel noch ausschließlich auf die sichere Speicherung und Bereitstellung der Inhalte ausgerichtet - eine Erweiterung um Funktionalitäten z. B. zur Visualisierung oder Analyse würden sie in die Nähe von virtuellen Forschungsumgebungen rücken. Häufig werden Repositorien und Repositoriensoftware zudem mit der digitalen [Langzeitarchivierung](#langzeitarchivierung) als Kernfunktionalität in Verbindung gebracht - aus Sicht von Repositorien-Betreiber\*innen und -Entwickler\*innen handelt es sich bei der digitalen Langzeitarchivierung allerdings eher um eine Spezialfunktionalität, die nur für bestimmte Nutzungsszenarien relevant und daher auch nicht in allen Repositorien gegeben ist. Mit dem proprietären System [Rosetta](https://exlibrisgroup.com/de/produkte/rosetta/) und der Open Source-Lösung [Archivematica](https://www.archivematica.org) seien hier nur zwei Beispiele für Systeme genannt, die sich auf den Erhalt von Informationen im Sinne der Langzeitarchivierung spezialisieren.

|  |
| --- |
| Definition |
| Der Begriff **Repository** bzw. **Repositorium** wird in diesem Kapitel vorwiegend verwendet für Plattformen, in denen Forschungsergebnisse dauerhaft archiviert, beschrieben, auffindbar und zugänglich gemacht werden. Darüber hinaus wird der Begriff Repository aber auch für Versionsverwaltungssysteme wie Git verwendet, die vorwiegend von Softwareprojekten genutzt werden, um es einer geschlossenen oder offenen Community zu ermöglichen, transparent und konfliktfrei zur Codebasis des jeweiligen Projekts beizutragen. Unter anderem [Zenodo](https://www.zenodo.org) erlaubt über eine Schnittstelle zum Code-Verwaltungssystem GitHub die Archivierung von Softwareprojekten und ähnlichem entsprechend den Konventionen von Repositorien für Forschungsergebnisse, einschließlich der Vergabe von DOIs. |

Das *Open Directory of Open Access Repositories (OpenDOAR)* stellt [Statistiken über die Verbreitung von Softwarelösungen](https://v2.sherpa.ac.uk/view/repository_visualisations/1.html) für Repositorien für Forschungsergebnisse bereit: Meistgenutztes System weltweit ist [*DSpace*](https://www.dspace.org), das auch in Deutschland zunehmend Verbreitung findet. Während DSpace vor allem von Universitäten eingesetzt wird, ist bei Fachhochschulen und HAWen [*OPUS*](https://www.opus-repository.org/) stark verbreitet, das meist durch den Kooperativen Bibliotheksverbund Berlin-Brandenburg (KOBV) oder das Bibliotheksservice-Zentrum Baden-Württemberg (BSZ) gehostet wird. Der Gemeinsame Bibliotheksverbund (GBV) bietet mit [*MyCoRe*](http://www.mycore.org) eine weitere Repositoriensoftwarelösung an. *OPUS* und *MyCoRe* finden bislang ausschließlich im deutschsprachigen Raum Anwendung, während DSpace von einer globalen Community getragen wird. Für Forschungsdatenrepositorien kommt zum Teil *DSpace*, zum Teil die auf Forschungsdaten spezialisierte Software [*Dataverse*](https://dataverse.org) zum Einsatz.

Referenzen

* Beate Rajski, Pascal Becker, "DSpace-Konsortium Deutschland – konsortial die Nachhaltigkeit sichern". *LIBREAS. Library Ideas*, 36 (2019). <https://libreas.eu/ausgabe36/rajski/>
* [Repositorien und das Semantic Web](https://doi.org/10.14279/depositonce-5015)

## Forschungsdatenmanagement

Die digitale Transformation hat Forschungsprozesse grundlegend verändert: In zahlreichen Fachdisziplinen entstehen an Forschungseinrichtungen heute täglich große Mengen digitaler Daten, die als Forschungsgegenstand dienen, angereichert, analysiert oder visualisiert werden. Dabei stehen Wissenschaftler\*innen vor der Herausforderung, diese Daten nicht nur zu verwalten, sondern sie auch langfristig und nachvollziehbar vorzuhalten und möglichst offen zur Nachnutzung zur Verfügung zu stellen. Maßgeblich sind hierfür die sogenannten **FAIR-Prinzipien**, denen zufolge Forschungsdaten auffindbar (findable), zugänglich (accessible), interoperabel (interoperable) und nachnutzbar (reusable) sein sollen (vgl. <https://doi.org/10.1038/sdata.2016.18> ).

|  |
| --- |
| Definition |
| **Forschungsdaten** sind alle digital vorliegenden Daten, die während des Forschungsprozesses entstehen (z. B. Messdaten, Laborwerte, Videoaufnahmen, Umfrageergebnisse). Klar davon abzugrenzen sind [Forschungsinformationen](#forschungsinformationssysteme). Der Lebenszyklus von Forschungsdaten beinhaltet die Erstellung, Speicherung, Archivierung bis hin zur Löschung aussortierter Daten. |

|  |
| --- |
| Abbildung 9.1: Darstellungen des Forschungsdatenlebenszyklus in unterschiedlichen Detailgraden helfen bei der Betrachtung der benötigten Werkzeuge für die forschungsnahen Dienste. |

Die Schnittmenge von Forschungsdaten und Forschungsinformationen liegt wie in [Abbildung 9.2](#fig-fis-fdm) dargestellt im Bereich Publikationen. Während sich allerdings Forschungsinformationen eher auf klassische, kontrolliert publizierte Dokumente beziehen, geht die Publikation von Forschungsdaten weit darüber hinaus und schließt jegliche Form von Aufzeichnungen wie Notizen und Zwischenergebnisse bis zu [Forschungssoftware](#forschungssoftware) mit ein.

|  |
| --- |
| Abbildung 9.2: Forschungsinformationen und ihre Sicht auf Forschungsdaten (CC-BY Franziska Mau) |

Services zum Forschungsdatenmanagement sollen Wissenschaftler\*innen beim Umgang mit ihren Forschungsdaten unterstützen, und zwar über den gesamten Forschungsdatenlebenszyklus hinweg, d. h. von der Datenplanung über die Datenerhebung und -analyse bis hin zur Datenarchivierung, -publikation und -nachnutzung. Bibliotheken nehmen beim Aufbau und Betrieb entsprechender Services eine zentrale Rolle ein - in aller Regel sind sie hierbei jedoch nicht die einzigen Akteure, sondern das Serviceportfolio wird in einer arbeitsteiligen Zusammenarbeit von Bibliotheken, Rechenzentren, Forschungsabteilungen und ggf. weiteren Akteur\*innen angeboten. Angesichts der großen Heterogenität an (disziplinspezifischen) Datentypen gelangen diese in aller Regel fachübergreifenden FDM-Dienste allerdings häufig an ihre Grenzen: Diese Erkenntnis ist konstitutiv für die seit 2020 im Aufbau befindliche Nationale Forschungsdateninfrastruktur ([NFDI](https://www.nfdi.de/)), in der fachspezifische, institutionsübergreifende Dienste entwickelt werden. In diesem Zusammenhang entwickeln sich derzeit auch neue Berufe wie Data Steward oder Data Librarian, die fachspezifische Unterstützung beim Forschungsdatenmanagement leisten und entweder zentral an den FDM-Servicestellen oder dezentral in Projekten oder Fachbereichen angesiedelt sind.

Die von Bibliotheken angebotenen Services zum Forschungsdatenmanagement umfassen in der Regel sowohl nicht-technische Services (z. B. Schulungs- und Beratungsangebote) als auch verschiedene technische Dienste. Zu den wichtigsten technischen Diensten für das Forschungsdatenmanagement, die von Bibliotheken (mit-)betrieben werden, gehören Forschungsdatenrepositorien. Diese ermöglichen die Veröffentlichung von Forschungsdaten als eigene Informationsobjekte gemäß den FAIR-Prinzipien (TODO: siehe hierzu ausführlich den Abschnitt [Repositorien für Forschungsergebnisse](#repositorien-für-forschungsergebnisse)). Daneben werden häufig weitere FDM-Tools angeboten, von denen einige im Folgenden vorgestellt werden sollen.

### Tools zur Erstellung von Datenmanagementplänen

Um den Umgang mit Forschungsdaten über ein komplettes Projekt zu beschreiben, hat sich der **Datenmanagementplan (DMP)** als geeignetes Format erwiesen. Derartige Pläne werden zunehmend von Forschungsförderern bei der Antragstellung oder in der Frühphase des Projekts erwartet und basieren häufig auf für die Förderlinie bzw. das Fachgebiet spezifischen Fragenkatalogen. Durch die einheitlichen Fragelisten und fest definierte Ausgabeformate lässt sich so ein menschen- und maschinenlesbares Dokument generieren. Um diese Fragenkataloge einheitlich zur Verfügung zu stellen und ggf. mit einrichtungs- oder programmspezifischen Daten anzureichern, wurden bereits einige Software-Werkzeuge entwickelt. In Deutschland verbreitet ist der aus einem DFG-Projekt entstandene [Research Data Management Organizer](https://rdmorganiser.github.io/) (RDMO), für den, auch unterstützt durch NFDI-Konsortien, ständig neue Fragenkataloge in einer Gemeinschaftsarbeit entwickelt werden. Obwohl einige öffentliche Instanzen der Software existieren, die z. B. einen Login über die ORCID ermöglichen, kann die Open-Source-Software auch selbst gehostet und inhaltlich sowie visuell auf die Bedarfe der jeweiligen Einrichtung zugeschnitten werden. DMP können innerhalb dieser Software auch kollaborativ erstellt werden, indem Personen als Mitarbeitende in das eigene Projekt eingeladen werden. Für EU-Projekte ist mit Stand 2023 die Software [ARGOS](https://argos.openaire.eu/) verfügbar, die eine direkte Einbindung der DMP in die [EOSC](https://eosc-portal.eu/) ermöglicht. Für Software als Forschungsdatum beginnen sich Softwaremanagementpläne zu etablieren.

### Elektronische Laborbücher

Die Dokumentation der Forschungsergebnisse ist ein zentraler Punkt im Forschungsdaten-Lebenszyklus. Digital entstandene Daten sollten ohne Medienbruch mit ihrer Dokumentation verknüpft und zugehörige Metadaten erfasst werden. Hierzu gibt es eine Reihe dedizierter Software-Lösungen, die unter dem Begriff elektronische Laborbücher (abgekürzt häufig *ELN* von *electronic laboratory notebook*) zusammengefasst werden. Neben kommerziellen Programmen wie [LabFolder](https://labfolder.com/de/), bei denen Bibliotheken eher in der Rolle des Vermittlers sind, gewinnen Open-Source-Lösungen zunehmend an Bedeutung. Diese können Bibliotheken auf eigenen Servern hosten und selbst administrieren, sind jedoch bei der Entwicklung neuer Features auf eine Community bzw. den Entwickler oder eigene Fachkräfte angewiesen. Beispiele sind hier [eLabFTW](https://www.elabftw.net/) für generische ELN bzw. [Chemotion](https://chemotion.net/) für eine eher fachspezifische Lösung. Eine [Handreichung zur Einführung eines ELN](https://doi.org/10.17192/bfdm.2023.5.8553) an der eigenen Einrichtung wurde 2023 von einer einrichtungsübergreifenden Autor\*innengruppe erstellt. Open-Source-ELN erfahren häufig auch Unterstützung durch NFDI-Konsortien. Der auch daraus entstandene Wunsch eines einheitlichen Transferformats der Laborbucheinträge und gemeinsamer Spezifikationen wird im [ELN Consortium](https://github.com/TheELNConsortium) adressiert. Hilfestellung bei der Auswahl eines passenden Produkts bietet z. B. der [ELN-Finder](https://eln-finder.ulb.tu-darmstadt.de/home).

### Git

Als freie Software zur Versionsverwaltung ist Git ein Standardtool der Softwareentwicklung geworden. Durch einfache Befehle auf der Kommandozeile oder zusätzlich installierte Software mit grafischem Interface lassen sich textuelle Daten auf dem eigenen System mit einem externen (Code-)Repositorium abgleichen, das die notwendigen Protokolle versteht. Neben dem großen Anbieter GitHub gibt es die lokal zu installierende Software GitLab, um ein solches Repositorium in der eigenen IT-Infrastruktur bereitzustellen. Durch die im Protokoll integrierte Versionskontrolle der Daten lassen sich Änderungen im Code einfach nachvollziehen und gegebenenfalls zurückrollen. Kollaborative Arbeit in verteilten Teams zum Beispiel wird über eine parallele Entwicklungsstruktur in „branches“ ermöglicht, die mit dem Hauptprojekt zu einem gewünschten Zeitraum zusammengeführt werden können. Bei diesem Funktionsumfang wird schnell klar, dass Git auch jenseits der forschungsnahen Dienste eine Vielzahl von Anwendungsmöglichkeiten hat. Hierfür sei auf das entsprechende Kapitel verwiesen. Eine Möglichkeit, die Funktion der Software spielerisch zu erkunden, bietet beispielsweise die vom Bundesministerium für Bildung und Forschung geförderte Webseite [ohmygit.org](https://ohmygit.org).

[*Forschungssoftware*](#forschungssoftware) *lässt sich, nicht nur wegen der Verwaltung mit Git oder ähnlichen Programmen, nicht einfach komplett analog zu den Forschungsdaten behandeln, sondern bedarf eines genaueren Blicks.*

## Forschungssoftware

Bei der Betrachtung von Forschungsprozessen setzt sich zunehmend die Erkenntnis durch, dass auch die dabei zum Einsatz kommende Software ein Teil der Forschungsdaten ist. Dies ist häufig kein kommerziell erhältliches Produkt, sondern ein speziell auf das Forschungsproblem zugeschnittener, selbst programmierter Code. Wird derartige Software nicht korrekt gesichert, versioniert und dokumentiert, leidet die Reproduzierbarkeit von Forschungsdaten. Komplexe externe Probleme wie prekäre Beschäftigungsverhältnisse können zusätzlich zum Verwaisen von Softwareprojekten führen, wenn diese nur lokal durch einzelne engagierte Personen vorangetrieben wurden. Zusätzlich führt eine Veröffentlichung der Forschungssoftware zu einer Auffindbarkeit und Zitierbarkeit, so dass diese zum wissenschaftlichen Output der Forschenden einen signifikanten Beitrag leisten kann. Aus der Wissenschaft getriebene Vereinigungen wie [de-RSE e. V.](https://de-rse.org/) treten als Vereinszweck für den Stellenwert von Forschungssoftware ein. Auch die FAIR-Prinzipien sollten für Forschungssoftware Anwendung finden (vgl. <https://doi.org/10.1038/s41597-022-01710-x>). Hier liegt es auch an den Bibliotheken, einerseits ein Bewusstsein dafür zu schaffen (z. B. durch dedizierte Policies), aber andererseits auch die benötigte Infrastruktur bereitzustellen.

Zur Zitierbarkeit von Forschungssoftware/Code dient die Generierung von entsprechenden Metadaten, etwa über <https://codemeta.github.io/> und [CITATION.cff](https://citation-file-format.github.io/), um menschen- und maschinen-lesbare Zitierinformationen für Software und Datensätze angeben zu können.

*TODO INFOBOX: CITATION.cff des Handbuchs als Beispiel*

Die Bereitstellung eines Coderepositoriums, in der Regel über Git, sollte zum Standardangebot zählen. Für eine interaktive wissenschaftliche Datenauswertung bietet sich zusätzlich der Betrieb eines [*JupyterHub*](https://jupyter.org/hub) an. Dieser ermöglicht die Nutzung von Jupyter Notebooks auf einem zentralen Server der Einrichtung und ist so nicht abhängig von den jeweiligen Rechenleistungen der verfügbaren Endgeräte. Zusätzlich sollte geprüft werden, inwiefern eine Archivierung der Software bzw. eines gesamten Repositoriums in der eigenen Infrastruktur nötig und sinnvoll ist. Anbieter wie [Software Heritage](https://www.softwareheritage.org/) bieten eine Archivierung auf externen Servern an. Zumindest in diesem Fall ist es als UNESCO-Projekt aber als geeignete Alternative zu betrachten.

Die Verknüpfung der im Gesamtprozess entstehenden Metadaten mit Systemen wie dem FIS oder auch Forschungsdatenrepositorien hinsichtlich der Auffindbarkeit dieser wissenschaftlichen Ergebnisse ist ein Punkt, der eine Betrachtung der kompletten [Toolchain](#toolchains) notwendig macht.

## Forschungsinformationssysteme

|  |
| --- |
| Definition |
| **Forschungsinformationen** sind Angaben über Aktivitäten, Ergebnisse und Infrastrukturen von Forschungsprozessen wie zum Beispiel Projekte, Publikationen und Forschungseinrichtungen. Davon zu unterscheiden sind [Forschungsdaten](#forschungsdatenmanagement). |

Neben Forschungsdaten gewinnt auch die strukturierte Erfassung von Forschungsinformationen an Bedeutung. Entsprechende Systeme werden **Forschungsinformationssysteme** (FIS) genannt. Dabei handelt es sich um Datenbanksysteme, die speziell für die Erfassung, Organisation, Speicherung und Verknüpfung von **Forschungsinformationen** konzipiert wurden. Sie können interne Anwendungen wie die leistungsorientierte Mittelvergabe unterstützen, aber auch für die Außendarstellung der Einrichtung genutzt werden. Eine Übersicht von Forschungsinformationen und ihre Sicht auf Forschungsdaten gibt [Abbildung 9.2](#fig-fis-fdm).

FIS führen Informationen zusammen, die dezentral in verschiedenen hochschulinternen Systemen (z. B. Drittmittelverwaltung, Personalverwaltungssysteme, Repositorien) und externen Quellsystemen (z. B. Scopus, ORCID) vorgehalten werden, um einen strukturierten und aktuellen Überblick über die Forschungsleistungen z. B. einer Einrichtung, eines (Bundes-)Landes oder einer Fachdisziplin zu gewinnen.

Die genauen Daten, die Nutzung der Daten und der Funktionsumfang eines FIS sind nicht festgelegt bzw. klar definiert. Verschiedene Softwarelösungen verfolgen unterschiedliche Ansätze: Einige legen den Schwerpunkt auf die Auffindbarkeit und Verknüpfung von Forschenden, andere Systeme haben ihren Schwerpunkt eher auf dem Berichtswesen und Monitoring und ggf. darauf basierenden Anreizsystemen. Wieder andere Systeme legen den Schwerpunkt darauf, die Forschungsaktivitäten zu präsentieren und öffentlichkeitswirksam bereitzustellen. Die Systeme passen sich zunehmend aneinander an, oft werden verschiedene Systeme aber auch in Kombination miteinander eingesetzt.

FIS sollten von Anfang an als Daueraufgabe einer Einrichtung betrachtet und entsprechende finanzielle und personelle Ressourcen eingeplant werden. Bei der Einführung eines FIS handelt es sich um ein langjähriges Organisationsentwicklungsprojekt, das eine Offenheit für Veränderungen in den Prozessen und Workflows der Einrichtung voraussetzt.

|  |
| --- |
| Herausforderungen beim Aufbau eines Forschungsinformationssystem (CC-BY Franziska Mau) |

Eine zentrale Herausforderung beim Aufbau eines FIS besteht darin, einen Überblick über die bestehenden Quellsysteme der Einrichtung zu gewinnen. In diesem Zusammenhang ist zum einen zu ermitteln, welche internen und externen Systeme relevant sind und wer die entsprechenden Ansprechpartner\*innen an der Einrichtung sind. Dies betrifft u. a. die Bibliothek (z. B. Repositorien), die Personalverwaltung (Identitätsmanagement), die Drittmittelverwaltung (Datenbank für Projekte), die Doktorand\*innenverwaltung oder die Patentverwaltung der Einrichtung.

Neben der Identifikation der relevanten Datenquellen stellt die Integration der Daten in das FIS meist die größte Herausforderung dar. So muss zum einen für fehlende oder ungeeignete Schnittstellen eine Lösung gefunden werden. Zum anderen variieren Qualität und Konsistenz der vorhandenen Daten mitunter stark, was zusätzliche Zeit für die Datenbereinigung und -konvertierung erfordert. Gleichzeitig ist die Sicherstellung der Datenintegrität und -qualität von entscheidender Bedeutung, um zu gewährleisten, dass das FIS korrekte und aussagekräftige Informationen liefert.

Der Markt für FIS-Software ist sehr dynamisch. Vor dem Hintergrund, dass sich gerade viele Forschungseinrichtungen in der Planungs- und Aufbauphase von Forschungsinformationssystemen befinden, kommen in Deutschland immer neue Softwarelösungen zum Einsatz. Es zeigt sich ein vielgestaltiges Bild aus kommerziellen Produkten (z. B. *PURE*, *Converis*, HISinOne-RES), Open Source-Lösungen (z. B. *DSpace-CRIS*, *VIVO*) und Eigenentwicklungen. An deutschen Forschungseinrichtungen wird mittlerweile häufig *HISinOne-RES* genutzt - befördert unter anderem durch Landesinitiativen wie *CRIS.NRW*, *HeFIS* oder *FIS-Thüringen* sowie den Umstand, dass es aktuell das einzige Produkt am Markt ist, dessen Datenmodell direkt am **Kerndatensatz Forschung (KDSF)** ausgerichtet ist. Obwohl sich ein Rückgang an Eigenentwicklungen andeutet, sind sie immer noch weit verbreitet. Des Weiteren gibt es die bereits lange etablierten kommerziellen Systeme *Converis* und *PURE*. Der Einsatz von Open Source-Lösungen wie *DSpace-CRIS* und *VIVO* nimmt erst in den letzten Jahren merklich zu – unter anderem befördert durch das Verbundprojekt Hamburg Open Science.

An vielen Einrichtungen besteht das Bestreben, dass das FIS zusätzlich auch die Funktionalität eines [Repositoriums](#repositorien-für-forschungsergebnisse) übernehmen soll. Ein Vorteil eines solchen vereinigten Systems wird zum einen in den geringeren Systemkosten gesehen - zum anderen erscheint es weniger aufwendig, die bibliographischen Einträge in einem FIS schlicht mit den dazugehörigen Dateien anzureichern statt einen Workflow für das Zusammenspiel zwischen FIS und Repositorium zu entwickeln. Dem entgegen stehen jedoch die verschiedenen Zielsetzungen beider Systeme: Während es bei einem FIS vor allem darum geht, möglichst alle Forschungsaktivitäten z. B. einer Einrichtung in einem System zu erfassen, steht bei einem Repositorium die nachhaltige Bereitstellung der Ressourcen selbst im Vordergrund (z. B. textuelle Publikationen oder Forschungsdaten). Ein Problem bei Mischsystemen ergibt sich auch hinsichtlich Retrieval und Zugriff: So werden Forschende bei einer Suche in externen Suchmaschinen z. B. erst im FIS feststellen, dass nur bei einem Teil der Treffer tatsächlich Zugang zu den Ressourcen selbst besteht, sie in den meisten Fällen jedoch lediglich Nachweise der Ressourcen finden. In der Praxis sind FIS-Repositorien-Mischsysteme derzeit dennoch aufgrund von Ressourcenknappheit nicht wegzudenken.

Um eine Interoperabilität der unterschiedlichen Systeme und eine gute Auffindbarkeit der enthaltenen Ressourcen zu ermöglichen, ist eine Standardisierung notwendig - z. B. über Zertifikate, Metadatenstandards und Schnittstellen. Die in diesem Zusammenhang wichtigen Grundlagen werden im folgenden Kapitel erläutert.

## Übergreifende Themen

### Zertifikate und Standards

Forschungsnahe Dienste bewegen sich von jeher an der Schnittstelle zwischen Wissenschaft und Infrastruktureinrichtungen, die die Dienste betreiben. Zertifikate erfüllen in diesem Spannungsfeld verschiedene Funktionen. Sie waren als Vertrauen gebende Maßnahmen angedacht, die Qualitätsmerkmale der Dienstleistungen sein sollten. Das [DINI-Zertifikat für Open Access Publikationsplattformen](https://dini.de/dienste-projekte/dini-zertifikat) versteht sich seit jeher auch als Ratgeber bei Einrichtung, Weiterentwicklung und Betrieb solcher Dienstleistungen, der „Maßstäbe, Richtlinien und Best Practices“ vermitteln will. Und letztlich dienen Zertifikate auch dem Schaffen von Standards, die die Interoperabilität der Dienste ermöglichen. Neben dem DINI-Zertifikat sind in Bezug auf forschungsnahe Dienste auch das [Core Trust Seal](https://www.coretrustseal.org) sowie das [Nestor-Siegel für vertrauenswürdige digitale Langzeitarchive](https://www.langzeitarchivierung.de/Webs/nestor/DE/Zertifizierung/nestor_Siegel/siegel.html) zu nennen. Während es mit der DIN Norm 31644 (auch als ISO Norm 16363 verbreitet) „Information und Dokumentation - Kriterien für vertrauenswürdige Langzeitarchive“ eine offizielle Norm für die Bewertung der Vertrauenswürdigkeit von Langzeitarchiven gibt, werden die meisten Standards in diesem Bereich eher als Best Practices oder Konventionen, denn als offizielle Normen eingeführt. Unabhängig von der Frage, ob Zertifikate als vertrauensstiftend eingeschätzt werden, lohnt es sich die Dokumentation der Zertifikate als Ratgeber oder Checkliste zu nutzen, sowohl beim Aufbau neuer Dienste, als auch zur regelmäßigen Überprüfung des eigenen Dienstes mit Blick auf neue Entwicklungen und Optionen eigene Dienste weiterzuentwickeln.

Schon im [Bethesda Statement on Open Access Publishing](http://legacy.earlham.edu/~peters/fos/bethesda.htm) taucht das Stichwort Interoperabilität in Zusammenhang mit Repositorien auf. Dazu gibt es verschiedene technische Ansätze (siehe unter anderem unten „[Schnittstellen](#schnittstellen)“). Neben den technischen Voraussetzungen, um Inhalte zu teilen, braucht es jedoch auch eine Einigung über die inhaltliche Aufbereitung der Informationen. Repositorien nutzen dazu strukturierte Metadaten. Für die Bezeichnung von Dokumententypen haben die DINI AG Elektronisches Publizieren und die DINI AG Forschungsinformationssysteme das [Gemeinsame Vokabular für Publikations- und Dokumenttypen](https://doi.org/10.18452/24147) herausgegeben. Im Sinne der Standardisierung enthält das DINI-Zertifikat weitere Vorgaben, wie zum Beispiel die Klassifizierung nach zumindest den DDC-Sachgruppen der Deutschen Nationalbibliografie und macht Vorgaben an die Ausgestaltung der OAI-PMH-Schnittstelle. Diese Standards ermöglichen es Diensten wie zum Beispiel der Bielefeld Academic Search Engine und anderen Aggregatoren Inhalte aus vielen verschiedenen Quellen einzubinden, Metadaten maschinenlesbar zu erhalten und nachzunutzen.

Referenzen

* [DINI-Zertifikat 2022 für Open Access Publikationsdienste](https://doi.org/10.18452/24678)
* <https://www.coretrustseal.org/>
* <https://www.langzeitarchivierung.de/Webs/nestor/DE/Zertifizierung/nestor_Siegel/siegel.html>

### Metadaten

Metadaten sind Daten struktureller, technischer, administrativer, bibliographischer und deskriptiver Natur, die Daten beschreiben (TODO: siehe Kap. **Daten & Metadaten** bzw. *Metadatenstandards*)*.* Metadaten werden oft in einer Schlüssel-Wert-Struktur genutzt, bei der der Schlüssel vorgibt, welche Angabe (z. B. Titel, Autorschaft, Erscheinungsdatum, …) im Wert zu finden ist (siehe Kap. **Daten & Metadaten** bzw*. Grundlegende Begrifflichkeiten*). Strukturell werden flache und hierarchische Metadatenschemata unterschieden. Flache Metadatenschemata beschränken sich auf eine einfache Struktur aus Schlüssel-Wert-Paaren. Hierarchische Metadatenschemata sehen vor, Werte aus anderen Werten zusammensetzen zu können, so dass z. B. die Autorenschaft in den Metadaten über Personen modelliert werden können und zu jeder Person Vorname, Nachname und weitere Angaben in untergeordneten Werten gespeichert wird. Hierarchische Metadatenschemata haben dabei oft etliche dieser Hierarchieebenen und Verschachtelungen.

Metadatenschemata definieren, welche Inhalte in den Metadaten erfasst werden, also welche Metadatenfelder existieren und mit Werten belegt werden können. [Dublin] [Core Element](https://www.dublincore.org/specifications/dublin-core/usageguide/elements/) (umgangssprachlich oft einfach *Dublin Core* genannt) ist als Metadatenschema für Inhalte im Internet entstanden (TODO: siehe Kap. **Daten & Metadaten** bzw. *Metadatenstandards*) . Es wird oft in Metatags auf HTML-Seiten verwendet und ist vorherrschend beim Austausch von Daten zwischen und mit Repositorien, auch wenn so gut wie jede Repositoriensoftware intern deutlich mehr Metadatenschemata unterstützt und nutzt. Es wird zwischen Simple und Qualified Dublin Core unterschieden. Simple Dublin Core besteht aus 15 Elementen, in Qualified Dublin Core können ergänzend Qualifier genutzt werden. So kann zum Beispiel das Element dc.contributor weiter spezifiziert werden als dc.contributor.author, dc.contributor.translator, dc.contributor.illustrator und so weiter. Von der Dublin Core Metadata Initiative wurde später auch das Metadatenschema DCMI Terms geschaffen, das Simple und Qualified Dublin Core zusammenfassen sollte. In der Praxis wird nach wie vor häufig Dublin Core Elements genutzt, in Schnittstellen oft ohne Qualifier. DCMI Terms werden in Repositorien dann oft ergänzend zu Dublin Core Elements verwendet.

Im Rahmen der DOI-Registrierung werden auch Metadaten erhoben. Das [DataCite Schema](https://schema.datacite.org) hat sich dabei als ein wichtiges Metadatenschema etabliert, das zum Teil auch losgelöst von DOIs zur Beschreibung von Forschungsdaten genutzt wird (TODO: siehe Kap. **Daten & Metadaten** bzw. *Metadatenstandards*).

Darüber hinaus entwickeln die verschiedenen Fachdisziplinen eigene domänenspezifische Standards (Übersicht: <https://fairsharing.org/search?fairsharingRegistry=Standard> ). Eine Aufgabe wissenschaftlicher Bibliotheken besteht in der Beratung bei der Auswahl geeigneter Metadatenschemata und Standards (Übersicht unter: <https://www.forschungsdaten.info/themen/beschreiben-und-dokumentieren/metadaten-und-metadatenstandards>-

### Persistent Identifier

Mit dem Aufkommen elektronischer Archive kam die Frage nach der Zitierbarkeit auf. Auch wenn die technischen Protokolle, auf denen das Internet basiert, sowohl in DNS als auch http(s) Mechanismen enthalten, um URLs weiterzuleiten, bekamen URLs schnell den Ruf flüchtig zu sein. Auch wenn URLs, die nicht mehr oder auf andere Inhalte auflösen, immer auf Managementprobleme zurückgehen, wurden Persistent Identifier Systeme geschaffen, die diese Probleme beim Zitieren elektronischer Quellen überwinden sollen. Dabei werden IDs geschaffen, die über einen sogenannten Resolver aufgelöst werden können. Der Resolver ist vergleichbar mit einem Melderegister: man kann nach der aktuellen Adresse einer ID fragen und erhält die jeweils aktuelle URL zurück, unter der sich die Ressourcen befinden sollen. Die Antworten können also zu unterschiedlichen Zeitpunkten unterschiedlich ausfallen.

Während die Deutsche Nationalbibliothek bis heute auf URN:NBN als Persistent Identifier setzt, haben sich im wissenschaftlichen Umfeld DOIs für die Identifikation von Artikeln, Daten und anderen Inhalten durchgesetzt. In Deutschland kommen dabei vor allem die DOI-Registrierungsagenturen [DataCite](https://datacite.org) und [CrossRef](https://www.crossref.org/) zum Einsatz. Beide vergeben DOIs sowohl für textuelle Publikationen als auch für Datensätze und andere Inhalte. Von der technischen Einbindung her ist DataCite moderner aufgestellt und leichter zu integrieren. Mit [ORCID](https://orcid.org/) und [ROR](https://ror.org) gibt es inzwischen weitere Persistent Identifier Systeme, die zunehmend Verbreitung finden und Personen bzw. Einrichtungen eindeutig identifizieren.

Die Vergabe von PIDs für Publikationen (z. B. Texte, Forschungsdaten) auf Publikationsservern bzw. Datenrepositorien wird teilweise von den wissenschaftlichen Bibliotheken gewährleistet. Somit werden Forschungsdaten nachhaltig unter entsprechenden Lizenzen öffentlich verfügbar gemacht (Berg-Weiß et al. 2022). Mit dem Ziel, eine nationale Beratungs- und Austauschplattform zu persistenten Identifikatoren aufzubauen, fördert die DFG seit 2023 das Projekt „[PID Network Deutschland](https://www.pid-network.de/) - Netzwerk für die Förderung von persistenten Identifikatoren in Wissenschaft und Kultur“.

### Langzeitarchivierung

Die dauerhafte Aufbewahrung und Lesbarkeit von digitalen Objekten zu gewährleisten, stellt auch für Bibliotheken, die zunehmend für die Archivierung von Open-Access-Publikationen, Forschungsdaten und andere elektronische Ressourcen verantwortlich sind, eine große Herausforderung dar. Die sogenannte **digitale Langzeitarchivierung (LZA)** beinhaltet neben der Speicherung zusätzliche Maßnahmen wie die regelmäßige Überprüfung der Datenintegrität, die Migration der Daten auf neue Speichermedien und die Anpassung an sich verändernde Technologien. Digitale Informationen (z.B. Forschungsdaten, digitalisiertes kulturelles Erbe) bleiben so langfristig erhalten und auch in der Zukunft zugänglich.

Bewahrt werden müssen also sowohl der Bitstream der Datei, als auch deren Eigenschaften und Semantik. Von großer Bedeutung ist zudem die Dokumentation, was mit den digitalen Objekten im Verlauf der LZA passiert ist. Zu den Aktivitäten und Prozessen werden daher spezielle Metadaten erfasst. Aktuell ist [PREMIS](https://www.loc.gov/standards/premis/) in der LZA der wichtigste Metadatenstandard. Das Datenmodell beinhaltet alle Informationen, die man sowohl über die digitalen Objekte selbst (z. B. Name, Dateiformat, Größe) als auch über Akteure, Rechte (z. B. AccessRights, Embargofristen) und Prozesse (z. B. Konvertierung, Migrationen, Reparatur, Formatvalidierung) wissen sollte.

Es gibt auf LZA spezialisierte Software wie Rosetta (ExLibris) oder Libsafe (libnova), die mittlerweile in vielen Bibliotheken, Museen und anderen kulturellen Einrichtungen weltweit zum Einsatz kommen. Diese Systeme basieren meist auf dem international anerkannten Referenzmodell für digitale Archivierung OAIS ([Open Archival Information System, ISO 14721:2012](https://www.iso.org/standard/57284.html)) und bieten neben den Standardfunktionen eines Archivsystems (z. B. bitstream-preservation, regelmäßige Integritätstests, Reduplizierung) auch Funktionen wie eine Format-Validierung und implementierbare Workflows.

Auf der Webpage [COPTR](https://coptr.digipres.org/) - Community Owned digital Preservation Tool Registry - werden diese und zahlreiche weitere Tools und Workflows zur Langzeitarchivierung vorgestellt. Als wichtige Anlaufstelle für Fragen rund um die digitale Langzeitarchivierung dient außerdem das [Kompetenznetzwerk nestor](https://www.langzeitarchivierung.de/), dessen Geschäftsstelle an der Deutschen Nationalbibliothek angesiedelt ist. Auch die NFDI behandelt „[Long-term Archival (LTA)](https://doi.org/10.5281/zenodo.6451456)“ in der [Sektion „Common Infrastructures“](https://www.nfdi.de/section-infra/) als Querschnittsthema.

|  |
| --- |
| Das als ISO 14721 verabschiedete Referenzmodell „Open Archival Information System” (OAIS), CC-BY Jørgen Stamp |

### Schnittstellen

Im Bereich von Repositorien hat sich das [Open Archives Initative Protocol for Metadata Harvesting](https://www.openarchives.org/pmh/) (OAI-PMH) für den Austausch von Metadaten durchgesetzt. Dieses Protokoll wird inzwischen auch im Zusammenspiel mit anderen forschungsnahen Diensten wie zum Beispiel FIS genutzt. Das Protokoll tauscht Metadaten in XML aus. Es unterstützt mehrere Metadatenformate, wobei die Spezifikation von OAI-PMH nur Dublin Core vorgibt und das Protokoll vorsieht, dass man eine Liste mit weiteren unterstützten Formaten abrufen kann.

Für das Einbringen von Daten in Repositorien hat sich das Protokoll [Simple Webservice Offering Repository Deposit](https://swordapp.org/) (SWORD) durchgesetzt, wobei auf die genaue Version dieses Standards geachtet werden muss. Einige Open-Access-Verlage bieten an, Dokumente über SWORD 1 direkt in Repositorien zu übertragen. [DeepGreen](https://info.oa-deepgreen.de/), ein Lieferdienst für Open-Access-Artikel, versorgt Repositorien über *SWORD* mit Verlagsinhalten.

Speziell für die Arbeit mit Bildern, Bildviewern und Bilddatenbanken wurde das *International Image Interoperability Framework (IIIF)* entwickelt. IIIF deckt umfangreich verschiedene Funktionen ab, wie die Ausgabe von Bildern in verschiedenen Formaten und Auflösungen oder Zoomstufen, die strukturelle Beschreibung von Bildern, Suchanfragen in einem Bildpool, Umgang mit Zugriffsbeschränkungen, Objektänderungen und so weiter.

OAI-PMH und SWORD werden zum Teil zwar auch über Repositorien hinaus verwendet, zum Beispiel bei Forschungsinformationssystemen. Verbreitung und Einsatz beschränken sich jedoch weitestgehend auf den Bereich der forschungsnahen Dienste. Als weiter verbreitete Prinzipien zur Bereitstellung von Informationen muss Linked Data gesehen werden. Die Bereitstellung der Repositorieninhalte als Linked Data wird zum Beispiel von DSpace unterstützt, in der Praxis jedoch leider noch viel zu selten aktiviert und genutzt.

Allgemein verbreiten sich derzeit REST-Schnittstellen, also Schnittstellen zur Einbindung von Diensten über das Internet in verschiedene Programme und Infrastrukturen. Auch forschungsnahe Dienste profitieren sehr von der Bereitstellung von REST-Schnittstellen, damit sie miteinander verschränkt und in andere Infrastrukturen eingebunden werden können.

Im Bereich von Repositorien, der in Bezug auf Schnittstellen und Standards oft Auswirkungen auf andere forschungsnahe Dienstleistungen hat, wurden von der [Coalition of Open Access Repositories](https://www.coar-repositories.org/) (COAR) weitere neue Protokolle vorgeschlagen. Hierzu zählt Signposting, das typisierte Links einsetzt, um von einer Ressource auf andere in Verbindung stehende Ressourcen zu verlinken und so die Auffindbarkeit durch Crawler und Bots erleichtern soll. Zum Zeitpunkt der Entstehung dieses Buchs ist offen, ob Signposting sich durchsetzt, wenn möglich sollte es in Repositorien aktiviert werden.

Eine ganz aktuelle Entwicklung ist das [Notify Project](https://www.coar-repositories.org/notify/), das ebenfalls von COAR betrieben wird. Notify soll es ermöglichen, dass verschiedene forschungsnahe Dienste sich Nachrichten über Aktivitäten senden und so auf Ressourcen aufmerksam machen. Als Beispiele gelten automatisiert ablaufende Aktivitäten zwischen einem textuellen Repositorium und einem Forschungsdatenrepositorium oder zwischen einem Repositorium und einem Service für die Organisation des Peer-Reviewing. Notify hat vor allem in den USA eine größere Förderung erhalten und wird derzeit in Softwarelösungen für verschiedene forschungsnahe Dienste integriert. Zum Zeitpunkt des Schreiben dieses Buchs ist zu erwarten, dass Notify sich in diesem Bereich verankern und helfen wird, Dienste dynamischer miteinander zu verknüpfen.

*Von diesen Entwicklungen wird auch die Veröffentlichung von Forschungsdaten profitieren. Bevor Daten allerdings in einem Zustand sind, dass sie veröffentlicht werden können, durchlaufen sie einen eigenen Lebenszyklus, bei dem unterstützende Dienste der Bibliotheken zunehmend gefragt sind.*

### Toolchains

|  |
| --- |
| Definition |
| Eine **Toolchain** („Werkzeugkette“) ist eine Reihe von miteinander verbundenen Anwendungen und Technologien, die gemeinsam eingesetzt werden, um spezifische Aufgaben oder Arbeitsabläufe zu optimieren und zu automatisieren. Die Einrichtung einer Toolchain hilft, den reibungslosen Informationsfluss zwischen Arbeitsschritten zu verbessern. |

Forschungsnahe Dienste müssen immer im Kontext des Forschungsprozesses und im Hinblick auf den Nutzen für die Forschenden betrachtet werden. Dazu ist es wichtig, für jeden der Dienste, die an einer Einrichtung genutzt werden, ein klares Profil zu erstellen. Hierbei muss insbesondere geklärt werden, wie sich die Dienste voneinander abgrenzen, damit Nutzer\*innen klar kommuniziert werden kann, welcher Dienst wofür verwendet wird.

Gleichzeitig muss aber auch das Zusammenspiel der einzelnen Dienste analysiert werden. Wie knüpfen die verschiedenen Dienste aneinander an? Wie kann der gebotene Mehrwert durch Verknüpfungen der Dienste gesteigert werden? Und welche Dienste bauen wie aufeinander auf? Diese Fragen sind nur lokal und konkret zu vorhandenen oder in Planung befindlichen Diensten zu beantworten.

Beispiel: An einer Bibliothek findet in Zusammenarbeit mit Fachwissenschaftler\*innen ein großes Retro-Digitalisierungsprojekt statt. Für die Verwaltung der [Digitalisierungsvorgänge](#digitalisierung) wird *Kitodo* verwendet. Die Werke sind im Bibliotheksmanagementsystem verzeichnet und die Metadaten werden über die SRU-Schnittstelle nach Kitodo importiert. Weitere Strukturdaten werden in Kitodo direkt eingetragen. Nach dem Scannen werden die Dokumente von Kitodo über die REST-API oder SWORD ins Repositorium exportiert. Das Repositorium ruft weitere Metadaten über SRU aus dem Bibliothekskatalog ab und vergibt DOIs, die auch wieder in das Bibliotheksmanagementsystem zurückgespeichert werden. Das Forschungsinformationssystem harvestet die Inhalte über OAI-PMH regelmäßig und weist die Digitalisate nach, die im Repositorium bereitgestellt werden.

*In diesem ideellen Bild wird nicht betrachtet, wo die nötigen Systeme stehen. Dies muss nicht immer ein lokal betriebenes System auf eigener Hardware sein, sondern ist oftmals als externer Service verfügbar.*

### Zusammenarbeit mit Dienstleistern

Selbstverständlich ist es nicht bei allen Problemstellungen möglich, IT-Services für forschungsnahe Dienste im eigenen Haus anzubieten. Bei bereits etablierten Anwendungen lohnt sich eine Kontaktaufnahme mit der jeweiligen Verbundzentrale. Häufig werden dort bereits Services angeboten, für die man kein zusätzliches Personal bzw. keine eigene Infrastruktur einplanen muss. Beispielsweise sind das der Repository-Service [Reposis](https://www.gbv.de/informationen/Verbundzentrale/serviceangebote/reposis-repository-service) des GBV oder das Langzeitarchiv [„Ewig“](https://ewig.zib.de/) des KOBV.

Wird die Verwendung einer bestimmten Software gefordert oder eine bestimmte Art, die Software einzusetzen, die nicht im Dienstleistungsportfolio der Verbundzentralen liegt, bietet sich die Zusammenarbeit mit externen, kommerziellen Dienstleistern an. Wo immer möglich, sollte Software von Open-Source-Communities Vorrang vor proprietärer Software genießen, da sie vor Abhängigkeiten von einzelnen Anbietern schützt und Weiterentwicklungen, die in die Community eingebracht werden, von anderen Einrichtungen wiederverwendet werden können. Vor allem im Bereich der forschungsnahen Dienste sind Open-Source-Lösungen oft vorherrschend. Ziel von Infrastruktureinrichtungen sollte es sein das so zu erhalten, anstatt den Aufbau proprietärer Software und neuer Oligopol- oder Monopolstellungen zuzulassen. Der Mehrwert, den IT-Dienstleister bieten, besteht ähnlich wie bei den Verbundzentralen darin, dass man selbst personelle und ggf. Infrastrukturressourcen sparen kann. In den meisten Fällen gibt es die Möglichkeit, entweder die Installation und Betreuung des Dienstes auf lokaler Infrastruktur oder auch ein „Rundum-Sorglos-Paket“ mit Hosting beim Dienstleister inklusive Betreuung einzukaufen.

Bei der Wahl des Anbieters gilt es, auf dessen Erfahrung im Umgang mit Open-Source-Projekten allgemein und mit der gewünschten Software im Speziellen zu achten. Genau wie nach Referenzen zu vergleichbaren Projekten zu fragen, lohnt es sich nach konkreten bereits geleisteten Beiträgen zu der jeweiligen Open-Source-Lösung zu fragen. Wenn die Open-Source-Lizenzen der jeweiligen Software keine Auflage machen, dass Weiterentwicklungen unter derselben Lizenz verbreitet werden müssen, sollte die Frage, unter welcher Softwarelizenz Weiterentwicklungen stehen, zwingend im Vertrag geklärt werden.

Besonders weit verbreitete Softwarelösungen haben große Entwicklergemeinschaften, die es einerseits zu unterstützen gilt, andererseits aber auch die communitygetriebene Entwicklungsrichtung der Software zu beachten ist. Ist der gewünschte Dienstleister hier noch neu, ist zumindest eine anderweitige Erfahrung in ähnlichen Projekten wünschenswert. Eine Recherche in öffentlichen Software-Repositorien der Projekte kann hier zielführend sein.

## Zusammenfassung und Ausblick

Forschungsnahe Dienste können Wissenschaftler\*innen grundsätzlich über den gesamten Forschungsprozess hinweg unterstützen - das Portfolio möglicher Services ist daher sehr groß. Bibliotheken setzen ihren Schwerpunkt hierbei insbesondere auf Services zur Unterstützung des Publikationsprozesses sowie des Forschungsdatenmanagements.

Wie das vorliegende Kapitel anschaulich gezeigt hat, umfassen diese Services auch eine Vielzahl an IT-Diensten, so z. B. Journal-Publishing-Systeme, Repositorien und Forschungsinformationssysteme. Der stabile und nachhaltige Betrieb solcher Dienste umfasst technische, organisatorische und inhaltliche Aspekte. Der Aufbau und Betrieb forschungsnaher Dienste an Bibliotheken bindet daher umfangreiche Ressourcen und erfordert ggf. auch eine Verlagerung von Ressourcen aus anderen Bereichen. Die Ausweitung des bibliothekarischen Serviceportfolios um forschungsnahe Dienste ist daher derzeit vor allem auch eine Frage der Organisations- und Personalentwicklung.

Zur vertieften Beschäftigung mit forschungsnahen Diensten an Bibliotheken wird folgende Literatur empfohlen:

* Konrad, Uwe, Förstner, Konrad, Reetz, Johannes, Wannemacher, Klaus, Kett, Jürgen, & Mannseicher, Florian (2020). Digitale Dienste für die Wissenschaft. Wohin geht die Reise? (Positionspapier), Zenodo. <https://doi.org/10.5281/zenodo.4301924>
* Farrenkopf et al. (2021). Forschungsunterstützung an Bibliotheken: Positionspapier der Kommission für forschungsnahe Dienste des VDB. O-Bib. Das Offene Bibliotheksjournal / Herausgeber VDB, 8(2), 1–19. <https://doi.org/10.5282/o-bib/5718>
* Berg-Weiß, A., et al. (2022). Openness in Bibliotheken : Positionspapier der Kommission für forschungsnahe Dienste des VDB. O-Bib. Das Offene Bibliotheksjournal / Herausgeber VDB, 9(2), 1–4. <(https://doi.org/10.5282/o-bib/5826>
* Azeroual, Otmane. „Untersuchungen zur Datenqualität und Nutzerakzeptanz von Forschungsinformationssystemen“. Dissertation. Magdeburg: Otto-von-Guericke-Universität Magdeburg, 2021.
* Schirrwagen, Jochen: Repositorien und Forschungsinformationssysteme bilden keine Dichotomie, Bibliothek - Forschung und Praxis 46, 2, 2022, 284–288.
* Management von Forschungsinformationen in Hochschulen und Forschungseinrichtungen. Eine Standortbestimmung 2022 (2022), hrsg. DINI AG Forschungsinformationssysteme, <https://doi.org/10.18452/25440>
* Druskat, Stephan und Bertuch, Oliver und Juckeland, Guido und Knodel, Oliver und Schlauch, Tobias (2022) *Software publications with rich metadata: state of the art, automated workflows and HERMES concept.* sonstiger Bericht. doi: <https://doi.org/10.48550/arXiv.2201.09015>; mith AM, Katz DS, Niemeyer KE, FORCE11 Software Citation Working Group. 2016. Software citation principles. PeerJ Computer Science 2:e86 <https://doi.org/10.7717/peerj-cs.86>
* Yves Vincent Grossmann und Michael Franke, Software ist kein Beiprodukt! Nachhaltige Forschungssoftware durch Software-Management-Pläne, b-i-t-online 26 (2023) Nr. 5, 457- 463, <https://www.b-i-t-online.de/heft/2023-05-fachbeitrag-grossmann.pdf>
* Cyra, M. A., Politze, M. ., & Timm, H. (2022). A push for better RDM: Erfahrungsbericht aus dem Einsatz von git für Forschungsdaten. *Bausteine Forschungsdatenmanagement*, (2). <https://doi.org/10.17192/bfdm.2022.2.8435>
* Barker, M., Chue Hong, N.P., Katz, D.S. *et al.* Introducing the FAIR Principles for research software. *Sci Data* **9**, 622 (2022). <https://doi.org/10.1038/s41597-022-01710-x>
* Putnings, Markus, Neuroth, Heike und Neumann, Janna. *Praxishandbuch Forschungsdatenmanagement*, Berlin, Boston: De Gruyter Saur, 2021. <https://doi.org/10.1515/9783110657807>

# Kommunikation

|  |
| --- |
| Wichtig |
| Dieses Kapitel befindet sich in einer ersten öffentliche Entwurfsversion. Feedback siehe Links im rechten Menü und [Hinweise zur Mitarbeit](#mitarbeit). |

## Einleitung

Im Verhältnis zwischen Intranet und Öffentlichkeitsarbeit bewegt sich die moderne Bibliothekslandschaft zunehmend weg von einem dualen Kommunikationsverständnis hin zu einem integrierten Ansatz, in dem Kommunikationsräume oft überlappen und Werkzeuge sowohl für die interne Kommunikation als auch für den Dialog mit der Öffentlichkeit genutzt werden.

Indem Bibliotheksmitarbeitende aktiv mit Intranet-Tools und -Plattformen arbeiten, bauen sie technische Fähigkeiten und Selbstvertrauen im Umgang mit ähnlichen Kommunikationswerkzeugen für die externe Kommunikation auf. Dadurch wird die Einführung von Intranetwerkzeugen zu einer Schlüsselstrategie der Organisationsentwicklung. Mitarbeitende verwandeln sich so in authentische Botschafter der digitalen Transformation, interagieren kompetent und auf Augenhöhe mit den Nutzenden und verkörpern eine innovative, nutzerzentrierte Kommunikationsweise:

|  |
| --- |
| Stufenmodell von Öffentlichkeit (Stockmann, 2019) |

## Ziele

Auch wenn anhand des obigen Stufenmodells deutlich wird, dass der Übergang von einem internen zu einem externen Publikum als Kontinuum verstanden werden sollte, können aber die [**Ziele von interner**](#ziele-interner-kommunikation) und [**Ziele von externer Kommunikation**](#ziele-externer-kommunikation) klar voneinander abgegrenzt werden.

### Ziele interner Kommunikation

**Interne Kommunikation** bezieht sich auf den gezielten Austausch von Informationen, Meinungen und Vorstellungen innerhalb einer Organisation. Sie umfasst sämtliche Kommunikationsprozesse und -instrumente, die darauf abzielen, die Mitarbeiter:innen eines Unternehmens oder einer Einrichtung miteinander und mit der Organisation insgesamt zu vernetzen. Hauptziele sind oft die Sicherstellung eines konsistenten Informationsflusses, die Unterstützung der organisatorischen Ziele und die Förderung einer positiven Unternehmenskultur. Im besten Falle vermeidet gelungene interne Kommunikation Wissensinseln und zu starke Wissensbindung an einzelne Personen.

**Wissensmanagement** beschreibt hingegen den Prozess des Erfassens, Organisierens, Bewahrens, Anwendens und Weitergebens von Wissen innerhalb einer Organisation. Dieses Wissen kann sowohl explizit (in Dokumenten, Datenbanken etc.) als auch implizit (in den Köpfen der Mitarbeitenden) vorhanden sein. Das Wissensmanagement zielt darauf ab, das im Unternehmen vorhandene Wissen effektiv zu nutzen, um Arbeitsprozesse zu verbessern, Innovationen zu fördern und die Wettbewerbsfähigkeit zu sichern.

Bezogen auf die interne Kommunikation bedeutet dies:

* Die **interne Kommunikation** ist ein zentrales Instrument des Wissensmanagements. Durch gezielte Kommunikationsmaßnahmen wird Wissen im Unternehmen verteilt, Mitarbeitende werden über Neuerungen informiert und der Austausch zwischen den Abteilungen wird gefördert.
* Das **Wissensmanagement** wiederum stellt sicher, dass die in der internen Kommunikation übermittelten Informationen von Relevanz und Qualität sind und zur richtigen Zeit am richtigen Ort ankommen.

In der Praxis sind interne Kommunikation und Wissensmanagement oft eng miteinander verknüpft, da eine effektive Kommunikation innerhalb einer Organisation es erleichtert, Wissen zu identifizieren, zu teilen und anzuwenden.

### Ziele externer Kommunikation

Bibliotheken bedienen eine Vielzahl verschiedener Zielgruppen. In den Wissenschaftlichen Bibliotheken sind es Studierende, Lehrende und Wissenschaftler. Öffentliche Bibliotheken bedienen sehr unterschiedliche Bedürfnisse und Gruppen, für die sie jeweils spezielle Angebote bereit halten. Dies gilt es bei der externen Kommunikation stets im Auge zu behalten.

**Externe Kommunikation** bezeichnet den systematischen und zielgerichteten Austausch von Informationen, Meinungen und Vorstellungen zwischen einer Organisation und ihren externen Stakeholdern. Dies sind in erster Linie die Nutzenden der Bibliothek, kann aber darüber hinaus auch andere Bibliotheken, Medien oder auch die breitere Öffentlichkeit bzw. die Politik umfassen. Externe Kommunikation kann sowohl zur Imagepflege als auch zur Informationsweitergabe über Produkte, Dienstleistungen oder andere relevante Entwicklungen in der Bibliothek dienen. Je nach Anspruch der Bibliothek kann die Kommunikation auch mehrsprachig sinnvoll sein.

**Öffentlichkeitsarbeit** (PR) ist ein spezifischer Bereich der externen Kommunikation, der sich darauf konzentriert, das Verständnis und das Image einer Organisation in der Öffentlichkeit zu fördern und zu pflegen. Die Ziele der Öffentlichkeitsarbeit können beinhalten:

* Aufbau und Pflege eines positiven Images und eines guten Rufs der Organisation in der Öffentlichkeit.
* Management von Krisensituationen und negativen Entwicklungen, die das Image der Organisation schädigen könnten.
* Erhöhung der Sichtbarkeit und Anerkennung der Organisation, ihrer Produkte oder Dienstleistungen.
* Einflussnahme auf die öffentliche Meinung oder politische Entscheidungsfindung.

Öffentlichkeitsarbeit verwendet eine Vielzahl von Kommunikationstools und -strategien, darunter Pressemitteilungen, Veranstaltungen, Social Media, Sponsoring und vieles mehr.

## Rahmenbedingungen

Die Einführung digitaler Werkzeuge und Verfahren ist eingebettet in verschiedene rechtliche und organisatorische Rahmenbedingungen:

* **Beteiligungsverfahren**: Etablierung klarer Dienstvereinbarungen und intensiver Dialog mit dem Personalrat und dem Kollegium, um eine transparente und mitarbeiter:innenorientierte Implementierung von digitalen Werkzeugen sicherzustellen.
* **DSGVO-Konformität**: Gewährleistung, dass alle eingesetzten digitalen Lösungen und Technologien die Anforderungen der Datenschutz-Grundverordnung (DSGVO) erfüllen, um den Schutz personenbezogener Daten sicherzustellen. Prüfung und Klarstellung, ob bei der Nutzung bestimmter Dienste ein Vertrag zur Auftragsverarbeitung (AV-Vertrag) notwendig ist sowie Aufnahme der Dienste in das Verfahrensverzeichnis
* **Betriebsmodell**: Für jedes Werkzeug ist die Entscheidung zu treffen, ob es a) im eigenen Rechenzentrum - so vorhanden - betrieben wird, b) bei einem externen Hosting-Dienstleister aufgesetzt wird oder c) als Software-as-a-Service (SaaS) von Dienstleistern eingekauft wird. TODO: Siehe auch den Abschnitt „Betriebsmodelle“ im Kapitel IT-Management.
* **Technische und organisatorische Maßnahmen (TOMs)**: Etablierung von TOMs, um sowohl den sicheren Betrieb der technischen Infrastruktur zu garantieren als auch bei möglichen Sicherheitsvorfällen vorbereitet und reaktionsschnell zu sein.
* **Weiterbildung und Kompetenzaufbau**: Sicherstellung, dass das Bibliothekspersonal durch kontinuierliche Weiterbildungen, Schulungen und Dokumentation in den kompetenten und souveränen Umgang mit den neuen digitalen Werkzeugen eingeführt wird.
* **Barrierefreiheit nach BITV 2**: Sicherstellen, dass alle digitalen Angebote und Dienste den Anforderungen der Barrierefreie Informationstechnik-Verordnung (BITV 2.0) entsprechen, um einen inklusiven und für alle zugänglichen digitalen Raum zu schaffen, der keine Nutzergruppe ausschließt und die Diversität der Gemeinschaft respektiert.
* **Institutioneller Rahmen:** Tools in Bibliotheken sind oft abhängig von der Institution / des jeweiligen Trägers. Software die beispielsweise innerhalb einer Universität angewendet wird, nutzt die Bibliothek dann ebenfalls zur Kommunikation. Ein Beispiel dafür wäre der Rocket-Chat oder Big Blue Button als Videokonferenztool.

Ein ausgeprägtes Maß an digitaler Souveränität (siehe Abschnitt im Kapitel IT-Management) kann die Auseinandersetzung mit den aufgeführten Rahmenbedingungen erheblich vereinfachen.

## Prozesse

Für welche Bibliotheksprozesse interne und externe Kommunikation notwendig ist, wird in den folgenden Abschnitten beschrieben.

### Markenbildung

Auch in Bibliotheken spielt das **Corporate Design** eine wichtige Rolle in der Öffentlichkeitsarbeit. Ein einheitliches Corporate Design stellt sicher, dass die visuelle Identität, einschließlich des Logos, der Farben und Schriftarten, in allen gedruckten und digitalen Materialien konsistent ist. Die Einhaltung des Corporate Designs über verschiedene Kommunikationskanäle hinweg sorgt für ein kohärentes Auftreten der Bibliothek und vermittelt den Eindruck von Stabilität und Seriosität. So stärkt ein gutes Corporate Design unmittelbar das Vertrauen der Nutzenden / der Öffentlichkeit in die Bibliothek.

Gepflegt werden kann ein Corporate Design anhand eines Styleguides. Dieser legt fest, wie visuelle Elemente in verschiedenen Kommunikationskanälen verwendet werden sollen, sei es in gedruckten Broschüren, auf der Website oder in sozialen Medien. Ggf. können auch sprachliche Vorgaben gemacht werden. Dies erleichtert die Erstellung von Marketingmaterialien und eine effektive Kommunikation der Bibliotheksangebote. Durch die Verwendung von vordefinierten Vorlagen und Design-Richtlinien können Bibliotheken nicht zuletzt auch die Kosten für die Gestaltung von Materialien senken und Ressourcen effizienter nutzen; das Rad muss (und sollte) nicht immer wieder neu erfunden werden.

Wenn möglich sollte man Pressematerial wie Logos, standardisierte Infotexte zur Bibliothek für Journalist:innen zum Download auf Website anbieten, so dass diese schnell auf grundlegende Infos zugreifen können.

Auch die interne Kommunikation sollte sich entlang der Vorgaben orientieren. Dies spart zum einen Zeit, da Vorlagen genutzt werden können und übt zum anderen den Gebrauch auch für externe Kommunikation.

### Redaktion

Für die Erstellung von internen wie externen Inhalten benötigt man Zeit und Ideen. Um nicht unter „Zugzwang“ zu geraten, hilft es, sich eine Übersicht über Inhalte zu machen, die man nach kommunizieren möchte und zu diesen entsprechende Formate zu entwickeln oder festzulegen. Denn: Meist mangelt es nicht an Themen, vielmehr hat man i.d.R. bereits sehr viele Themen, die nur noch „verpackt“ werden müssen.

Dabei helfen Fragen zum Beispiel nach dem Umfang und der Tiefe des Inhaltes. Nicht jedes Thema eignet sich, um ein ausführliches Video zu drehen. Andersherum kann es sein, dass ein Thema nicht in Form eines einzelnen Social-Media-Posts verständlich aufbereitet werden kann. Es geht also darum, sich über die unterschiedlichen medialen Darstellungsformen - Video, Text, Text und Bild, Grafik, etc. - klar zu werden und Themen zuzuordnen.

Auch innerhalb einer Darstellungsform gibt es unterschiedliche Formate: informative Social Media Posts, ausführliche Blogbeiträge oder verschiedene Videoformate. Gleichzeitig sollten die Voraussetzungen, Möglichkeiten und Konventionen der Plattformen berücksichtigt werden.

Auch empfiehlt es sich (grafische) Templates zu nutzen (zum Beispiel bei Canva Templates für Infoposts, Veranstaltungshinweise, Personenvorstellung etc. anlegen). So braucht man nur noch die vorgefertigten Templates mit Inhalten zu füllen. Auch bei der Produktion von Videos ist die Arbeit mit Formaten (Interviews, Erklärvideos, Tutorials), nach deren Muster die Videos erstellt werden, von Vorteil.

Außerdem ist es hilfreich bei der Erstellung eines konkreten Inhalts zu überlegen, welche weiteren Formate sich anbieten. So braucht man die inhaltliche Recherche/Arbeit nur einmal machen, kann diese aber in mehrfacher Form und an verschiedenen Stellen wieder aufbereiten und nachnutzen. Wird zum Beispiel ein ausführlicher Blogbeitrag zu einem Thema verfasst, so lassen sich daraus häufig mehrere Social-Media-Posts machen. Besonders bei der Produktion eines Videos bietet es sich an, "Hinter den Kulissen” Content zu produzieren, z.B. indem man den Entstehungsprozess dokumentiert durch kurze Videos und Fotos oder zu einem ausführlichen Video eine Kurzversion dreht - wenn die Technik schon mal steht.

* **Vorproduzieren und Redaktionsplan nutzen**:  
  Neben tagesaktuellem Content, wie zum Beispiel Hinweise auf Veranstaltungen, gibt es auch den so genannten Evergreen Content. Bei WBs können das z.B. Inhalte zum wissenschaftlichen Publizieren, Forschungsdatenmanagement o.ä. sein, die sich auf verschiedenen Unterseiten des Webauftritts verbergen. Diese lassen sich gut für andere Plattformen aufbereiten, man kann sie gut vorproduzieren, beliebig erweitern und mit Hilfe eines Redaktionsplans ausspielen.
* In einem Redaktionsplan wird festgehalten, welche Inhalte wann (und von wem) publiziert werden. So behält man einen Überblick. Auf diese Weise kann man auch über einen längeren Zeitraum im Voraus planen, vor allem Evergreen Content lässt sich so gut verteilen. Es bietet sich an, eine Regelmäßigkeit für solchen Content festzulegen, wie etwa an einem bestimmten Tag in der Woche o.ä. In ÖBs eignen sich hierfür beispielsweise Buchempfehlungen oder regelmäßig wiederkehrende Veranstaltungen.
* **Content anteasern/ Social Storytelling**: vor allem „größeren“ Content, wie zum Beispiel aufwändigere Videoproduktionen, vorab auf den Social Media Plattformen durch kurze Begleitbeiträge, wie etwa Behind-the-scenes-Berichte oder allgemeine Infos zum Thema anteasern. So kann die Neugier der Nutzenden geweckt werden und sie bleiben dran.
* **Prozesse optimieren**: Häufig wird in Bibliotheken die externe Kommunikation von mehr als einer Person übernommen. Zur Strukturierung und zur Absprache innerhalb eines Teams eignen sich verschiedene Kommunikationstools, vom 1:1 Gespräch über eine Chatgruppe bis hin zur Nutzung eines Redaktionsplanes, um gemeinsam Inhalte zu planen. Wichtig ist hierbei Konsistenz und die Nutzung der, für den vorliegenden Zweck, effektivsten Tools. Dabei sollten nicht zu viele unterschiedliche Kanäle genutzt werden, um nicht den Überblick zu verlieren. Es sollte vorab klar sein, welche Voraussetzungen erfüllt sein müssen, um bestimmte Formate zu kommunizieren. Dazu kann es hilfreich sein, einen Zeitplan sowie eine Checkliste zu erstellen, um sicherzugehen, dass alle relevanten Punkte abgedeckt sind (Wo wird Was Wie kommuniziert).

### Wissensmanagement

Explizites Wissen ist gesichert in Texten, Datenbanken und anderen Dokumenten, doch wie sichern wir das implizite Wissen, das nur in den Köpfen von Mitarbeitenden vorhanden ist? Für diesen Teil des Wissensmanagements gibt es verschiedene mündliche und schriftliche Verfahren. Dabei helfen unter anderem strukturierte Interviews, Erzählungen (Storytelling), Visualisierungen, Verschriftlichung mündlicher Anleitungen, Videobotschaften und nicht zuletzt das Vormachen von Tätigkeiten. Zur kontrollierten Umsetzung des Wissenstransfers gibt es mehrere Werkzeuge (Mittelmann, 2011):

* **Jobmap**: ein Instrument, das dazu dient, das Wissen ausscheidender Mitarbeiter\*innen systematisch festzuhalten und zu dokumentieren. Sie kann händisch oder mit Mindmapping-Werkzeugen erstellt werden. Sie kann als Protokoll eines moderierten Gesprächs zwischen Wissensgeber (Ausscheidende\*r) und Wissensnehmer (Nachfolger\*in) befüllt oder von Mitarbeiter\*innen selbst erstellt werden.
* **Wissensträgerkarte:** veranschaulichen, bei welchen Wissensträger\*innen welche Kompetenzen in welchen Wissensgebieten vorhanden sind. Sie vermitteln keine Wissensinhalte, sondern zeigen, welche Kompetenzen ein bestimmter Mitarbeitender hat und wer Wissensträger für ein bestimmtes Gebiet ist.
* **persönlicher Wissenstransfer:** Gespräch
* **kodifizierter Wissenstransfer:** mittels Informations- und Kommunikationssysteme wie z.B. Datenbanken, Laufwerk, CMS, Wiki
* **Infokarte:** Themenkarte oder Mikroartikel welcher eine Aufgaben beschreibt
* **Projektplan:** fasst den Vollständigen Ablauf des Wissenstransfers zusammen. Er enthält Informationen über die Zeitplanung, den Umfang, die Fälligkeitsdaten und die Ergebnisse.

|  |
| --- |
| Jobmap nach Mittelmann (2011) |

### Evaluation

Sind Tools/Kanäle einmal eingeführt, wird leider oft nicht weiter verfolgt, wie sie angenommen und genutzt werden. Regelmäßige Evaluation der Nutzung und auch Fehlerauswertung helfen dabei, Schwachstellen zu erkennen, Verbesserungen einzubinden oder Tools/Kanäle ggf. auch wieder abzuschaffen bzw. durch andere abzulösen.

Viele Tools und Werkzeuge bringen ihre eigenen Statistiken und Auswertungen mit, aber auch eine einfache, aber regelmäßige Abfrage der Nutzungsgewohnheiten der Mitarbeitenden hilft bereits, die Praxistauglichkeit einzuschätzen. Daher ist es sinnvoll, die Evaluation bereits bei der Planung, spätestens aber beim Start mit einzuplanen.

Beispiele:

* Nutzungsstatistiken Webseiten und Werkzeuge (z.B. mittels des Open Source Tools „Matomo“)
* Usability Studien
* Social Media Plattformen haben in der Regel eigene integrierte Analysewerkzeuge

Neben herkömmlichen Nutzungsstatistiken können einheitliche Fragebögen bei der Einschätzung helfen, ob Kanäle oder Tools geeignet sind, die gesetzten Ziele zu erreichen.

Plattformen und Kanäle sollten regelmäßig evaluiert werden, um festzustellen, ob die Zielgruppe über diesen Weg tatsächlich erreicht wird.

## Kanäle

Es gibt verschiedene Online-Kanäle, mit Hilfe derer man die Nutzenden informieren kann. Diese werden in den folgenden Abschnitten beschrieben.

### Website

Die Website, im allgemeinen Sprachgebrauch häufig auch als Homepage bezeichnet, dient der Vorstellung der Bibliothek, ihrer verschiedenen Angebote und Kontaktmöglichkeiten. Hier werden u.a. Neuigkeiten über Veranstaltungen, digitale Angebote oder Kundenservices (z.B. Anmeldemodalitäten) präsentiert.

Websites werden meistens auf Basis eines Content Management Systems (CMS) erstellt. Mit Hilfe von Content Management Systemen können redaktionelle Inhalte dargestellt, organisiert und bearbeitet werden, ohne dass sich die Websiteredakteur\*innen mit dem Aufbau und dem Design der Website beschäftigen müssen.

Öffentliche Bibliotheken sind oft Teil der Kommunen-Infrastruktur, weshalb die Website häufig eine Unterseite der Seite darstellt. Hier ist das CMS vorgegeben und der Gestaltungsspielraum oft begrenzt. Im besten Fall hat die Bibliothek eine eigene Domain, kann ein eigenes CMS auswählen und hat damit auch die Gestaltungsfreiheit, beliebige Unterseiten anzulegen und eigene Inhalte zu beschreiben. WBs hingegen sind oft als Teil der Universitätswebsite angegliedert und dementsprechend limitiert in der Ausgestaltung.

In beiden Fällen - ob eigene Domain oder eingegliedert in die übergeordnete Einrichtung - gibt es in der Regel Designvorgaben wie eine Corporate Identity. Dennoch sollten insbesondere bei der Planung und späteren Konzeption gängige [Usability-Gesichtspunkte](#was-beeinflusst-den-nutzungseindruck) berücksichtigt werden.

Gängige CMS sind *Typo3*, *Wordpress*, *Drupal*. Ob eine CMS vorgegeben ist oder selbst festgelegt werden kann, ist von Institution zu Institution unterschiedlich und von den Rahmenbedingungen abhängig. Das gleiche gilt für das Hosting und den technischen Support der Website. Dies kann entweder hausintern passieren, über einen Verbundpartner oder bei einem kommerziellen Drittanbieter.

Ist eine Bibliothek in der komfortablen Lage, ein eigenes CMS auszuwählen, sollten auch immer die anzubindenden Portale wie Discovery-Systeme oder Repositorys berücksichtigt werden.

### Blog

Ein Blog ist eine spezielle Form einer Website und dient in erster Linie als Plattform für die Veröffentlichung redaktioneller (Text-)Beiträge. Blogbeiträge eignen sich vor allem für die Darstellung von Inhalten, die ausführlicher und tiefer in eine Thematik einsteigen und über den rein informativen Charakter hinausgehen. Ein Blog kann einerseits als ein Bereich innerhalb eines Webauftritts angesiedelt werden oder als eigene Website existieren. Für das Aufsetzen eines eigenständigen Blogs eignet sich vor allem das CMS *Wordpress*.

### Social Media

Über Social-Media-Kanäle kann man die Nutzenden / die Zielgruppe schnell erreichen, eigene Inhalte mit ihnen teilen und mit ihnen in Kontakt treten. Auch dienen Social-Media-Kanäle der Markenbildung. Jede Social-Media-Plattform hat eigene Konventionen und Bedingungen, sowohl was die Gestaltung der Beiträge angeht als auch die zugrunde liegenden Algorithmen, mit denen man sich vertraut machen sollte. Ziel sollte es sein möglichst viele Nutzende zu erreichen und neue Nutzende dazu zu gewinnen. Hier auch wichtig: nicht nur die eigenen Kanäle zu bespielen, sondern auch mit anderen Kanälen zu interagieren.

Je nach Zielgruppe sind die einzelnen Social-Media-Kanäle mehr oder weniger geeignet. Ist eine Zielgruppe auf einem Netzwerk nicht aktiv, so lohnt es sich nicht, dieses Netzwerk zu bespielen und man sollte sich eher auf die Netzwerke fokussieren, in denen die eigene(n) Zielgruppe(n) anzutreffen sind.

* **Twitter/X, Mastodon, Bluesky:** Kurznachrichten, v.a. textbasierte Inhalte; auch gerne genutzt, um auf Inhalte auf anderen Plattformen (Fachartikel, Blogbeiträge, Veranstaltungen..) zu verweisen. Für das deutschsprachige Bibliothekswesen ist insbesondere die Mastdon-Instanz [openbiblio.social](https://openbiblio.social) etabliert.
* **Facebook, Instagram:** Inhalten in Form von Bildern/Grafiken, Texte, (Kurz-)Videos;
* **TikTok:** Videoportal für kurze Videos im Hochformat
* **YouTube, Vimeo:** Plattformen für die Veröffentlichung von Videos
* **LinkedIn:** Längst hat sich LinkedIn von einem reinen Berufsnetzwerk für den beruflichen Einstieg oder die berufliche Weiterentwicklung/ Veränderungen hin zu einem Plattform entwickelt, auf der Inhalte geteilt und diskutiert werden. Vor allem für Themen aus Wissenschaft und Forschung eignet sich LinkedIn als Plattform. Allerdings ist hier zu beachten, dass persönliche Accounts vom Algorithmus bevorzugt behandelt werden im Gegensatz zu Unternehmensseiten, so dass über persönliche Accounts geteilte Inhalte eine deutlich höhere Reichweite erzielen.

### Newsletter

Newsletter sind redaktionell aufbereitete Texte und Grafiken, die an einen bestimmten Verteilerkreis (Studenten, Kund\*innen) in einem regelmäßigen Rhythmus versendet werden. Ein Beispiel sind Veranstaltungsnewsletter der ÖBs. Dabei ist insbesondere der Datenschutz zu beachten. Vor dem Versand ist eine Einwilligung jedes Empfängers einzuholen, der durch Double-Opt-In verifiziert wurde. Es ist ratsam, für den Versand einen Newsletter-Dienst zu nutzen, um technische Fallstricke zu umgehen. Diese sind u.a. auf große Versandmengen, Umgang mit zurückkommenden Mails oder Einwilligungsverwaltung spezialisiert.

### Mailinglisten

Verteilung von Inhalten an eine (größere) Gruppe von E-Mail-Empfänger:innen. Personen einer Mailingliste können sich über diese untereinander austauschen.

### Offene Schnittstellen

Offene Daten verändern Wissens- und Wissenschaftskommunikation in Bibliotheken. Dies umfasst Daten digitalisierter Werke wie auch deren ggf. offene Metadaten. Diese Daten können in neuen digitalen Publikationen und Social Media im Detail verknüpft werden – als Zitate, Referenzen, Linkkontexte im Inhalt sowie auf der Metaebene in Metadaten einer Publikation. Offene Zitationsdaten ermöglichen bibliografische Analysen und Datenvisualisierungen, die selbst als Gegenstand in der internen und externen Wissenschaftskommunkation genutzt werden können. Dabei ermöglichen offene Werkzeuge wie Wikidata gewissermaßen eine ‚Demokratisierung‘ der Metadatenproduktion und -pflege unabhängig bzw. in Ergänzung von traditionellen Bibliothekssystemen. Wikidata fungiert dabei zudem als ‚Linked Open Data‘-Knoten (Datenhub) für Identifikatoren verschiedener Datenquellen, die jeweils identische Entitäten beschreiben (GND, PPN, …).

RSS-Feeds (Rich Site Summary) sind Dateiformate für Web-Feeds. Nach dem Abonnieren werden Änderungen von Webseiten, Blogs, Foren usw. im Feed-Reader angezeigt.

### Digitale Anzeigen/Poster

Unter digitalen Anzeigen werden alle Arten von elektronischen Displays verstanden, die Informationen zu Institutionen oder Angeboten und Veranstaltungen anzeigen. Beispiele sind Infotafeln und Wegweiser.

### Wiki

Ein Wiki ist eine webbasierte Sammlung von Informationen, Artikeln, Beiträgen zu bestimmten Themen. Diese können von Nutzenden selbst bearbeitet werden. Es kann frei im Internet oder nur für einen festgelegten Nutzerkreis (Mitarbeitende) verfügbar sein. (Bsp: PmWiki, Confluence…)

## Werkzeuge

Der bekannte Sinnspruch „Wer als Werkzeug nur einen Hammer hat, sieht in jedem Problem einen Nagel“ gilt auch im digitalen Raum. Die Positionierung zwischen den Extremen „wir bilden alles über das Dateisystem ab“ (ich habe nur einen Hammer) und „wir haben für jede Anforderung ein spezialisiertes Werkzeug“ (man schleppt einen großen Werkzeugkasten durch die Gegend) ist alles andere als trivial.

Einen ersten Überblick bietet dieses Schaubild, was die populären Werkzeuge entlang der Achsen Intern/Extern sowie Solitär/Kollaborativ positioniert:

|  |
| --- |
| Quadrantenmodell von Kommunikations-Werkzeugen |

Im Folgenden werden die grundlegenden Funktionen, Werkzeuge und Kanäle kurz beschrieben:

### Kommunikation

Synchrone und asynchrone Kommunikation sind zwei verschiedene Ansätze für den Informationsaustausch, die jeweils ihre eigenen Vorteile und Anwendungsgebiete haben. Bei synchroner Kommunikation tauschen die Kommunikationspartner Informationen in Echtzeit aus. Sie erlaubt sofortiges Feedback, d.h. Fragen können direkt gestellt und beantwortet werden. Es müssen allerdings beide Parteien gleichzeitig verfügbar sein. Das kann zu Ablenkungen bzw. Druck führen, da eine sofortige Reaktion erwartet wird. Bei asynchroner Kommunikation findet der Informationsaustausch zeitlich versetzt statt. Ein Teilnehmer sendet eine Nachricht, und der andere antwortet, wenn es passt. Die Kommunikationspartner sind also nicht an eine gemeinsame Zeit gebunden. Das ermöglicht eine größere Flexibilität in der Zeitplanung und eröffnet die Möglichkeit zur gründlichen Reflexion vor einer Antwort. Es ist außerdem oft weniger störend für den Arbeitsfluss. Allerdings kann asynchrone Kommunikation zu Verzögerungen und fehlendes sofortiges Feedback zu Missverständnissen führen.

* **Telefon** ist das klassische Medium mit synchroner 1-zu-1 Kommunikation für schnelle Absprache; mittlerweile durch die Verbreitung von Videokonferenzen, Chat etc. in der internen Kommunikation eher in den Hintergrund geraten, in der externen Kommunikation nach wie vor ein wichtiger Kanal insbesondere für den Erstkontakt.
* **E-Mail** ist ein Kommunikationsmedium, welches in den meisten Fällen vom Anbieter der IT-Infrastruktur angeboten wird. Ideal ist der Zugang per IMAP/SMTP und nicht nur über proprietäre Protokolle, so dass beliebige Clients genutzt werden können.
* E-Mail ist bis heute das klassische Kommunikationsmedium und allgegenwärtig. Ursprünglich als elektronische Variante des Briefes für die asynchrone 1:1 Kommunikation entwickelt, wird sie entgegen diesem Zweck häufig für die Gruppenkommunikation zweckentfremdet; außerdem wird sie häufig als synchron verstanden und eine sofortige Antwort erwartet. Das führt dazu, dass man meist (1) zu viele davon hat und (2) in den entscheidenden Unterhaltungen am Ende doch nicht im CC ist. Im besten Fall finden sich bessere Alternativen, die je Situation einsetzbar sind.
* **Videokonferenzsysteme** dienen dem synchronen Austausch mittels Bild - und Tonübertragung. Kommerzielle Beispiele sind Microsoft Teams, Webex und Zoom. Daneben stehe freie Alternativen wie [jitsi](https://jitsi.org/) und [DFNconf](https://www.conf.dfn.de/) zur Verfügung. Zunehmend sind Videofunktionen auch in anderen Werkzeugen integriert.
* Um virtuelle Besprechungsräume so auszustatten, dass auch Menschen an Besprechungen teilnehmen können, die nicht anwesend sind, ist technische Infrastruktur erforderlich. Die Anwesenden müssen für die Anwesenden sichtbar und hörbar sein, ohne dass die Anwesenden Headsets tragen müssten. Dies erreicht man mit einer Art überdimensionalen Webcams, die automatisch auf die sprechende Person fokussieren und im Idealfall auch Störgeräusche (Echos, Rauschen, Rascheln) ausblenden. Die gängigen Systeme im knapp vierstelligen Eurobereich genügen für Konferenzen mit bis zu sechs anwesenden Personen um einen Tisch herum. Sind mehr Personen anwesend, steigt der technische Aufwand stark an, wenn man häufig hybrid arbeiten möchte und die anwesenden Personen nicht benachteiligen möchte.
* **Chat** kann als synchrone und asynchrone Kommunikation 1-zu-1 und in Teams sowohl in der externen als auch in der internen Kommunikation verwendet werden:
  + Chats können in Bibliothekswebsiten (Homepage, OPAC) ohne großen Aufwand integriert werden. Die Nutzung ist anonym oder unter Angabe von personenbezogenen Daten (z.B. E-Mail-Adresse) möglich. Chat-Anfragen werden zentral bearbeitet und können bei Bedarf weitergegeben werden. Ein Pool von Mitarbeitenden kann sich anmelden, die Anfragen werden vom System auf die Mitarbeitenden verteilt. Steht kein Mitarbeitender zur Verfügung, leitet die Software z.B. auf ein Mailformular um. Je nach Verwendung ist der Abschluss eines Vertrages zur Auftragsverarbeitung mit dem Anbieter notwendig. Chat-Systeme werden i.d.R. als SaaS-Lösung angeboten
  + Chats zwischen Mitarbeiter\*innen dienen der niedrigschwelligen Kommunikation, oft als Alternative zu Telefon und E-Mail. Gruppenchats in Abteilungen können für einfache Fragen und Absprachen genutzt werden und Menschen, die mobil arbeiten, können mit einem Chatsystem besser integriert werden. Die Nutzung von kommerziellen Systemen, die noch dazu außerhalb von Europa gehostet werden, ist aus Datenschutzgründen nur in Ausnahmefällen zulässig (WhatsApp u.a., Teams, Slack, …). Sofern die Möglichkeit besteht, ist ein selbst gehostetes, ggf. Open Source System vorzuziehen (Beispiel: RocketChat).
* Eine Alternative ist für Bibliotheken mit IT-Abteilung oder RZ-Unterstützung die Nutzung von Open-Source-Software wie [mibew.org](https://mibew.org/). [rocket.chat](https://www.rocket.chat/secure-messaging-platform) bietet sich als selbst betriebene Alternative an. Der Chat ist häufig aber auch Bestandteil von anderen Lösungen wie bspw. Microsoft Teams oder Webex. Auch ein Forensystem wie Discourse bietet eine Chat-Komponente, sodass ernsthaft abgewogen werden sollte, ob es wirklich eine dezidierte Chat-Lösung braucht.

### Contenterstellung

Unter Contenterstellung versteht man das Produzieren von Inhalten (Werbung, Informationen…) z.B. für das Internet. Mit Content wird immer eine bestimmte Zielgruppe angesprochen. Inhalte sind zu planen, zu erstellen und zu veröffentlichen.

* **Bild-/Grafikbearbeitung:** dient der Nachbearbeitung von Bilddateien (Schärfen, Farben, Dateigröße, Belichtung, Kontraste)
  + Bildbearbeitungs-Software: *Affinity Photo*, *Adobe Photoshop*, *Canva*
  + freie Bilddatenbanken: *Pixabay*, *CC-Search*…
  + Grafikdesign: [*Canva*](https://www.canva.com/de_de/)
* **Audiobearbeitung:** hilft beim Schnitt von Audiodateien und dient der Klangoptimierung mit Effekten wie Equalizer, Kompressor oder Hall
  + [*Audacity*](https://www.audacity.de/)
  + Professionelle Podcast-Produktion: [*Ultraschall*](https://ultraschall.fm/)
* **Videobearbeitung:** dient der Erstellung von bewegten Bildern aus vorhandenem Ausgangsmaterial wie digitalisierte Fotos, Scans, Grafiken, Texte, Audioaufnahmen, Musik
  + Bearbeitungssoftware (*DaVinci Resolve*): Zur Nachbearbeitung von Roh-Videomaterial gibt es spezielle Software, die Video- und Tonspuren zusammenführen können. Ebenso sind Veränderungen durch Filter oder Zusammenführen von verschiedenen Bild- und Videoaufnahmen möglich. In der Regel bieten die Tools diverse Exportmöglichkeiten in verschiedenen Qualitäten und Formaten.
  + Aufnahmesoftware (*Camtasia*, *Screenflow*, [*OBS*](https://obsproject.com/de)): dient zur Aufnahme von Bildmaterial über Kameras oder z.B. der Anzeige auf Bildschirmen (Screencasts). Je nach Aufnahmegerät wird der Ton gleichzeitig aufgenommen oder muss separat eingespielt werden.
    - *OBS* kann zudem verwendet werden, um Bilder/Videos auf Videoplattformen zu streamen (z.B. Youtube Livestream).
    - *SimpleShow* (Mix aus Aufnahme / Bearbeitungssoftware) dient zur Erstellung von Lernvideos mit einzelnen Folien, auf denen ein Skript anhand von Schlagwörtern von der Software mit einfachen Grafiken illustriert wird. Diese kann man dann nachbearbeiten und durch eigene Bilder ersetzen sowie Untertitel einfügen und das Ganze vertonen.
* **Hardware**: Bei der Anschaffung von Hardware für Videokonferenzen, Podcasts und Video-Tutorials in Bibliotheken ist es wichtig, Qualität und Funktionalität im Auge zu behalten. Mit der richtigen Ausrüstung können Bibliotheken professionelle Ergebnisse erzielen, die Anschaffung muss dabei nicht zwangsläufig teuer sein. Es lohnt sich immer, in eine hochauflösende externe Kamera zu investieren. Auch ein qualitativ hochwertiges Mikrofon bzw. Headset ist unerlässlich für einen überzeugenden Klang. Die in übliche Laptops integrierte Kameras und Mikrofone genügen üblicherweise nicht, um eine zumindest halbwegs professionelle Ton- und Bildqualität zu erzeugen. Sollte überhaupt kein Budget zur Verfügung stehen, ist die entsprechende Hardware in einem guten Smartphone meist besser. Hilfestellung - für nahezu jeden Budgetrahmen - erhält man etwa im Podcast-Forum [*Sendegate*](https://sendegate.de) oder im [Wiki des Netzwerkes „Tutorials in Bibliotheken“](https://bibtutorials.miraheze.org/wiki/Hauptseite#Technik)
* **Umfragen/Abstimmungen** können mittels Tool erstellt werden (*SurveyMonkey*, *Limesurvey*, *Mentimeter*, *Lamapoll*…)
* **URL-Shortener:** verkürzt lange Internetadressen und gibt sie in verkürzter Fassung über einen Drittanbieter wieder: Bit.ly, Goo.gl, Yourls, t1p.de
* **E-Learning-Software:** dienen der Erstellung von Lernformen, die durch elektronische, technische oder digitale Medien unterstützt werden (*Capterra*, *Moodle*)
* **Gamification:** ist die Anwendung spielerischer Elemente im spielfremden Kontext
  + Multimedia-Guide für interaktive Stadtralleys und Handy-Schnitzeljagd [de.actionbound.com](https://de.actionbound.com/)
  + spielbasierte Lernplattform für Quiz usw. https://kahoot.com/

### Planung & Koordination

Die in diesem Absatz beschriebenen Werkzeuge dienen der Aufgaben-Koordination in der Bibliothek. Je mehr Personen in einem Team an gemeinsamen Aufgaben arbeiten, desto empfehlenswerter ist der Einsatz folgender Instrumente:

**Ticketsystem**: Ticketsysteme dienen der strukturierten, regelbasierten (z.T. automatischen) Abarbeitung von Anfragen, Wünschen, Fehlermeldungen von Nutzenden. Diese Systeme können für alle anfallenden Aufgaben innerhalb der Bibliothek verwendet werden, z.B. in der [Softwareentwicklung,](https://it-in-bibliotheken.de/anforderungen.html) bei der Bearbeitung von Kundenanfragen oder internen Prozessen.

Vorteil ist die transparente Bearbeitung von Vorgängen, da die Stati und Bearbeitende jederzeit einsehbar und nachvollziehbar sind. Statusänderungen werden dem Autor je nach Einstellung des Ticketsystems automatisiert mitgeteilt. Alle Bearbeiter:innen haben zu jeder Zeit Einblick und können bei Bedarf übernehmen. Eine Vertretungsregelung oder E-Mail-Weiterleitung wird damit obsolet. Nach abschließender Bearbeitung werden Tickets in der Regel archiviert und können später zu Statistik- und Dokumentationszwecken ausgewertet werden. Auch eine Art FAQ kann automatisiert durch entsprechend markierte Tickets erstellt werden.

Beispiele für Ticketsysteme:

* *Redmine* (Open Source)
* *Jira* (kommerziell)
* Integrierte Issue-Tracker im Rahmen der Softwareentwicklung, beispielsweise in *GitLab* (Open Source) und *GitHub* (kommerziell)

**Kanban-Board**: ein Kanban-Board visualisiert jegliche Arbeitsabläufe anhand einer in Spalten unterteilten Tafel (z.B. Backlog, in Arbeit, fertig). Physisch ist dies mit Post-Its und Whiteboard machbar. Ein Kanban-Board kann auch als Grundlage für ein Projektmanagement nach SCRUM dienen.

Beispiele für Kanban-Software:

* [Deck als Nextcloud-Plugin](https://apps.nextcloud.com/apps/deck) (Open Source)
* [*Kanboard*](https://kanboard.org/) (Open Source)
* [*Trello*](https://trello.com/) (kommerziell)

**Interne und externe Kalender:** Kalender können unterschiedliche Aufgaben erfüllen. Intern zwischen Mitarbeitenden geteilte Kalender helfen bei der Terminplanung und Abstimmung.

Auf der Bibliothekswebsite veröffentlichte Kalender können z.B. buchbare Veranstaltungen für Kund\*innen und Besucher:innen der Bibliothek enthalten. Ebenso ist eine Reservierung von Räumen oder Technik denkbar.

Verbunden mit einem Ticketverkauf oder einer Ticketreservierung können sie die umständliche Kommunikation über E-Mails ablösen und ersparen damit einen großen Verwaltungsaufwand.

Beispiele für Kalender:

* [*SuperSaaS*](https://www.supersaas.de/) (kommerziell)

**Social Media Planung:** Betreibt eine Bibliothek mehrere Social Media Kanäle, kann ein einziges Tool zur Planung und Veröffentlichung von Beiträgen sinnvoll sein. Damit kann der Content zeitlich und inhaltlich vorgeplant werden. Integrierte Analysewerkzeuge helfen zudem bei der Zeitplanung, indem die vergangenen Veröffentlichungen ausgewertet werden und der beste Zeitpunkt für die Zielgruppe ermittelt wird.

Beispiele: [*Hootsuite*](https://www.hootsuite.com/), (kommerziell, in der Basisversion frei nutzbar), [*Buffer*](https://buffer.com/), (kommerziell, in der Basisversion frei nutzbar) [*Trello*](https://trello.com) (kommerziell, in der Basisversion frei verfügbar)

**Wissenslandkarte**: Wissenslandkarten, auch Wissenskarten (*Knowledge Maps*), sind grafische Darstellungen von Wissen in Organisationen. Als Wissenslandkarten werden im Wissensmanagement graphische Verzeichnisse von Wissensträgern, Wissensbeständen, Wissensquellen, Wissensentwicklung, Wissensstrukturen oder Wissensanwendungen bezeichnet. Sie dienen vor allem der Identifikation von Wissen in Unternehmen, um Arbeitsabläufe effektiver und effizienter zu gestalten und referenzieren auf Expertenwissen, Teamwissen, Wissensentwicklungsstationen sowie organisationale Fähigkeiten und Abläufe. Bei dieser Methode wird lediglich der Verweis auf das verankerte Wissen geliefert und nicht das Wissen selbst dort abgelegt.

Quelle: <https://de.wikipedia.org/wiki/Wissenslandkarte>

Beispiele: [*MindManager*](https://www.mindmanager.com/de/) (kommerziell)

**Terminfindung:** Abseits von Kalendern in E-Mail-Programmen können Terminfindungstools sinnvoll sein, insbesondere bei der Planung mit externen Beteiligten.

Beispiele: [*DFN terminplaner*](https://terminplaner.dfn.de/), [*nuudel*](https://nuudel.digitalcourage.de/)

### Echtzeit-Kollaboration

Kollaborative Online-Tools ermöglichen es Teams, in Echtzeit an Dokumenten, Tabellen und Präsentationen zu arbeiten und diese gemeinsam zu nutzen. Sie können die Zusammenarbeit innerhalb einer Bibliothek (und darüber hinaus) deutlich beflügeln und effizienter gestalten, insbesondere wenn man es mit dem oftmals mühsamen Zusammenführen von Dokumentversionen, die in klassischen Office-Lösungen in Einzelarbeit entstanden sind, vergleicht. Das Spektrum reicht dabei vom digitalen Notizzettel bis zum gemeinsam erarbeiteten Förderantrag, vom Brainstorming bis zur fertig ausgearbeiteten Präsentation.

* Für den ad hoc Einsatz bietet sich HedgeDoc als Open-Source-Tool für kollaboratives Schreiben in Echtzeit mit Markdown-Unterstützung an.
* Für die gemeinsame Bearbeitung von klassischen Office-Formaten ist OnlyOffice eine veritable Alternative zu Google Docs, insbesondere unter Gesichtspunkten der Digitalen Souveränität (s.o.).
* Cloudspeicher: die zeitgemäße Variante eines Netzwerklaufwerkes; vgl. z.B. [Nextcloud](https://nextcloud.com/) bzw. [ownCloud](https://owncloud.com/de/) das gut mit *OnlyOffice* zusammenspielt (s.o.)
* Online-Whiteboards sind digitale Tools, die ein traditionelles physisches Whiteboard emulieren. Sie ermöglichen es, in Echtzeit gemeinsam zu brainstormen, indem etwa Post-Its in den virtuellen Raum geklebt werden, aber auch um zu zeichnen und multimediale Inhalte zu integrieren. Damit eignen sie sich gut sowohl für die kreative Zusammenarbeit innerhalb der Bibliothek als auch für die Durchführung von Workshops und anderen Formaten mit Nutzenden. Verbreitet sind [*Conceptboard*](https://conceptboard.com/de/) und [*miro*](https://miro.com/), ernstzunehmende offene Alternativen gibt es leider noch nicht.

## Social Intranet

Bei der Vielzahl der oben genannten Werkzeuge und Methoden besteht die Gefahr, nach und nach eine immer weiter fragmentierte Softwarelandschaft zu erzeugen, mit entsprechend hohen Anforderungen an Betrieb, Dokumentation und Schulungen. Ein sinnvoller Ansatz ist es deshalb, zunächst das Verhältnis der verschiedenen Lösungen auf konzeptioneller Ebene klar festzulegen, um Redundanzen zu vermeiden. Dieses Konzept kann sich, soweit möglich, aber auch in einer technisch integrierten Lösung für möglichst viele der genannten Aspekte widerspiegeln: einem Social Intranet.

Ein Social Intranet ist ein internes, web-basiertes Netzwerk, das speziell darauf ausgelegt ist, die Kommunikation, Zusammenarbeit und Informationsverbreitung innerhalb einer Organisation zu fördern. Es kombiniert die klassische Funktion eines Intranets, die Informationserstellung und -bereitstellung, mit eher sozialen, kommunikativen Funktionen und ermöglicht den Mitarbeitenden so den einfachen Austausch von Informationen und Ideen. Ein gut geplantes und eingerichtetes Social Intranet ist in der Lage, die meisten der oben genannten Aspekte und Ideen abzudecken. Nicht immer in der vollen Funktionstiefe, dafür aber in einer einheitlichen, gut nutz- und administrierbaren Oberfläche. Auch sonst nur aufwändig implementierbare Features wie eine Volltextsuche über alle Materialien wird problemlos möglich.

Um im Bild vom Anfang des Kapitels zu bleiben: ein Social Intranet ist weder ein Hammer, noch ein schwerer Werkzeugkasten, sondern am ehesten ein Schweizer Taschenmesser.

Kern der Idee eines Social Intranets ist, dass neben Dokumenten, die Menschen und deren Austausch untereinander im Fokus stehen, weshalb jede\*r Mitarbeiter\*in ein Profil pflegen kann, um seine Rolle, Fachgebiete und Kontaktinformationen zu teilen. Je nachdem, ob eher die sozialen Features bzw. themenspezifische Diskussionsräume oder die Informationsbereitstellung im Vordergrund steht, bieten sich zur Implementierung Wiki-Software (etwa: Confluence, kommerziell) oder Foren-Software (etwa: Discourse, Open Source) an.

## Zusammenfassung und Ausblick

Die externe Kommunikation dient dazu, Beziehungen zu externen Stakeholdern zu pflegen. Dafür werden Webseiten, E-Mail-Marketing, Social Media, Medienproduktionstools und weitere genutzt. Die interne Kommunikation konzentriert sich auf den Informationsaustausch innerhalb einer Organisation. Hier kommen Tools wie Social Intranets, Kollaborationstools und Videokonferenzsysteme zum Einsatz. Die aktuellen Entwicklungen im Bereich der Large Language Models (LLM) haben das Potenzial, die Art und Weise, wie Organisationen sowohl intern als auch extern kommunizieren, tiefgreifend zu verändern. Sie können bspw. in Chatbots und andere Support-Systeme integriert werden, um Anfragen von Nutzenden - zumindest initial - in Echtzeit zu beantworten und so Mitarbeitende zu entlasten. Auch bei der Erstellung von Content, können sie helfen, diesen schneller und kohärenter zu gestalten. Im Bereich des Wissensmanagements könnten Mitarbeiter\*innen LLMs nutzen, um spezifische, auch komplexe Fragen zu beantworten oder Hintergrundinformationen zu bestimmten Themen zu erhalten ohne dabei ständig auf menschliche Experten zurückgreifen zu müssen. Es ist jedoch wichtig, dabei ethische Überlegungen und Fragen zur Datenvertraulichkeit zu berücksichtigen, insbesondere wenn LLMs in sensiblen Kommunikationsbereichen eingesetzt werden.

# Literaturverzeichnis

Hier nur die explizit im Handbuch referenzierten Publikationen. Die vollständige Bibliographie wird [in einer gemeinsamen Zotero-Gruppe](https://www.zotero.org/groups/4673379/it_in_bibliotheken) verwaltet.

Assfalg, Rolf. 2023. „Metadaten“. In *Grundlagen der Informationswissenschaft*, herausgegeben von Rainer Kuhlen, Dirk Lewandowski, Wolfgang Semar, und Christa Womser-Hacker, 7. Ausgabe, 245–56. Berlin, Boston: De Gruyter Saur. <https://doi.org/10.1515/9783110769043-021>.

Bach, Nicolas. 2022. „Das Handbuch IT in Bibliotheken: Einblicke in den ersten bibliothekarischen Book Sprint Deutschlands“. *Informationspraxis* 8 (1). <https://doi.org/10.11588/ip.2022.1.94475>.

Borgman, Christine L. 1997. „From Acting Locally to Thinking Globally: A Brief History of Library Automation“. *The Library Quarterly: Information, Community, Policy* 67 (3): 215–49. <https://www.jstor.org/stable/40039721>.

Breeding, Marshall. 2022. „How to Secure Library Systems From Malware, Ransomware, and Other Cyberthreats“. 2022. <https://www.infotoday.com/cilmag/jan22/Breeding--How-to-Secure-Library-Systems-From-Malware-Ransomware-and-Other-Cyberthreats.shtml>.

Breeding, Marshall. o. J. „Library Technology Industry Mergers and Acquisitions. Library Technology Guides“. Zugegriffen 28. April 2022. <http://librarytechnology.org/mergers/>.

Christensen, Anne. 2022. „Wissenschaftliche Literatur entdecken: Was bibliothekarische Discovery-Systeme von der Konkurrenz lernen und was sie ihr zeigen können“. *LIBREAS*, Nr. 41. <https://doi.org/10.18452/24798>.

Christensen, Anne, und Frank Seeliger. 2022. „„Wie schreiben wir gemeinsam ein nützliches Buch?”“. *b.i.t.online* 25 (6): 509–10. <https://www.b-i-t-online.de/heft/2022-06-nachrichtenbeitrag-christensen.pdf>.

Deutsches Institut für Normung e. V. (DIN). 2020. „DIN EN ISO 9241-110 Ergonomie der Mensch-System-Interaktion - Teil 110: Interaktionsprinzipien (ISO 9241-110:2020)“. <https://www.din.de/de/mitwirken/normenausschuesse/naerg/veroeffentlichungen/wdc-beuth:din21:320862700>.

DFG. 2021. „Datentracking in der Wissenschaft: Aggregation und Verwendung bzw. Verkauf von Nutzungsdaten durch Wissenschaftsverlage. Ein Informationspapier des Ausschusses für Wissenschaftliche Bibliotheken und Informationssysteme der Deutschen Forschungsgemeinschaft“. Deutsche Forschungsgemeinschaft. <https://www.dfg.de/download/pdf/foerderung/programme/lis/datentracking_papier_de.pdf>.

Freyberg, Linda, und Sabine Wolf, Hrsg. 2019. *Smart Libraries: Konzepte, Methoden und Strategien*. b.i.t.online Innovativ 76. Wiesbaden: b.i.t. Verlag.

Gesetzliche, Deutsche, und Unfallversicherung e.V. (DGUV). 2019. „Bildschirm- und Büroarbeitsplätze: Leitfaden für die Gestaltung“. <https://publikationen.dguv.de/widgets/pdf/download/article/409>.

Gould, J. D., und C. Lewis. 1987. „Designing for usability: Key principles and what designers think“. In *Human-computer interaction: a multidisciplinary approach*, 528–39. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

„Government Design Principles. GOV.UK“. 2012. 2012. <https://www.gov.uk/guidance/government-design-principles>.

Hanson, Cody. 2015. „Opinion: Libraries are Software“. 2015. <https://www.codyh.com/writing/software.html>.

Holländer, Stephan. 2023. „Cyberangriffe auf Universitäten, Fachhochschulen und deren Bibliotheken – ein unterschätztes Problem?“ *B.I.T. online* 26 (3). <https://www.b-i-t-online.de/heft/2023-03-nachrichtenbeitrag-hollaender.pdf>.

Jakob Nielsen. 2000. „Why You Only Need to Test with 5 Users. Nielsen Norman Group“. 2000. <https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>.

Jetter, Hans-Christian. 2023. „Informationsvisualisierung und Visual Analytics“. In *Grundlagen der Informationswissenschaft*, herausgegeben von Rainer Kuhlen, Dirk Lewandowski, Wolfgang Semar, und Christa Womser-Hacker, 7. Ausgabe, 295–306. Berlin, Boston: De Gruyter Saur. <https://www.degruyter.com/document/doi/10.1515/9783110769043-025/pdf>.

Kern, Christian. 2011. *RFID für Bibliotheken*. Springer.

Kling, Rob, und Susan Leigh Star. 1998. „Human centered systems in the perspective of organizational and social informatics“. *ACM SIGCAS Computers and Society* 28 (1): 22–29. <https://doi.org/10.1145/277351.277356>.

Kluge, Matthias. 2022. *Anbieter von Bibliothekssoftware*. Landesfachstelle für das öffentliche Bibliothekswesen. <https://www.oebib.de/bau-einrichtung-it/it-und-internet/bibliothekssoftware>.

Kost, Michael, Bastian Loibl, Peter Reuter, und Matthias Stenke. 2022. „#JLUoffline. Der Cyber-Angriff auf die Justus-Liebig-Universität Gießen im Dezember 2019“. *ABI Technik* 42 (1). https://doi.org/<https://doi.org/10.1515/abitech-2022-0005>.

Kuhlen, Rainer, Dirk Lewandowski, Wolfgang Semar, und Christa Womser-Hacker, Hrsg. 2023. *Grundlagen der Informationswissenschaft*. 7. Ausgabe. Berlin, Boston: De Gruyter Saur. <https://doi.org/doi:10.1515/9783110769043>.

Matthews, Joseph R., und Carson Block. 2020. *Library information systems*. Second edition. Library and information science text series. Santa Barbara, California: Libraries Unlimited.

MDR. 2023. „Vermehrt Hackerangriffe auf Hochschulen und Universitäten“. 25. Januar 2023. <https://www.mdr.de/wissen/vermehrt-hackerangriffe-auf-hochschulen-und-universitaeten-100.html>.

Michaelis, Barbara. 2014. In *RFID für Bibliothekare: ein Vademecum*, herausgegeben von Frank Seeliger, 3. Auflage, 145–50. Verlag News & Media. <https://doi.org/10.15771/RFID_2014_13>.

Müller, Heiko. 2023. „Vorsorge für den Angriffsfall: Der Weg in die Cloud.“ *iX Magazin für professionelle IT* 11.

Rehm, Stefan-Marc. 2023. „IT-Sicherheitsbeauftragten bestellen. www.haufe.de“. 23. August 2023. <https://www.haufe.de/compliance/management-praxis/cybersicherheit-it-sicherheitsbeauftragter_230130_447256.html>.

Rölke, Heiko, und Albert Weichselbraun. 2023. „Ontologien und Linked Open Data“. In *Grundlagen der Informationswissenschaft*, herausgegeben von Rainer Kuhlen, Dirk Lewandowski, Wolfgang Semar, und Christa Womser-Hacker, 7. Ausgabe, 257–70. Berlin, Boston: De Gruyter Saur. <https://www.degruyter.com/document/doi/10.1515/9783110769043-022/pdf>.

Schmidt, Caroline. 2023. „Wann die Bestellung eines Datenschutzbeauftragten für Ihr Unternehmen unabdingbar ist. www.e-recht24.de“. 12. Juni 2023. <https://www.e-recht24.de/datenschutz/10744-datenschutzbeauftragter-dsgvo.html#>.

Schweitzer, Roswitha. 2016. „Anforderungen an ein Bibliothekssystem der neuen Generation - der Kriterienkatalog von hbz und VZG“. Köln. <https://docplayer.org/61296444-Anforderungen-an-ein-bibliothekssystem-der-neuen-generation.html>.

Seeliger, Frank, Hrsg. 2014. *RFID für Bibliothekare: ein Vademecum*. 3. Auflage. Verlag News & Media. <https://doi.org/10.15771/978-3-936527-32-2>.

———. 2019. „Smart Services als Marketinginstrument“. In *Praxishandbuch Informationsmarketing: Konvergente Strategien, Methoden und Konzepte*, herausgegeben von Frauke Schade und Ursula Georgy, 343–57. De Gruyter Saur. <https://doi.org/10.1515/9783110539011-023>.

Shneiderman, Ben, und Catherine Plaisant. 2005. *Designing the user interface. Strategies for effective human-computer interaction*. 4th Aufl. Pearson.

Siems, Renke. 2022. „Das Lesen der Anderen: Die Auswirkungen von User Tracking auf Bibliotheken“. *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB* 9 (1): 1–25. <https://doi.org/10.5282/o-bib/5797>.

Steilen, Gerald. 2012. „Discovery-Systeme - die OPACs der Zukunft?“ Hamburg. <https://www.slideshare.net/steilen/discoverysysteme-die-opacs-der-zukunft>.

„Tracking in der Wissenschaft: So können Bibliotheken Daten und Wissenschaftsfreiheit schützen“. 2022. <https://www.zbw-mediatalk.eu/de/2022/01/tracking-in-der-wissenschaft-so-koennen-bibliotheken-daten-und-wissenschaftsfreiheit-schuetzen>.

„UHF RFID Almanach 2022“. 2022. EECC. <https://eecc.info/rfidalmanach.html>.

Voß, Jakob. 2021. „Datenqualität als Grundlage qualitativer Inhaltserschließung“. In *Qualität in der Inhaltserschließung*, 167–76. De Gruyter Saur. <https://doi.org/10.1515/9783110691597-010>.

———. 2022. „Einführung in die Verarbeitung von PICA-Daten“. 2022. <https://pro4bib.github.io/pica/>.

# Abkürzungsverzeichnis

[API](https://de.wikipedia.org/wiki/Programmierschnittstelle)

Application Programming Interface, Programmierschnittstelle

[Avram](https://format.gbv.de/schema/avram/specification)

Schemasprache für feldbasierte Datenformate, insbesondere MARC und PICA+

BMS

Bibliotheksmanagementsystem, gebräuchlich auch als LMS (Library Management System)

[COUNTER](https://www.projectcounter.org/)

Counting Online Usage of NeTworked Electronic Resources

[EDIFACT](https://de.wikipedia.org/wiki/EDIFACT)

Electronic Data Interchange for Administration, Commerce and Transport

[ERM](https://en.wikipedia.org/wiki/Electronic_resource_management)

Electronic Resource Management, Verwaltung von Lizenzinformationen zu E-Ressourcen

[ERP](https://de.wikipedia.org/wiki/Enterprise-Resource-Planning)

Enterprise-Ressource-Planning, Ressourcen wie z.B. Finanzen unternehmerisch überwachen und planen

FID

Fachinformationsdienst

[FOLIO](https://www.folio-bib.org/)

Open-Source Bibliotheksmanagementsystem

[GOKb](https://gokb.org/)

Global Open Knowledgebase, Freie Datenbank zu E-Ressourcen, insbesondere für ERM

ID

Eindeutige Referenz auf einen Datensatz (Identifikator)

[IDM](https://en.wikipedia.org/wiki/Identity_management)

Identity Management, das Speichern von Metadaten zu Personen

[KBART](https://www.niso.org/standards-committees/kbart)

Knowledgebases and related tools, Datenformat zum Transfer von Metadaten

[MARC](https://de.wikipedia.org/wiki/Machine-Readable_Cataloging)

MAchine-Readable Cataloging, ältestes und noch immer wichtigstes bibliothekarisches Austauschformat

[OCR](https://de.wikipedia.org/wiki/Texterkennung)

Automatische Erkennung von Texten in Bildern (Optical Character Recognition)

[OPAC](https://de.wikipedia.org/wiki/OPAC)

Online Public Access Catalogue, Katalog einer Bibliothek

[SIP2](https://en.wikipedia.org/wiki/Standard_Interchange_Protocol)

Protokoll zur Anbindung von Selbstbedienungsautomaten an Bibliothekssysteme

[SaaS](https://de.wikipedia.org/wiki/Software_as_a_Service)

Software as a Service, Software und Hardware bei externem Dienstleister

[WCAG](https://www.w3.org/WAI/standards-guidelines/wcag/)

Web Content Accessibilty Guidelines

Z39.50

Protokoll zur Abfrage von bibliografischen Daten

# Abbildungsverzeichnis

*Hier fehlt noch ein vollständiges Abbildungsverzeichnis mit Quellenangaben*

* media/Sketch\_FIS\_FDM\_CC\_BY\_Mau.png Sketchnote: Forschungsinformationen und ihre Sicht auf Forschungsdaten. <https://doi.org/10.5281/zenodo.4899000> CC-BY 4.0 [Franziska Mau](https://orcid.org/0000-0001-7701-0301)
* Sketchnote: Auf der Suche nach dem heiligen Gral - Forschungsinformationssysteme. <https://doi.org/10.5281/zenodo.4388855> CC-BY 4.0 [Franziska Mau](https://orcid.org/0000-0001-7701-0301)
* media/sso-approaches.svg Three aproaches to Single Sign On. <https://commons.wikimedia.org/wiki/File:Single_sign_on_aproaches.svg> Public Domain

# Hinweise zur Mitarbeit

Dieses Handbuch wird in einem offenen Prozess erarbeitet und verbessert. Mitarbeit ist dabei grundsätzlich sehr willkommen und in verschiedener Form möglich. Voraussetzung ist ein Verständnis des [Arbeitsablauf](#arbeitsablauf), Einhaltung des [Styleguide](#styleguide) und die freie [Lizenzierung](#lizenz) der eigenen Beiträge.

## Arbeitsablauf

Alle Kapitel wurden in Booksprints vor Ort geplant und in Form von Google Docs Dokumenten geschrieben. Diese wurden anschließend einmalig ins Markdown-Format übertragen und auf [it-in-bibliotheken.de](https://it-in-bibliotheken.de) veröffentlicht (siehe [Technik](#technik)).

Weitere Änderungen und Beiträge zum Handbuch sind auf folgenden Wegen möglich:

1. **Bearbeitungsvorschläge und Kommentare** in Google Docs („Kapitel Kommentieren“): im Zweifelsfall die einfachste Methode
2. Direkte Bearbeitung der Markdown-Quellen durch einen **Pull-Request** („Seite editieren“): erfordert einen GitHub-Account und rudimentäre Markdown-Kenntnisse
3. Öffentlicher **Kommentar als** [**GitHub-Issue**](https://github.com/pro4bib/handbuch-it-in-bibliotheken/issues) („Problem melden“): erfordert einen GitHub-Account
4. Persönlicher Hinweis an [die Autor\*innen](#autorinnen)

|  |
| --- |
| Feedback-Links unter dem Inhaltsverzeichnis rechts |

Die Bearbeitungsvorschläge und Kommentare aus Google Docs werden in unregelmäßigen Abständen von der Redaktion in die Markdown-Quellen übernommen und die Google Docs Dokumente anschließend aktualisiert. Für [Bilder](#mitarbeit-medien), [Literaturverzeichnis](#mitarbeit-literatur) und [Glossar](#mitarbeit-glossar) gibt es eigene Abläufe. Der aktuelle Workflow nach Ende der Booksprints beinhaltet folgende Schritte:

Lektorat

Fertige Kapitel stehen auf der Webseite und als Google Docs Dokument zum Korrekturlesen und Kommentieren bereit. Je Kapitel sind zwei externe (bisher nicht am Schreiben beteiligte) Lektor\*innen eingeteilt, um die Einhaltung des [Styleguide](#styleguide) zu überprüfen und Änderungen vorzuschlagen.

Redaktion

Hinweise und Änderungsvorschläge werden von den Autor\*innen diskutiert und entschieden. Falls die daraus resultierenden Änderungsvorschläge nicht als Pull-Request vorliegen, müssen sie in Google-Docs bestätigt und parallel in die Markdown-Quellen eingepflegt werden.

Publikation

Bestätigte Änderungen an der Markdown-Dateien im git-Repository führen direkt dazu, dass das Buch automatisch aktualisiert wird (siehe [Technik](#technik), bisher nur HTML-Version, Druckversion ist geplant).

## Styleguide

Dieser Styleguide soll eine einheitliche Form trotz [unterschiedlicher Autor\*innen](#mitarbeit-autorinnen) gewährleisten. Dazu gibt es Hinweise zu [Zielgruppe](#zielgruppe), [Stil und Aktualität](#schreibstil), [Schreibweise](#schreibweise) und [Struktur](#struktur) sowie Vorgaben zu besonderen Inhalten wie [Bildern und anderen Medien](#mitarbeit-medien), [Literaturverzeichnis](#mitarbeit-literatur) und [Abkürzungsverzeichnis](#mitarbeit-glossar).

Bei der Bearbeitung sollten folgende Grundsätze beachtet werden:

* Wir verzichten auf individuelle Autor\*innenschaft an einzelnen Textteilen. Alle können an allen Teilen mitarbeiten.
* Das Handbuch ist keine wissenschaftliche Arbeit, sondern soll einen Überblick geben. Für Details kann auf weiterführende Quellen verwiesen werden.

### Zielgruppe

Zur Klärung der Zielgruppe dieses Handbuchs wurden einige sogenannte *Personas* definiert:

* **Janine** Buchinger: Janine leitet die Stadtbibliothek in einer Stadt mit 250.000 Einwohnern. Die Bibliothek besteht aus einer Zentrale und zwei Zweigstellen. Mit den Schulbüchereien besteht eine Kooperation für fachliche Beratung und gemeinsame Aktivitäten bei der Informationskompetenz-Vermittlung.
* Dr. **Tillmann** Schuppe: Tillmann ist Leiter einer Fachhochschulbibliothek mit 500.000 Medieneinheiten. Die Bibliothek gehört einem Bibliotheksverbund an. Die Bibliothek plant einen Neubau, der gemeinsam mit dem Rechen- und Medienzentrum bezogen werden soll.
* **Magda** Olsowski: Magda ist studierte Informatikerin und leitet die Gruppe Forschungsdatenmanagement an einer großen Universitätsbibliothek. Sie hat keine bibliothekarische Vorbildung.
* **Alicia** Meyer: Alicia studiert Bibliotheksmanagement und plant eine Masterarbeit, in der sie die Implementierungsprozesse von Software analysieren möchte.
* **Robert** Pohlmann: Robert leitet die IT-Abteilung einer mittelgroßen Universitätsbibliothek und ist nebenberuflich Lehrbeauftragter für einen bibliothekarischen Studiengang.

### Stil und Aktualität

* **Journalistische oder enzyklopädische Neutralität sind nicht oberstes Prinzip dieses Buches.** Es soll vielmehr fundiert und praxisorientiert informieren und beraten und darf dabei auch parteiisch sein.
* **Dieses Buch ist keine wissenschaftliche Forschungsveröffentlichung.** Nicht jede Aussage muss mit einer Quelle belegt werden. Für die Anwendung in der Hochschullehre reicht es, wenn das Buch zentrale Aussagen belegt bzw. auf die wichtigsten aktuellen Studien verweist und somit auch Tipps zur weiterführenden Lektüre bietet.
* **Das Buch sollte in 2-5 Jahren noch aktuell und verständlich sein, aber nicht mehr unbedingt in 10 Jahren.**
* **Das Buch soll als Nachschlagewerk dienen, das nicht vollständig durchgelesen werden muss.** Dabei helfen Redundanz und Querverweise (siehe [Hinweise zur Struktur](#struktur)).
* **Zumindest die einzelnen Kapitel sollen sich gut durchlesen lassen.** Insbesondere Brüche in Stil und Aussage, Widersprüche oder unnötige Redundanzen sind innerhalb der Kapitel unbedingt zu vermeiden. Unter Berücksichtung des Nachschlagewerk-Charakters streben wir auch einen guten Lesefluss innerhalb der Kapitel an.

### Struktur

#### Struktur des Texts

* Wir verwenden **kurze, unverschachtelte Sätze**.
* Wir erzeugen **Sinnabschnitte**, die möglichst für sich stehend verständlich sind.
* Wir schreiben **stark strukturiert**, also
  + mit vielen **Zwischenüberschriften**, bis maximal zur vierten Gliederungsebene,
  + wo es inhaltlich passt, in stichpunktartigen **Listen** und
  + mit Hervorhebung **wichtiger Begriff durch Fettdruck** als Gliederungshilfe.
* Wir verwenden [**Infoboxen**](https://quarto.org/docs/authoring/callouts.html), die auch unabhängig vom übrigen Text lesbar sind.
* Wir liefern **wichtige Informationen zusätzlich** zum Text in Form von Bildern, Tabellen, Listen, Infoboxen und/oder Zusammenfassungen.

#### Struktur der Hauptkapitel

Das Handbuch behandelt [aufgeteilt in Hauptkapitel](#inhalt) die wesentlichen Themen rund um IT in Bibliotheken.

* Jedes Kapitel beginnt mit einer **Kurzfassung** als Infobox gefolgt von einer **Einleitung** und endet mit einem Abschnitt **Zusammenfassung und Ausblick**.
* Kapitel haben normalerweise einen Umfang von rund 4.000 Wörtern. Deutlich längere Kapitel sind darauf zu prüfen, ob sie sich in mehrere Kapitel trennen lassen, und wenn das nicht möglich ist, müssen sie sorgfältig in Unterkapitel aufgeteilt werden.
* Die Strukturierung in Unterkapitel sollte stimmig und ausgewogen sein, insbesondere mit Blick auf die Ziele des Buches.
* Die Übergänge zwischen einzelnen Unterabschnitten sollten stimmig sein.
* Jedes Kapitel beinhaltet ein aussagekräftiges Metadatenfeld description [für Suchmaschinen](https://developers.google.com/search/docs/advanced/appearance/snippet) (maximal 158 Zeichen)

### Schreibweise, Fachbegriffe und Verweise

* Wir verwenden im gesamten Buch gendergerechte Schreibweise mit Sternchen (\*). In Markdown ist es sicherer dem Sternchen einen Backslash voranzustellen, z.B. Autor\\*innen.
* Eine Schreibweise für häufig verwendete Fachbegriffe sollte quer durch das Buch eingehalten werden, so z.B. BMS für Bibliotheksmanagementsysteme
* Fachbegriffe (z.B. Bibliotheksverbund) werden dort verwendet, wo sie wiederholt relevant sind, und werden bei ihrer ersten Erwähnung definiert. Abkürzungen werden zusätzlich [im Abkürzungsverzeichnis](#mitarbeit-glossar) erfasst.
* Wir vermeiden IT-Jargon.
* Wir vermeiden Substantivierung („Digitalisierungswürdigkeit“) und unnötige Passiv-Formen.
* Quellen sollten nur dann genannt werden wenn in der jeweiligen Textpassage auch wirklich paraphrasiert oder wörtlich zitiert wird.

#### Typografie

* Abkürzungen werden durch geschützte Leerzeichen getrennt (z. B.)
* Gerade Anführungszeichen "..." werden automatisch durch die Deutschland und Österreich übliche Anführungszeichen („…“) ersetzt.

Externe Links, die nur auf Anbieter oder andere Websites verweisen, werden inline verlinkt.

* Links, die auf später entstehende Kapitel verweisen, werden durch eckige Klammern kenntlich gemacht.
* *Kursive* Hervorhebung sollte nur für Namen und [Glossareinträge](#mitarbeit-glossar) verwendet werden.
* Wir verzichten auf Fußnoten.

### Bilder und andere Medien

* Bilder und andere Mediendateien kommen in das Verzeichnis media im [git-Repository](#technik). Alternativ können sie von externen Quellen per URL eingebunden werden wenn die Quelle voraussichtlich dauerhaft verfügbar ist.
* Bilder sollten möglichst als Vektorgrafik (SVG) bereitgestellt werden.
* Standardschriftart ist [Source Sans Pro](https://de.wikipedia.org/wiki/Source_Sans_Pro) und Standard-Farbe für Hervorhebungen ist Blau mit dem Farbcode #2780e3 und darauf aufbauende Hellblau-Töne.
* Bitte nutzt sprechende Dateinamen!

### Literaturverzeichnis

Die zitierte und weiterführende Literatur wird in einer Zotero-Gruppe unter <https://www.zotero.org/groups/4673379/it_in_bibliotheken> verwaltet. Der BibLaTex-Export dieser Bibliographie wird mit Aufruf von make refs von dort heruntergeladen und unter references.bib gespeichert. Diese Datei sollte also nicht direkt bearbeitet werden! Innerhalb des Markdown-Quelltext kann mittels [Pandoc-Citation Syntax](https://quarto.org/docs/authoring/footnotes-and-citations.html#sec-citations) und dem jeweiligen Citekey aus references.bib auf Literatur verwiesen werden.

### Glossar

Das Abkürzungsverzeichnis in der Datei glossar.yml enthält erklärungswürdige Begriffe mit Kurzbeschreibung und optionalem Link auf eine weiterführende Quelle (meist Wikipedia). Es werden \_keine\_ Firmennamen aufgenommen, auch wenn sie Akronyme sind. Die Glossarbegriffe werden in den Textdateien zur Hervorhebung *kursiv* gesetzt (in Markdown so ein \*Fachbegriff\*). Bei Erzeugung der HTML-Version des Handbuchs wird die Hervorhebung in einen Tooltip umgewandelt.

### Autor\*innen-Verzeichnis

Wenn Du etwas beigetragen hast und möchtest, dass Du im [Verzeichnis der Autor\*innen](#autorinnen) auftauchst, trage Dich in der YAML-Datei [contributors.yml](https://github.com/pro4bib/handbuch-it-in-bibliotheken/blob/main/contributors.yml) ein. Die Einträge sollten nach Nachname sortiert werden. Die Felder email, position und orcid sind optional.

## Technik

Die Master-Version des Handbuch liegt in einem git-Repository unter <https://github.com/pro4bib/handbuch-it-in-bibliotheken>. Die Ergebnisdateien werden automatisch via GitHub und einen Server der *VZG* mit der Software [quarto](https://quarto.org/) aktualisiert, so dass unter <https://it-in-bibliotheken.de/> immer der aktuellste Stand einsehbar sein sollte.

### Verzeichnisstruktur

Die Markdown-Dateien im Wurzelverzeichnis (\*.md) sind die Masterdateien.

* metadata.yml bibliographische Metadaten (Titel, Abstract…)
* \_quarto.yml zentrale Konfigurationsdatei zur Anpassung der [Konvertierung mit Quarto](#konvertierung)
* <contributors.yml> [Autor\*innen-Verzeichnis](#autorinnen-verzeichnis)
* references.bib [Literaturverzeichnis](#literaturverzeichnis) (bitte nicht direkt bearbeiten!)

Weitere Unterverzeichnisse:

* media/ [Bilder und andere Medien](#bilder-und-andere-medien)

Die Dateien in folgenden Verzeichnissen sollen nicht per Hand geändert werden:

* docs/ aus den Masterdateien mit quarto erzeugte Publikation
* \_gdrive/ von bzw. nach Google-Drive importierte bzw. exportierte Kapitel (siehe [README.md](_gdrive/README.md))

### Konvertierung

Zur Anpassung der Konvertierung des Handbuchs mit [quarto](https://quarto.org/) muss das Repository lokale geklont und Quarto installiert werden. Für die DOCX-Ausgabe muss außerdem rsvg-convert installiert werden (Paket librsvg2-bin bzw. libsrvg).

Die Aufrufe sind zur Vereinfachung in [Makefile](../Makefile) zusammengefasst:

* make preview konvertiert das Handbuch nach HTML und ermöglicht eine Vorschau unter <http://localhost:15745/> (der Port ist die Postleitzahl von Wildau). Die HTML-Ansicht wird automatisch aktualisiert wenn die Quelldateien lokale geändert werden.
* make build konvertiert das Handbuch in alle konfigurierten Formate und legt die Ergebnisse im Verzeichnis \_book ab. Dieser Schritt wird auch automatisch nach jedem Push auf GitHub ausgeführt.
* make all ruft make build und docx auf und kopiert die DOCX-Dateien ins Publikationsverzeichnis \_book.
* make html erzeugt nur HTML in \_book.
* make docx erzeugt nur DOCX in \_book.
* make refs aktualisiert das [Literaturverzeichnis](#literaturverzeichnis) in references.bib von Zotero.

## Lizenz

Alle Beiträge werden unter der Lizenz *Creative Commons Namensnennung 4.0 International* ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/legalcode.de)) veröffentlicht. Für Abbildungen kann auch eine CC-BY-Lizenz (kein -NC oder -ND) verwendet werden.