

Handbuch zur Erstellung diskriminierungssensibler Metadaten für historische Quellen und Forschungsdaten

Schlüsselbegriffe und Erfahrungen aus der Praxis

Levyn Bürki

Moritz Mähr

Noëlle Schnegg

2024-10-31

Dieses Handbuch ist ein Leitfaden zur Erstellung von diskriminierungssensibler Metadaten für historische Quellen und Forschungsdaten. Es richtet sich an Historiker*innen, Archivar*innen, Bibliothekar*innen und alle, die sich mit Datenethik und Open Research Data in den Geschichtswissenschaften beschäftigen. Die Autor*innen Levyn Bürki, Moritz Mähr und Noëlle Schnegg führen durch die praktischen Aspekte der Erstellung von Metadaten, basierend auf den FAIR- und CARE-Prinzipien, um Forschungsdaten auffindbar, zugänglich, interoperabel und nachnutzbar zu machen. Durch praktische Anleitungen und illustrierte Fallbeispiele zeigt das Handbuch, wie maschinenlesbare Metadaten Forschung und Lehre bereichern und die Interpretation historischer Quellen beeinflussen können. Als öffentlich zugängliches “Living Document” ist es auf eine kontinuierliche Weiterentwicklung durch die Community ausgelegt und verpflichtet sich zu einer inklusiven und diskriminierungsfreien Darstellung historischer Inhalte. Das Handbuch ist eine grundlegende Ressource für alle, die sich mit moderner digitaler Geschichtswissenschaft und Open Research Data beschäftigen wollen.



Content Note

Dieses Dokument enthält Abbildungen von historischen Quellen, die diskriminierende Sprache, Bilder oder Darstellungen enthalten. Sie sind Ausdruck von Vorurteilen, Stereotypen oder Gewalt gegen bestimmte Gruppen in der Vergangenheit.

1 Einleitung

- Warum dieses Handbuch?

- *Aktualität Thema:*
- *Abgrenzung zu anderen Projekten/Leitfäden/Handbüchern:*
- *Zielgruppe:* Das Handbuch richtet sich an Historiker*innen, Archivarinnen, Bibliothekar*innen und alle, die sich mit Open Research Data in den Geschichtswissenschaften beschäftigen.

- **Entstehungskontext**

- Eine erste Auflage dieses Handbuchs entstand im Rahmen des historischen Forschungsprojekts *Stadt.Geschichte.Basel* und wurde im Juni 2024 von Moritz Mähr und Noëlle Schnegg publiziert. Gemeinsam mit Levyn Bürki wurde das Handbuch ab Herbst 2024 schliesslich neu strukturiert und überarbeitet.
 - * Einarbeitung des Feedbacks aus der Erstveröffentlichung sowie von Diskussionen, welche seither an Workshops und Konferenzen geführt.
 - * Titeländerung
 - * Eigene Positionalität

- **Problemdefinitionen**

- **Diskriminierungsbegriff und Datenethik:**
 - * Diskriminierung kann viele Formen annehmen und geht über diskriminierendes Vokabular hinaus.
 - * Darunter Rassismus, Antisemitismus, Sexismus, Ableismus, Transphobie und viele andere, und sie kann mehrere Diskriminierungsformen gleichzeitig enthalten (Intersektionalität).
 - * Sie kann explizit oder implizit sein und ist oft tief in den Kontext und die Interpretation dieser Ressourcen eingebettet.
 - * Implizite oder strukturelle Formen von Diskriminierung finden sich in vielen bestehenden Thesauri und Schlagwortverzeichnissen wie der Gemeinsamen Normdatei (GND), der am weitesten verbreiteten Normdatei im deutschsprachigen Raum.
 - * Durch Basisfelder, Klassifizierung nach hegemonialer Praxis, ... Bei einem Textfeld geht es immer auch darum, was es *kann*: wie wird es ausgelesen? Interpretiert? Dargestellt? Übermittelt? Dokumentiert?
 - * Sie stecken aber auch in institutionellen Strukturen/Bias, kolonialen Sammlungskonzepten, ...
 - * Schliesslich sind auch Aspekte des Datenschutz und Selbstbestimmung ein wesentlicher Teil eines diskriminierungssensiblen Umgangs mit Daten und Metadaten. Im Kontext indigener Daten existieren dazu bereits die CARE-Prinzipien, welche auch für andere Kontexte, etwa den Umgang mit Daten aus nationalsozialistischen Kontexten wegweisend sein können. Datenethische Grundsätze (Do-no-harm Prinzip, Schadensminimierung, Collectiv benefit).
 - * Unter diskriminierungssensiblen Arbeiten an und mit Archiv- und Forschungsdaten verstehen wir unmittelbar auch eine Arbeit an Forschungsdatenqualität.

- **Beispiele, Fragen und Konfliktlinien:** Inhaltlich, strukturell, technisch und politisch:
 - * Wie können Metadaten transparent gestaltet (etwa hinsichtlich Versionierung und Autor:innenschaft) und wie können Entscheidungsprozesse dokumentiert werden?
 - * Wie soll mit Zielkonflikten (etwa mit Förderrichtlinien) umgegangen werden?
 - * Wie gelingt diskriminierungssensibler Umgang mit Metadaten mit Blick auf limitierte Ressourcen?
 - * Umgang mit KI im Kontext der Metadatenerstellung und -verarbeitung
 - * DMPs
 - * Wie können Arbeitsprozesse diskriminierungssensibel gestaltet werden – von der Sensibilisierung, Schulung und Diversifizierung von Teams, der Entwicklung von Tools, ...
 - * Wie kann diskriminierungssensible Arbeit in Forschung, Sammlungen und Archiven vor politischer Einflussnahme geschützt und verteidigt werden?
 - * Was muss bei der Veröffentlichung und Vermittlung beachtet werden um diskriminierender Strukturen nicht zu reproduzieren?
 - * Kontrollverlust der Darstellung beim Publizieren der Daten auf externen Plattformen und Repositorien
 - * Wie kann diskriminierendes Vokabular überwunden werden ohne historische Quellen zu verfälschen und auch künftig historische Forschung nach bestimmten Begriffen zu ermöglichen?
 - * Umgang mit Lücken, Unsicherheiten und Widersprüchen in Daten, insbesondere bei Massenerschliessungen, wo es auch darum gehen muss, systematische Fehler zu limitieren.
 - * ...
- **Aufbau, Fokus und Verwendungshinweise:** Was das Handbuch *ist* – und was es *nicht ist*. Wie verwende ich das Handbuch? Was kann ich erwarten? Ist dieses Handbuch für mich?
 - **Ziel:** Ziel muss es sein, diskriminierende Strukturen nicht nur aufzeigen, sondern diese zu überwinden. Dazu braucht es einen Austausch über Lösungsansätze genauso wie über bestehende Hürden und Unsicherheiten. Hierzu möchte das vorliegende Handbuch einen Beitrag leisten. Dazu sammelt es Erfahrungswissen zum diskriminierungssensiblen Umgang mit Metadaten in Forschungskontexten und gibt diese weiter. Um möglichst vielen Menschen zugänglich zu sein, werden in einem ersten Teil wichtige Schlüsselbegriffe eingeführt. Ein zweiter Teil versammelt schließlich Praxiswissen aus verschiedenen thematischen, strukturellen und institutionellen Kontexten.
 - **Metadaten-Lebenszyklus:** Erstellung / Verwendung / Vermittlung usw.
 - **Diskriminierungs-Kontexte:** Essenzialisierung, Bias, Lücken und Unsichtbarmachung, Fehlende Perspektiven durch Unwissen/Ignorance, ethische Konflikte bei

sensiblen Daten, institutionalisierte Strukturen im Arbeitsalltag,...

- **Umgang:** Wünschenswert sind auch Beiträge, wie diskriminierende Metadaten nicht nur erkannt, sondern auch angesprochen und beseitigt/vermieden/überwunden werden können – nicht zuletzt im Kontext von Förderrichtlinien oder Teaminternen Spannungen.
- **Spezifität und Pragmatismus:** Die im Handbuch versammelten Empfehlungen müssen freilich auf den Kontext spezifischer Institutions-, Sammlungs- und Forschungskontexte angepasst werden. Durch einen modularen Aufbau mit verschiedenen Perspektiven und Beispielen versucht das Handbuch den heterogenen Ansprüchen und Verwendungskontexten Rechnung zu tragen.
- **Boxen mit Erfahrungen und Fallbeispielen:** Der modulare Aufbau soll es ermöglichen, verschiedene Stimmen aus dem Forschungs- und Sammlungskontext, sowie aus Public History Projekten zu versammeln und Erfahrungen und Lösungsansätze zu teilen.
- **Living Document:** Das Handbuch ist als kostenloses und öffentlich zugängliches Living Document konzipiert und soll von der Community auf dem öffentlichen Code-Repository weiterentwickelt werden.
- **Code of Conduct:** [Contributor Covenant](#) [Code of Conduct](#)

Dank: - Ggf. hier Personen verdanken, die Inputs und Beispiele geliefert und gegengelesen haben.

2 Schlüsselbegriffe

2.1 Diskriminierung in Daten und Datensätzen

2.2 Forschungsdaten

2.3 Metadaten

2.3.1 Was sind Metadaten?

- Beispiele und Abgrenzung zu *Daten* und *Paradaten*.
- Was ist ein Metadatenobjekt?
- Was ist eine zugeordnete Ressource?
- Können verschiedene zueinander im Verhältnis stehen – in beiden Fällen kann es zu Diskriminierung/Bias kommen.

2.3.2 Verhältnis von Metadatenobjekten und zugeordneten Ressourcen

2.3.2.1 Ein Metadatenobjekt mit einer zugeordneten Ressource

2.3.2.2 Ein Metadatenobjekt mit drei zugeordneten Ressourcen (Triptychon)

2.3.2.3 Ein Metadatenobjekt mit drei verschiedenen zugeordneten Ressourcen

2.3.2.4 Zwei Metadatenobjekte mit derselben zugeordneten Ressource

2.3.3 Standards

- Festhalten: Es existieren verschiedene Standards... Fokus hier (vorerst) auf Dublin Core

2.3.3.1 Schemas / Strukturstandard

- Dublin Core Metadata Element Set

 Tipp der Stadt.Geschichte.Basel

Verweis auf Anhang mit Tabellen der Metadaten der Stadt.Geschichte.Basel

2.3.3.2 Formatstandard

- XML/RDF

2.3.3.3 Inhaltsstandard

- Welche Felder wie befüllt werden (z. B. Reihenfolge Vor- und Nachnamen)
- RSWK

2.3.3.4 Wertstandard

- Inhalt der Felder (z. B. GND)

2.4 Opendata-Prinzipien

2.4.1 Open Data und Langzeitarchivierung

- Grundsätzliches
- Zenodo und andere Aggregatoren und Publikationsorte

2.4.2 FAIR-Prinzipien

- FAIR open (in gewissen Kontexten kann Einschränkung sinnvoll sein und ist in FAIR eingeschlossen!)
- ...

2.4.3 CARE-Prinzipien

- Die CARE-Prinzipien wurde innerhalb www.rd-alliance.org gegründet.
 - C:
 - A:
 - R:
 - E:
- Stand Umsetzung aktuell
 - Noch relativ neu und kommen deshalb erst nach und nach in Umsetzung.
 - Ihre hohe Anziehungskraft (gerade als Komplement zu den FAIR-Prinzipien) verweist jedoch auf ein grosses Bedürfnis.
- Fokus:
 - CARE-Prinzipien adressieren dezidiert indigene Daten, bietet jedoch keinen allgemeinen Rahmen für ethischen Fragen in Bezug auf Sammlungsdaten.
 - Entsprechend stellt sich die Frage, inwieweit die CARE-Prinzipien “fliessend” auf andere Kontexte ausgeweitet werden können, oder ob die Formulierung separater Prinzipien je nach Kontext sinnvoller wäre.
 - Die CARE-Prinzipien liefern wichtige Impulse für FDM um ethische Fragen im Umgang mit Forschungsdaten und -materialien zu reflektieren und systematisch zu berücksichtigen.

2.4.3.1 Communities

- Bei der konkreten Umsetzung stellen sich Fragen wie: “Wer sind genau die *Communities*?” oder “Wie genau bemessen wir *collective benefit*?” Je nach Kontext sind diese Fragen schwer zu beantworten.
 - * Kritische Einordnung des oftmals pauschalisierend eingesetzten Community-Begriffs (vgl. z.B. Kaiser et al. 2023)
 - * Verliert eine Gruppe irgendwann die Rechte über ihre Vorfahren? Wer bestimmt, woran festmachen?
 - * ...
- Beispiele nennen für andere Kontexte, e.g. Datenschutzthematik, NS-Register, heterogene Community, ...

- TK-Labels (Fokussieren auf Selbstbestimmung bei der Label-Vergabe im Kontext von Institutions-Gatekeeping)

2.4.3.2 Diskriminierungsfrei vs. Diskriminierungssensibel

- “Diskriminierungsfrei” im alten Titel aufgreifen
 - * Mit dieser Differenzierung soll kein Relativismus betrieben werden!
 - Streben hin zu Diskriminierungsfreiheit: ja, unbedingt; Diskriminierungsfreiheit behaupten: nein
 - * Begründung
 - Diskriminierungsverständnis ist geografisch und communityabhängig heterogen
 - Diskriminierungsverständnis ändert sich – auch in Zukunft
 - Auch “Normdaten” verändern sich (auch die GND)
 - Dekolonialisierung nie abgeschlossen
 - Archiv-Lücken betonen
 - Archiv-Geschichte bleibt
 - Weit mehr als bloss Vokabular

3 Praxis: Diskriminierungssensibler Umgang mit Metadaten

3.1 Erstellung von Metadaten

- Abgrenzen (1) Primärererschließung, (2) Arbeiten mit Subsets bestehender Metadaten -> in der Praxis oft eine Mischung aus beidem.
- Folgende Arbeitsschritte sind dabei stets wichtig: Inhaltsangabe Folgekapitel

3.1.1 Forschungsrahmen abstecken

- Was beschreibe ich?
- Umfang der Metadatenerfassung (Ziel, Produkt, Ressourcenmanagement (Notwendiges vs. nice to have, etc., für wen (Zielgruppen)?)
- Prioritätensetzung und Zeitmanagement
- Erfassen und Identifizieren der Ressource(n)
- Welche Informationen werden benötigt, um die Ressourcen in den Kontext zu setzen

3.1.2 Umgang mit bestehenden Metadaten

- **Situationen:** Datenrecherche mit Archivkatalog/Aggregator/...
- **Woher stammen Metadaten:** Archive, Literatur, ...
- **Herausforderungen:** Heterogenität, Umgang mit bestehenden Diskriminierungen
- **Strategien:** Zusammenführung vs. Anreicherung vs. Ausdünnung vs. komplett neu vs. ...

💡 Erfahrungen von Stadt.Geschichte.Basel

Beispiele

3.1.3 Kontextualisierung der Ressource

- Entstehungskontext der Ressource
 - Sozialer und politischer Kontext
 - Quellenbeschreibung
 - Kontexte der Primär-Autor:innen
- Verwendungskontext der Ressource
 - In welchem Kontext/Umfang/Zustand/Verschlagwortung wird sie heute zur Verfügung gestellt?
 - Kontexte der betreuenden Archivar:innen etc.
 - Interpretation und Rezeption in der Forschung usw. (beeinflusst etwa Auffindbarkeit)
- Praxis: Reflexion und Dokumentation von obigem

💡 Beispiele

Diverse Beispiele

3.1.4 Metadatenfelder festlegen

- Was sind die Optionen?
- Identifizieren relevanter Felder für die Bedürfnisse des Projekts
- Übernehmen bestehender Felder
- Ergänzen bestehender Felder mit anderen Standards und Entscheidungen dokumentieren
- Extrinsisch vs. intrinsisch (vgl. auch Wertstandards)

💡 Dublin Core

...

3.1.5 Wertstandards wählen und verschlagworten

- Was sind die Optionen?
- Wertstandard wählen und mit Ressourcen abgleichen
- Anpassungen abwägen und vornehmen und Entscheidungen dokumentieren
- Mappings: Verknüpfen verschiedener Wertstandards / Äquivalenzklassen
- Extrinsisch vs. intrinsisch (vgl. auch Metadatenfelder)

💡 Limitationen

- Das Verwenden und Ausgeben von Objektbezeichnungen in einer Lokalsprache setzt voraus, dass eine 1:1-Übersetzung existiert, was nicht immer der Fall ist.

i Ausschlüsse durch Mapping

- kumulative Äquivalenz

i Stolpersteine bei Begriffsauswahl

- Sprachspezifische Unterschiede (Rasse/Race)
- Kontext (Zwerg)
- Wiederaneignungen und Selbstbeschreibungen (Crip)
- Stehende Begriffe (e.g. *Lycaon krebsi*)
- ...

💡 Stolpersteine bei der Verschlagwortung

- Umgang mit Darstellungen von Nacktheit. Malega-ze/Objektivierung/Sexualisierung/Prüdeheit/Selbstbestimmung/... (“Bad zu Leuk”, ...)
- Werte-Urteile von aussen ohne Insider-Wissen
- Gut gemeint aber...
 - Neue Diskriminierungen schaffen durch...
 - * Monofokus (Keltin)
 - * Alarmglocken (Nacktheit)

- Übersetzungen (vgl. oben)

💡 ‘Neue’ Konzepte und historische Quellen

- Beim Verschlagworten und Beschreiben historischer Quellen sollten Begriffe stets so akkurat wie möglich gewählt werden.
- Zuweilen wird dieses Argument jedoch verwendet, um bspw. die Auszeichnung potenziell nichtbinärer Menschen als ideologisch oder unwissenschaftlich zu diskreditieren. Nichtbinarität sei ein zeitgenössisches Konzept, heisst es dann, welches in historischen Daten nichts verloren habe.
- Dekonstruktion dieser Behauptung.

💡 Systematische Lücken bei Schlagwortindexen

(ggf. in Kapitel 2-Schlüsselbegriffe verschieben) - Beispiele: Gender-Open-Pendent für Kolonialitäts-Kontext, Gender-Open-Pendent e.g. für Italienisch - Gründe für Fehlen: ... - Bedarf und Chancen: ... - Hürden: ...

💡 Sollen wir eine eigene Ontologie erstellen?

- Erfahrungen der Stadt.Geschichte.Basel zum Einsatz von OpenGender
 - nur OpenGender Schlagwortindex?
 - interne Strategie
- Chancen: ...
- Gefahren: ...
- Notwendigkeit: es fehlen viele Indexe (vgl. Box “systematische Lücken bei Schlagwortindexen”)
- Strategien und Alternativen
 - Kombination/Mapping verschiedener Indexe (vgl. nächste Box)
 - Subset erstellen

💡 Best Practices für Mappings

- Tipps, Ressourcen und Aggregatoren
- GND-OpenGender-Mapping

3.1.6 Transparente Metadaten: Versionierung und Dokumentation

- Relevanz
 - Wer hat wann und wieso welche Metadaten erstellt/geändert?
 - *“Wie dokumentieren wir, wie wir selber erschliessen? (In 50 Jahren wird auch das wieder interessieren und in Frage gestellt werden.)”* Tobias Wildi (Rückmeldung per LinkedIn)
 - *“Veraltete Metadaten sollen nicht einfach gelöscht, sondern durch neue aktuelle Beschreibungen überlagert werden. Es muss möglich sein, die Entwicklung verschiedener Generationen von Metadaten über die Zeit nachzuverfolgen, denn an ihnen manifestiert sich, wie wir Dinge benennen und über sie denken. Und die Welt steht wie gesagt nicht still, die heute angefertigten Metadaten werden nicht die letzten sein. Zukünftige Generationen werden ihrerseits aus unseren Metadaten auf unsere gesellschaftlichen Verhältnisse und Wertvorstellen rückschliessen. Diese Nachvollziehbarkeit ist aber nur möglich, wenn alte Metadaten nicht gelöscht, sondern versioniert werden und ihr Entstehungskontext dokumentiert wird.”* Tobias Wildi (Rückmeldung per LinkedIn)
 - Historische Nachvollziehbarkeit: *“Unter welchen Umständen und durch welche Personen sind Metadaten mit Begrifflichkeiten und Weltvorstellungen entstanden, von denen wir uns heute abgrenzen?”* Tobias Wildi (Rückmeldung per LinkedIn)
- Umsetzung
 - Technische Integration in bestehende Sammlungsmanagement-Tools
 - Automatisierbarkeit
- Abwägungen
 - Detailgrad
 - Sichtbarkeit nach aussen (siehe unten)

💡 Beispiel für technische/konzeptuelle Umsetzung Versionierung

- ...
- Mindestanforderungen
 - Zugriff via (Volltext)suche (vgl. unten)
 - “This content was machine translated”

💡 Beispiel für technische/konzeptuelle Umsetzung Dokumentation

- ...

💡 Wohin mit der Dokumentation

- Gehört eine Dokumentation in jede einzelne Ressource oder reicht es, eine zentrale Dokumentation für einen ganzen Datensatz anzulegen?
- Wie kann nebst den Daten auch die Dokumentation interoperabel gestaltet werden?

3.1.7 Umgang mit Nicht-Wissen

3.1.7.1 Lücken und Unsicherheiten in Quellen

- Sammlungslücken und fehlende Metadaten/Informationen sind auch eine wichtige Information!
 - Beispiele aus e.g. Militär-Archiven sowie aus Provenienzforschung bringen.
 - Problematik der Unsichtbarmachung im Archiv (hier bloss Verweis auf Schlüsselbegriffe)
- Leerstellen: Zum Umgang beim Leerlassen von Metadatenfeldern:
 - ggf. sollte Absicht hinter dem Leerlassen dokumentiert werden, denn Info existiert nicht Info nicht gefunden Datensatz nicht/nur teilweise bearbeitet.
 - analoges gilt für Tagging von Beständen als un/problematisch.
- Qualifikation der Metadaten: Zum Umgang mit Unsicherheiten:
 - z.B. unsichere oder unleserliche Quelle widersprüchliche, fragwürdige oder umstrittene Quellenlage

3.1.7.2 Expertisen einholen und einbinden

- Expertisen identifizieren.
 - Wer hat Autorität, (mit)zubestimmen? (vgl. *2-Schlüsselbegriffe/Communities*)
- Eigene Positioniertheit reflektieren
- Wechselseitige Machtdynamiken, Interessen, Bedürfnisse und Verletzlichkeiten reflektieren (auf beiden Seiten)
- Potenziale, Formen (auch Grenzen) von Dialog und Mitbestimmung gemeinsam ausloten
- Settings herstellen um Expertise optimal einzuholen
- Umgang mit sich widersprechenden Sichtweisen bzw. Bedürfnissen (ggf. auch innerhalb einer “Community”)
 - Zur Problematik des oftmals pauschalisierend eingesetzten Community-Begriffs, siehe *2-Schlüsselbegriffe/Communities*
 - Interessensabwägung

- Dokumentieren in Freitextfeld
- ...

3.1.8 Umgang mit LLMs

(Nicht im Fokus von Handbuch 2.0 → Verweisen auf Handbuch 3.0)

- Dringlichkeit, dieses Kapitel ins Handbuch aufzunehmen
 - Vorstellung ist verbreitet, AI könne Verschlagwortung übernehmen
 - Sensibilisierung für Biases in Trainingsdaten
 - Metadaten für LLM optimiert publizieren als Teil von FAIR-Data? (machinereadable/LLM-readable data)
 - Umgang mit bereits bestehenden AI-generierten Metadaten
 - * wie erkennen?
 - * worauf achten?
 - * wie markieren/dokumentieren?
- Publizieren von Metadaten
- AI-Tools
 - Diskussion bestehender Tools
 - * DE-BIAS-Tool
 - * ...
 - Chancen von und Anforderungen an AI-Tools
 - * man-in-the-loop
 - Grenzen von AI-Tools
 - * Umgang mit Halluzinationen.
 - * Autor:innenschaft / Haftungsprobleme mit vollautomatisch eingesetzter KI

Positive Erfahrungen aus der Praxis

- Beispiele sammeln

Negative Erfahrungen aus der Praxis

- Beispiele sammeln

3.2 Zugang und Publikation

- Metadaten machen Anschein, alles sei suchbar und da. Doch nein!

- Bei Publikation von Metadaten muss stets auch Rezeptionsebene mitgedacht werden – ganz besonders z.B. auf Sammlungsportalen.
- Im Kontext sensibler Daten geht es dabei auch um Fragen von Zugang
- Und im Kontext von Open Data und Aggregatoren auch um die Frage, inwiefern eine diskriminierungssensible Weitergabe und Präsentation von Daten gefördert werden kann.

3.2.1 Rahmen abstecken

- Metadaten als Findhilfe? Oder mehr?

3.2.2 Umgang mit sensiblen und/oder diskriminierenden Inhalten

- **Was:** Sensible vs. diskriminierende Daten
- **Wirkung:**
 - Trigger vermeiden vs.
 - Schutz vulnerabler Einzelpersonen/Menschengruppen/Kulturgut vs.
 - Reproduktion diskriminierender Narrative
- **Ebenen:** Rechtslage vs. Ethik (es geht um Sicherheit, aber auch um Würde)
- **Abwägung:** Potenzieller Interessenskonflikt zwischen (1) Zugang erleichtern und (2) Schaden vermeiden (nicht immer einfach)

3.2.2.1 Identifizieren sensibler Daten

- Beispiele
 - **Geodaten:** So kann es etwa Sinn machen, Fundorte archäologischer Objekte nicht öffentlich zu machen, um Risiko von Plünderungen zu reduzieren
 - **Personendaten im Archiv:** sensible Informationen lebender/toter Personen Objekte (z.B. Aufenthaltsorte/Interviewtranskripte mit Nachfahren, Zeitzeug:innen, ...)
 - **Personendaten der Archivar:innen:** Transparenz liefert wohlverdiente Sichtbarkeit für Sachbearbeiter:innen, exponiert sie aber gleichzeitig auch. Ggf. Kürzel statt Klarnamen verwenden
- Rechtliche/Ethische Grundsätze
 - ...

3.2.2.2 Identifizieren diskriminierender Daten

- Beispiele
 - **Diskriminierende Beschreibungen:** ...
 - **Zuschreibungen und Stereotypisierungen:** ...
- Rechtliche/Ethische Grundsätze
 - ...

3.2.2.3 Strategien zum Zugangs-Management

- Grundsätzlich: Was zeigen wir wem wie weshalb und unter welchen Bedingungen?
 - Betrifft ein Datenportal genauso wie dieses Handbuch
 - **Wer:** Unterschiedliche Zugänge für verschiedene Zielpublika (Menschen mit spezifischer Forschungsfrage, Online-Datenbank, Dialog mit “Herkunftsgesellschaften”, Web-Crawlers, ...)
 - **Worauf:** Vollzugriff vs. Kontextualisierung vs. Teilzugriff vs. Versionsgeschichte
 - **Kontextualisierung:** Unter welcher Frage/Suchbegriffen soll ein bestimmter Inhalt (z.B. ein Bild) gefunden werden? (z.B. “Antisemitismus” oder “jüdische Menschen”?)
- **Umgang mit Bildern**
 - Relevant für Portale, Ausstellungen, Slides auf Vorträgen
 - Entsprechende Überlegungen sollten auch an einer Fachtagungen gemacht werden (z.B. beim zeigen rassistischer Praktiken auf Bildern)
- **Umgang mit Text**
 - Strategien: Verfremdung, Verzerrung, Durchstreichung, Sternchenbarriere (N***), Änderung, Anführungszeichen
 - Diskussion: Rücksicht mit historischem Kontext
 - Diskussion zu Titeln
- **Umgang mit Biased Perspektiven**
 - Reproduktion von Blickregimen (z.B. in der [Deutschen Fotothek](#))
 - Auswahl von Touren / Eingangsnarrativen / Kontextualisierungen (und Kontextualisierungsprinzipien!!) / bewussten Umwegen / Bildminiaturen
- **Umgang mit Lücken**
 - Explizieren, dass bspw. Namen von Personen unbekannt (bzw. nicht archiviert/notiert/als erwähnenswert befunden worden)
- **Zugangsbeschränkung**

- API-Key als (symbolische) Hürde
- Content-Notes/Disclaimer
 - * MARC-Feld 500
 - * Kommentar zu Länge, Detailgrad, Inhalt
 - * User wählen lassen, ob sie spezifische Inhalte sehen wollen (CSS-Media-Feature)
- **Schutz von Datenbanken vor unbefugtem Zugriff**
 - Verantwortung

💡 Erfahrungen

- Umgang mit Zensurvorfällen beim Markieren/Kontextualisieren
 - Beispiel Stadtbücherei Münster: [Link](#)
- GND
 - Umgang mit veralteten/alternativen Begriffen
 - RSWK erwähnt Datenethik nicht
 - Problematik Verwendungshäufigkeit: z.B. “I—”, welches noch häufig verwendet wird?
- ...

💡 Umgang mit Zitaten und Titeln

- Vorstellung, dass Werk/Ressource nur *einen* Titel haben kann, hat grundsätzlich keinen ontologischen Grund. D.h. es ist durchaus möglich, problematische Titel zu ändern!
 - Beispiele
- Sternchenbarrieren (N***) in Titeln
 - möglich solange kommentiert und falls wissenschaftliche Volltextsuche auf Volltext (und ggf. Versionen) zugreifen kann.

3.3 Handlungsfähig bleiben

- *“Es ist halt schlicht unmöglich, komplett diskriminierungsfrei zu sein.”*
- *“Dann können wir ja gar nichts mehr veröffentlichen”*
- Plädoyer für konstruktiven und proaktiven Umgang aus der “Komplett biasfreie Daten sind doch eh nicht möglich”-Blockade raus, ohne Extremszenarien gegeneinander auszuspielen.
- Gleichzeitig Warnung vor “es sich mit Kompromissen zu einfach machen”.

3.3.1 Typische Widerstände und Konflikte

- Es kann herausfordernd und persönlich belastend sein, Widerstand zu leisten und kritische Positionen konstruktiv einzubringen.
 - Dieses Kapitel des Handbuchs dient dem Teilen oft gehörter Widerstände und Konflikte, und Strategien, wie diese aufgelöst werden können.

3.3.1.1 Widerstände im Team/Institution

“Komplett biasfreie Daten sind doch eh nicht möglich”

“2015 waren die Diskurse noch nicht da”

“Ich finde es ja toll, was ihr hier zu Datenethik diskutiert. Aber in der Praxis uns fehlt einfach die Zeit dazu”

“Welche Missbrauchsfälle sind euch bekannt?”

- Falsche Frage, wenn es um Reproduktion von Narrativen geht. Hier geht es nicht um böse Absichten, sondern um schleichende aber machtvolle Prozesse. Entsprechend sollten nicht Einzelfälle gesucht (und nicht gefunden) werden.

“Wir wissen nicht, wen wir fragen sollten. Und die Ursprungsgemeinschaften erreichen wir ja eh nicht.”

“Das ist technisch halt nicht möglich”

“Wir haben jetzt schon zu viele Metadatenfelder.”

“Gender ist ein neues Konzept und gehört nicht in historischen Metadaten”

- vgl. Box oben in “Erstellen von Metadaten”

“Als Museum sind wir der Objektivität verpflichtet. Wir machen keinen politischen Aktivismus”

“Als Doktorand:in finde ich es abschreckend, eine lange Checkliste vorgesetzt zu kriegen.”

- CARE-Prinzipien als anregende Frage- und Reflexionsliste einführen

i Wir sammeln weitere Statements

- bitte melden per Email oder Pullrequest

3.3.1.2 Konflikte auf Förderebene

“Wir hätten Projektfinanzierung nicht erhalten, hätten wir die Bilder nicht vollumfassend offen publiziert.”

“Wir brauchen ein Endprodukt”

“Uns sind die Hände gebunden”

i Wir sammeln weitere Statements

- bitte melden per Email oder Pullrequest

3.3.2 Wie weiter mit abgeschlossenen Projekten?

- Bestehende Plattformen werden (aktuellen) Ansprüchen zu diskriminierungssensiblem Umgang oft nicht gerecht.
- Leider fehlt es oft an Zeit und Finanzierung, um Änderungen vorzunehmen. Oder es können technische, personelle oder institutionspolitische Hürden bestehen.
- Kritische Auseinandersetzung mit bestehenden Projekte kann sehr produktiv sein, da es auf potenzielle Hürden zukünftiger Projekte sensibilisiert.
- Gleichzeitig stellt sich die Frage, wie mit abgeschlossenen Projekten umzugehen ist. Denn wenn datenethische Grundsätze verpflichten zu Aktion zur Reduktion von Schaden.
- Konkret kann dies heissen: ...

4 Literatur

- Australian Research Data Commons (ARDC). 2020. «ARDC Metadata Guide», März. <https://doi.org/10.5281/ZENODO.6459832>.
- Baca, Murtha. 2016. *Introduction to Metadata*. Herausgegeben von Murtha Baca. 3. Aufl. Los Angeles: Getty Publications. <http://www.getty.edu/publications/intrometadata>.
- Carroll, Stephanie Russo, Edit Herczog, Maui Hudson, Keith Russell, und Shelley Stall. 2021. «Operationalizing the CARE and FAIR Principles for Indigenous Data Futures». *Scientific Data* 8 (1): 108. <https://doi.org/10.1038/s41597-021-00892-0>.
- Davis, Edie, und Bahareh Heravi. 2021. «Linked Data and Cultural Heritage: A Systematic Review of Participation, Collaboration, and Motivation». *Journal on Computing and Cultural Heritage* 14 (2): 1–18. <https://doi.org/10.1145/3429458>.
- Doğtaş, Gürsoy, Marc-Paul Ibitz, Fatima Jonitz, Veronika Kocher, Astrid Poyer, und Laurenz Stapf. 2022. «Kritik an rassifizierenden und diskriminierenden Titeln und Metadaten – Praxisorientierte Lösungsansätze». *027.7 Zeitschrift für Bibliothekskultur / Journal for Library Culture* 9 (4). <https://doi.org/10.21428/1bfadeb6.abe15b5e>.

- Freire, Nuno, Pável Calado, und Bruno Martins. 2019. «Availability of Cultural Heritage Structured Metadata in the World Wide Web». In *Connecting the Knowledge Commons — From Projects to Sustainable Infrastructure*, herausgegeben von Leslie Chan und Pierre Mounier, 121–33. OpenEdition Press. <https://doi.org/10.4000/books.oep.9024>.
- Gabay, Simon, Tobias Hodel, Moritz Mähr, Stefan Nellen, Barbara Roth-Lochner, Pascale Sutter, Andrea Voellmin, und Karin von Wartburg. 2023. «Datenstandards Für Die Historische Forschung – Ein White-Paper Der Schweizerischen Gesellschaft Für Geschichte». Herausgegeben von Schweizerische Gesellschaft für Geschichte. *Whitepaper*, November. <https://doi.org/10.5281/ZENODO.10122052>.
- Gruber, Andrea. 2022. «Vom Knüpfen feministischer Begriffsnetze: Ariadnes Faden & geschlechtersensible Normdaten». *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare* 75 (1): 262–88. <https://doi.org/10.31263/voebm.v75i1.7213>.
- Heinrich, Andreas, und Anita Runge. 2018. «GenderOpen: Ein Repositorium für die Geschlechterforschung». <https://doi.org/10.25595/584>.
- Jaffe, Rachel. o. J. «Library Guides: Metadata Creation». Guide. Zugegriffen 5. Mai 2024. <https://guides.library.ucsc.edu/c.php?g=618773&p=4306381>.
- Lampe, Moritz. 2021. *Diskriminierende Begriffe und Wissensordnungen im Bildarchiv*. Berliner handreichungen zur bibliotheks- und informationswissenschaft 481. Berlin: Institut für Bibliotheks- und Informationswissenschaft der Humboldt-Universität zu Berlin. <https://doi.org/10.18452/23766>.
- Sparber, Sandra. 2016. «What’s the frequency, Kenneth? – Eine (queer)feministische Kritik an Sexismen und Rassismen im Schlagwortkatalog». *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare* 69 (2): 236–43. <https://doi.org/10.31263/voebm.v69i2.1629>.
- Staunton, Ciara, Carlos Andrés Barragán, Stefano Canali, Calvin Ho, Sabina Leonelli, Matthew Mayernik, Barbara Prainsack, und Ambroise Wonkham. 2021. «Open Science, Data Sharing and Solidarity: Who Benefits?» *History and Philosophy of the Life Sciences* 43 (4): 115. <https://doi.org/10.1007/s40656-021-00468-6>.
- Team MusIS. o. J. «Regelwerke, Thesauri, Klassifikationen, Systematiken und Begriffslisten». Wiki. BSZ Wiki. Zugegriffen 5. Mai 2024. <https://wiki.bsz-bw.de/display/MUSIS/Regelwerke%2C+Thesauri%2C+Klassifikationen%2C+Systematiken+und+Begriffslisten>.
- The Turing Way Community. 2022. «The Turing Way: A Handbook for Reproducible, Ethical and Collaborative Research». <https://doi.org/10.5281/ZENODO.3233853>.
- Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, u. a. 2016. «The FAIR Guiding Principles for Scientific Data Management and Stewardship». *Scientific Data* 3 (1): 160018. <https://doi.org/10.1038/sdata.2016.18>.
- Zhang, Lei. 2022. «Empowering Linked Data in Cultural Heritage Institutions: A Knowledge Management Perspective». *Data and Information Management* 6 (3): 100013. <https://doi.org/10.1016/j.dim.2022.100013>.

5 Anhang

5.1 Checkliste

- ...

5.2 Nützliche Tools

5.2.1 Handbücher und Leitfäden

- Reader for Black Lives Philadelphia

5.2.2 GND-Viewers

- Kommentar Vorteile und Nachteile (z. B. des GND-Explorers)
- <https://swb.bsz-bw.de/>
- GND-wiki der dnb
- Regelwerk
- Theoriebox

5.2.3 Ontologien und Mappings

- [LIDO](#)
- Iconclass (ruft Community aktiv auf, sich antidiskriminierend zu beteiligen)
- <https://wiki.dnb.de/display/GND/GND-Mappings+zu+externen+Thesauri>
- Link zum Opengender-Mapping (externe Publikation)

5.3 Metadaten aus dem Projekt *Stadt.Geschichte.Basel*

Erfahrungen Stadt.Geschichte.Basel

- Metadatenobjekte (Eltern)
- Zugeordnete Ressourcen (Kinder)
- Relation von Objekt und Media