



Die CARE-Prinzipien im Forschungsdatenmanagement

Überblick und konkrete Anwendung

Dr. Holger Gauza
Universität Tübingen
Zentrum für Datenverarbeitung
Digital Transformation Lab
holger.gauza@uni-tuebingen.de

Veröffentlichung von Forschungsdaten

Fokus auf Publikation von Forschungsdaten am Ende des Lebenszyklus

- Publikation wird gefordert (as open as possible), sofern keine berechtigten Einwände bestehen, u.a. [Horizon Europe](#)
- Quantitative Forschungsdaten
 - Messdaten, Gensequenzen von Pflanzen
 - Gensequenzen von Menschen
- Qualitative Forschungsdaten
 - Politische Einstellung, Interviews, Sozialforschung

Veröffentlichung von Forschungsdaten

- FAIR-Prinzipien
 - Datenzentrisch
 - Open Science und Open Data
 - Datenaustausch
 - Formate, Schemata, Metadaten, Schnittstelle
 - PIDs / DOIs
 - Lizenzen

Lücken hinsichtlich:

- Positive & negative Auswirkung einer Veröffentlichung
- Nutzen und Gefahren für die mit den Daten verbundene Gruppe
- Fairness bei der Datennutzung

Veröffentlichung von Forschungsdaten

Die CARE-Prinzipien füllen diese Lücke durch:

- Fokus auf indigene Personen und Gemeinschaften und Stärkung derer Rechte
- Beteiligung an geschaffenen Werten bzw. Nutzen für die Datengeber
- Maximierung des Nutzens, Minimierung möglichen Schadens

Etwas
überspitzt:

FAIR-Pol

- Immer frei
- Alles
- Jeder

CARE-Pol

- Fairer Anteil
- Reflektierter Zugang
- Abwägung von Auswirkungen



Veröffentlichung von Forschungsdaten

In die Erstellung der CARE-Prinzipien sind mehrere Dokumente und Vereinbarungen eingeflossen, u. a.:

- [United Nations Declaration on the Rights of Indigenous Peoples](#)
- [Principles of Māori Data Sovereignty](#)
- [Indigenous Data Sovereignty Communique](#)

[Quelle](#)



Agenda

- Ausflug in die Forschungsethik als Anker für die CARE-Prinzipien
- Vorstellung der CARE-Prinzipien
- Umsetzung der CARE-Prinzipien anhand von Beispielen
- Beitrag der CARE-Prinzipien für nicht indigenes RDM?

Forschungsethik?

Auswirkungen der Forschung auf Individuen und die Gesellschaft im Spannungsfeld von Forschungsinteressen und Wertesystemen

- Tier- und Menschenversuche
- Stammzellen- und Genforschung
- Militär- und Rüstungsforschung
- Ethikkommissionen und Ethikstatements von Fachgesellschaften
- Dual-Use
 - Kernforschung: Bombe *oder* Energie
 - Sequenzierung: Analyse schlechter Proben *oder* Wiederbelebung der Spanischen Grippe



Forschungsethik?

Beispiel 1: *Y chromosome haplotype reference database* (YHRD)

- Verwendet von ForscherInnen und Strafverfolgungsbehörden
- Betreut von Mitgliedern der Charité
- Fragwürdiger Umgang mit genetischen Daten von Minderheiten
 - Uiguren in China (Xinjiang) und Tibeter
 - Roma in Europa
- Gab es wirklich immer ein informiertes Einverständnis?
- Gefahr von *genetic profiling* u.a. in autoritären Regimen

[Quelle](#)

Forschungsethik?

Beispiel 2: *HeLa Zellen*

- Epithelzellen eines Zervixkarzinoms
- Ohne Einwilligung der Afroamerikanerin Henrietta Lacks 1951 in einem Krankenhaus in Baltimore entnommen
- Potentiell unsterbliche Zelllinie und Grundlage bedeutender Forschung, u.a.:
 - Impfstoff gegen Kinderlähmung
 - Grundlage für vier Nobelpreise

Forschungsethik?

Beispiel 2: *HeLa Zellen*

- Kommerzielle Vermarktung der Zelllinien ohne Beteiligung der Familie Lacks
- 2013 Veröffentlichung des Genoms eines Stammes der Zelllinie, nach Protesten Löschung der Daten. Etablierung einer Governance Struktur für die Familie Lacks

[Quelle](#)

Forschungsethik?

Gemeinsamkeit beider Beispiele:

- Ungleiche Machtverhältnisse
- Rolle des informierten Einverständnisses unklar oder nicht gegeben
- Auswirkungen auf die Betroffenen direkt und deren Familien
- Keine Beteiligung an der Wertgenerierung
- Veröffentlichung von Forschungsdaten

Die CARE-Prinzipien

CARE-Prinzipien 2020 als Ergänzung zu den FAIR-Prinzipien vorgestellt

- Basieren auf der [Declaration on the Rights of Indigenous Peoples](#)
- Entwickelt von der [Global Indigenous Data Alliance \(GIDA\)](#)

Indigene Völker: potentiell marginalisierte Urbevölkerung (vereinfacht)

Indigene Daten: Daten, Informationen und Wissen in jeglichem Format, das indigene Völker, Nationen und Gemeinschaften auf individueller oder kollektiver Ebene betrifft. Das umfasst auch Daten über deren Ressourcen und Umwelt sowie Daten über sie als Einzelne oder als Kollektiv

Die CARE-Prinzipien

CARE-Prinzipien 2020 als Ergänzung zu den FAIR-Prinzipien vorgestellt

- Fokus auf Wohl indigener Gesellschaften, deren informationelle Selbstbestimmung, Machtausgleich sowie fairer Beteiligung an der Wertschöpfung
- Dokumentation der Provenienz, kulturelle Metadaten, Etablierung indigener Governance, Einschränkung der Verfügbarkeit von Daten
- Auszeichnung indigenen Ursprungs
- Aufbau von Infrastruktur und Ausbildung zur Ausübung von Governance



Die CARE-Prinzipien: Collective Benefit



Das Datenökosystem soll so aufgebaut sein, dass indigene Gesellschaften einen Nutzen daraus erzeugen können

C1: For inclusive development and innovation

- Förderung der **(Nach)nutzung von Daten durch indigene Gesellschaften** u.a. durch die Unterstützung lokaler, selbstbestimmter Entwicklungsprozesse durch Regierungen und Einrichtungen

C2: For improved governance and citizen engagement

- Verwendung von Daten als **Entscheidungsgrundlage für und zur Unterstützung idigener Entscheidungsprozesse**. Besseres Verständnis der eigenen, indigenen Gesellschaft

Die CARE-Prinzipien: Collective Benefit

Das Datenökosystem soll so aufgebaut sein, dass indigene Gesellschaften einen Nutzen daraus erzeugen können

C3: For equitable outcomes

- **Gerechte Beteiligung** an geschaffenen Werten auf Basis indigener Daten und Beitrag zu indigenen Bestrebungen und Wohlergehen



Die CARE-Prinzipien: Authority to Control



Kontrolle über Daten indigenen Ursprungs sowie über die Darstellung indigener Gesellschaften, Ressourcen, Territorien, Wissen etc.

A1: Recognizing rights and interests

- Indigene Gesellschaften haben das Recht auf **informierte Einwilligung** zur Verwendung und Sammlung indigener Daten einschließlich der Entwicklung von Policies und Sammelprotokollen

Die CARE-Prinzipien: Authority to Control

Kontrolle über Daten indigenen Ursprungs sowie über die Darstellung indigener Gesellschaften, Ressourcen, Territorien, Wissen etc.

A2: Data for governance

- **Recht auf Zugang** zu indigenen Daten sowie deren Bereitstellung zur Unterstützung indigener Selbstbestimmung

A3: Governance of data

- **Recht auf die Entwicklung eigener Policies** im Umgang mit indigenem Wissen, indigenen Daten und das Recht auf eine active Rolle in der Datastewardship



Die CARE-Prinzipien: Responsibility



Wer mit indigenen Daten arbeitet, muss nachweisen, dass die Nutzung dieser Daten zum Wohle der indigenen Gesellschaft erfolgte

R1: For positive relationships

- Die Nutzung indigener Daten beruht auf **Respekt, Gegenseitigkeit, Vertrauen und gegenseitigem Verständnis**. Der Datennutzer muss sicherstellen, dass die Würde indigener Gesellschaften gewahrt wird

R2: For expanding capability and capacity

- **Stärkung der Datenkompetenz** sowie Aufbau von Infrastruktur zur Sammlung, Erstellung, Anwendung und Verwaltung von Daten durch indigene Gesellschaften

Die CARE-Prinzipien: Responsibility

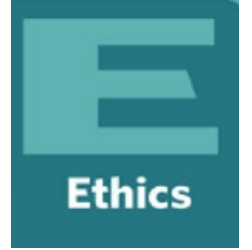
Wer mit indigenen Daten arbeitet, muss nachweisen, dass die Nutzung dieser Daten zum Wohle der indigenen Gesellschaft erfolgte

R3: For Indigenous languages and worldviews

- Bereitstellung von **Ressourcen zur Erstellung von Daten** basierend auf den Sprachen und Weltanschauungen indigener Gesellschaften



Die CARE-Prinzipien: Ethics



Die Rechte und das Wohlbefinden indigener Gesellschaften ist das primäre Interesse über den gesamten Datenlebenszyklus und Datenökosystem hinweg

E1: For minimizing harm and maximizing benefit

- Ethische Daten sind Daten, die indigene Gesellschaften **weder stigmatisieren noch defizitär darstellen** und im Einklang mit gesammelt und verwendet wurden. Die Bewertung dessen obliegt der indigenen Gesellschaft, welche die Daten betrifft

Die CARE-Prinzipien: Ethics

Die Rechte und das Wohlbefinden indigener Gesellschaften ist das primäre Interesse über den gesamten Datenlebenszyklus und Datenökosystem hinweg

E2: For justice

- Ethische Prozesse adressieren **ungleiche Macht- und Ressourcenverhältnisse** und wie diese sich auf Menschenrechte und die Rechte indigener Gesellschaften auswirken. Ethische Prozesse umfassen die Beteiligung indigener Gesellschaften

Die CARE-Prinzipien: Ethics

Die Rechte und das Wohlbefinden indigener Gesellschaften ist das primäre Interesse über den gesamten Datenlebenszyklus und Datenökosystem hinweg

E3: For future use

- Der Umgang mit Daten sollte die **mögliche zukünftige Verwendung sowie zukünftigen Schaden** auf Grundlage eines ethischen Rahmens auf Basis der Werte und Prinzipien der betroffenen indigenen Gesellschaft berücksichtigen. Metadaten sollten die Herkunft und den Zweck sowie Beschränkungen und Verpflichtungen einschließlich Fragen der Zustimmung erfassen



Die Umsetzung der CARE-Prinzipien

Auszeichnung indigenen Ursprungs durch TK Label

- Kontextualisierung von Daten
- Ausschließlich in Zusammenarbeit mit indigenen Gesellschaften
- Information für Dritte über Bedingungen für die weitere Nutzung oder den Zugang

[Biocultural Labels \(BC\)](#) für den Umgang mit Gensequenzen und Forschung



[Quelle](#)



Die Umsetzung der CARE-Prinzipien



- Korrektur von Fehlern
- Sicherstellung der korrekten Attribuierung



- Nutzung für Bildungszwecke
- Außendarstellung einer Gruppe oder Individuums



[Access image](#)

[Image Metadata](#)

PROTOCOL:
Passamaquoddy Public Access

COMMUNITY:
Passamaquoddy

CATEGORY:
Historic Events

COLLECTIONS:
DAR 1916 Revolutionary War Memorial Dedication

ORIGINAL DATE
DESCRIPTION:
06/14/1916

ORIGINAL DATE:
1916 June 14th

CREATOR:
DAR - Individual Unknown

TRADITIONAL KNOWLEDGE LABELS:
 
TKA TKO

[Quelle](#)



Die Umsetzung der CARE-Prinzipien

Umgesetzt von der Library of Congress für die Sammlung *Ancestral Voices*

- Tonaufnahmen der Passamaquoddy, gesammelt 1890 von Jesse Walter Fewkes
- Digitalisiert zwischen 1983-1988, Kopie an die Passamaquoddy übergeben. 2015 erneute Digitalisierung
- Zusammenarbeit mit den Passamaquoddy, verbessertes Verständnis der eigenen Kultur

[Quelle](#)

The screenshot shows a web browser displaying a digital archive page from the Library of Congress. The page title is "Passamaquoddy story of the Fisher and the Sable, part 4" and it is part of the "Jesse Walter Fewkes collection of Passamaquoddy cylinder recordings SR02". The main content features a large audio player for a digital restoration of the recording, with a progress bar and a "Share" button. Below this, there are two smaller audio player thumbnails for alternative versions: a digital preservation master file from the original cylinder and a track from a digital preservation copy of an AFC preservation tape. The "About this Item" section provides metadata including the title, other titles, contributor names (Jesse Walter Fewkes, Peter Selmore, and the Passamaquoddy Tribe), and the creation/publishing date (1890-1930). A "Traditional Knowledge Labels" section is also present, listing labels such as Attribution - Elihtasik, Outreach - Ekehkimkewey, and Non-Commercial - Ma yut monuwasiw.

[Quelle](#)



Die Umsetzung der CARE-Prinzipien

Mukurto – Das CARE CMS

- Unterstützt indigene Gesellschaften beim Aufbau digitaler Sammlungen bzw. Hubs
- Unterstützt TK-Label
- Unterstützt Zugangsbeschränkungen / Protokolle
- Ermöglicht mehrere Aufzeichnungen an den Objekten



[Quelle](#)



Die Umsetzung der CARE-Prinzipien

Aktiver Aufruf von Institutionen zur Kollaboration mit indigenen Gesellschaften




Schulung und Ausbildung indigener Gesellschaften:

- AI & Algorithmen für Māori
- Datenhoheit ausüben
- Datenkompetenz ausbauen

[Quelle](#)



Open to Collaborate Notice [Quelle](#)

 <p>Getting Started Getting started and additional information about GEOME</p>	 <p>Query Search and download sequence files and associated metadata</p>	 <p>Workbench View and manage project specific data</p>
--	--	---

"The Genomic Observatories Metadatabase (GEOME): A new repository for field and sampling event metadata associated with genetic samples", John Deck, Michelle R. Gaither, Rodney Ewing, Christopher E. Bird, Neil Davies, Christopher Meyer, Cynthia Riginos, Robert J. Toonen, Eric D. Crandall Published: August 3, 2017 <https://doi.org/10.1371/journal.pbio.2002925>



[Quelle](#)



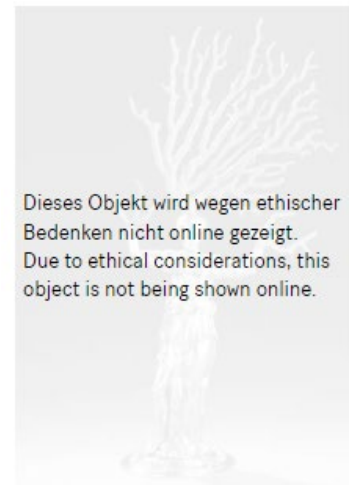
Die Umsetzung der CARE-Prinzipien

Umsetzung durch [Journals](#)

[...] we will reject submitted papers that concern Indigenous communities but do not provide evidence of the care taken towards engagement with Indigenous communities including appropriate attribution, appropriate access, and ideally Indigenous authorship.[...]

Umsetzung durch Kultureinrichtungen

- Staatliche Kunstsammlung
Dresden
- Austausch mit indigener
Gesellschaft und Entfernung
von Material



Dieses Objekt wird wegen ethischer Bedenken nicht online gezeigt.
Due to ethical considerations, this object is not being shown online.

Porträt einer Ehrendame der hawaiianischen Königin Emma Kamehameha

Ambrotypie, koloriert, Blechpassepartout (legiert) mit ovalem Ausschnitt, Preserver (legiert), mit geprägtem Papier i...

GRASSI Museum für Völkerkunde zu Leipzig

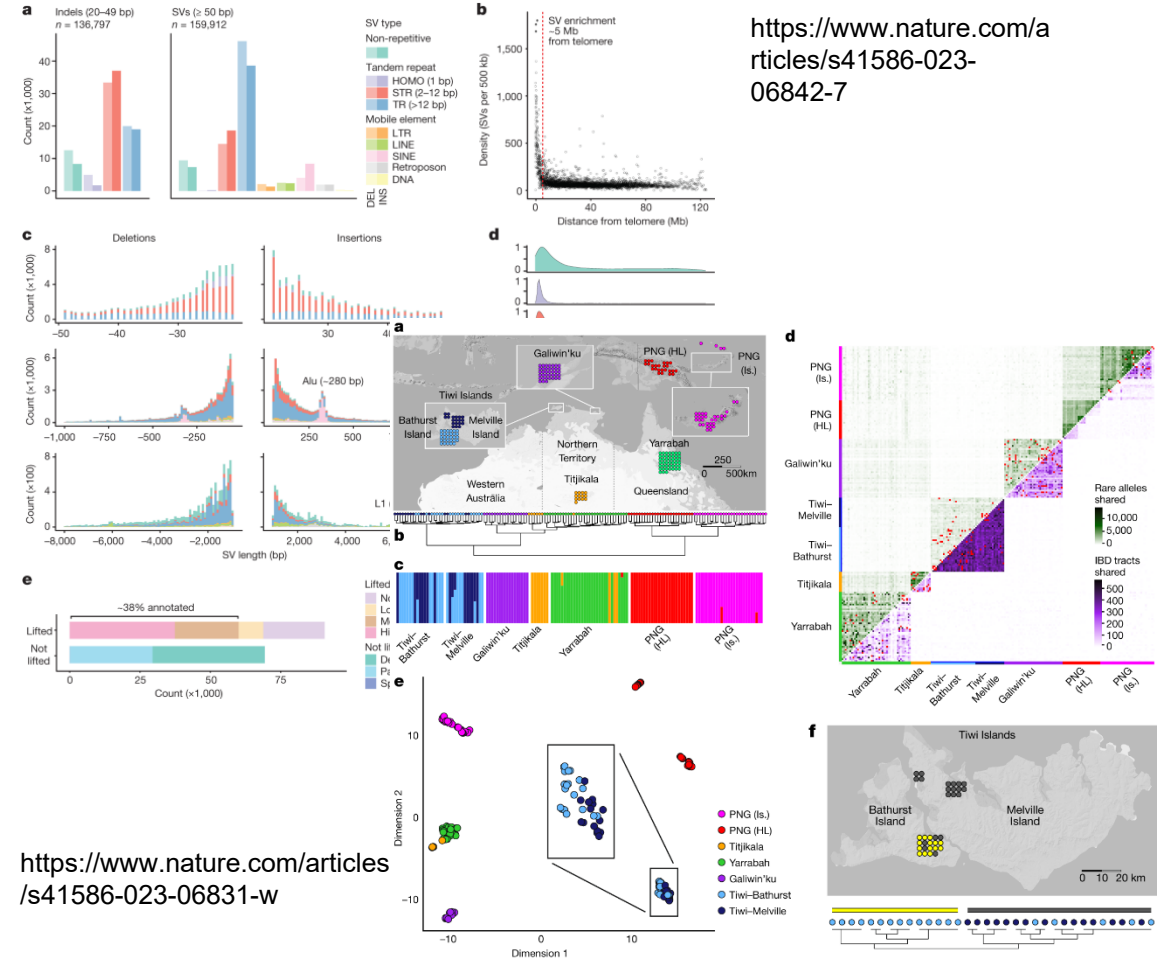
[Quelle](#)



Die Umsetzung der CARE-Prinzipien

Beispiele gegenwärtiger Forschung:

- Genetische Untersuchung indigener Gemeinschaften in Australien -> Referenzgenom
- Vorteile durch Erforschung: Personalisierte Medizin
- Kontrolle über Daten liegt bei National Centre for Indigenous Genomics



<https://www.nature.com/articles/s41586-023-06842-7>

<https://www.nature.com/articles/s41586-023-06831-w>



Die Umsetzung der CARE-Prinzipien

NCIG als Kontrollgremium

- Mehrheitlich außeruniversitär und indigen besetzt
- Entdecken von Behandlungsmöglichkeiten und Präventionsmaßnahmen
- Aufbau guter Beziehungen zu indigenen Gesellschaften durch Information und Beteiligung



Indigenous
leadership



Research that
matters



World-leading
community
consultation

[Quelle](#)

Übertragung der CARE-Prinzipien

Überspitzt: „Ja gut, aber ich habe ja nichts mit indigenen Daten zu tun.“

Fiktives Beispiel: Datenbank mit Interviews von Gewalttätern.

1. Welche Risiken entstehen durch die Veröffentlichung und was kann dagegen unternommen werden?
-> Gezielte Selektion arabischstämmiger Männer und Publikation im politischen Kontext -> E3
2. Welchen Nutzen hat die untersuchte Gruppe von der Publikation?
-> Datenbank als Ursachenforschung und zur Gewaltprävention
-> Einrichtung eines Blogs als Beitrag zur Destigmatisierung -> E1
3. Besondere Sorgfalt bei der Erfassung des Einverständnisses zur Datenverwendung und der informierten Zustimmung -> A1

Letzte Worte

Hier sehe ich die CARE-Prinzipien im DACH-Raum als Impulsgeber für einen bewussten, reflektierten und ethischen Umgang mit Forschungsdaten, siehe auch [hier](#).

Anregung für gegenwärtige Forschung und zukünftige Nutzung, beispielsweise im Bereich KI und Big Data.

Primär Umsetzung der CARE-Prinzipien auf organisatorischer Ebene und mittels Metadaten.

CARE-Prinzipien werden auch im [NFDI4Culture](#) aufgegriffen, siehe auch [hier](#).

Letzte Worte

Die CARE-Prinzipien stehen im Kontext weiterer Initiativen, die auf einen fairen Ausgleich hinarbeiten:

Die Regenwaldapotheke - [Das Nagoya-Protokoll](#)

- Regelt den Zugang zu genetischen Ressourcen
- Regelt die Ausgewogene Nutzung und Aufteilung von Vorteilen
- Völkerrechtliches Übereinkommen

Letzte Worte

Vielen Dank für die Aufmerksamkeit!