

# Digitale Langzeitarchivierung

## Die Nutzbarkeit von Forschungsdaten langfristig erhalten

Dr. Astrid Recker  
GESIS – Datenarchiv für Sozialwissenschaften  
[astrid.recker@gesis.org](mailto:astrid.recker@gesis.org)

Helmholtz Open Science Webinare zu Forschungsdaten: Webinar 28 – 16./21.07.15

# Gliederung

## 1. Einführung

- Herausforderungen und Probleme bei der langfristigen Nutzung digitaler Objekte
- Ziele und Strategien der digitalen Langzeitarchivierung (dLZA)

## 2. Archivierung digitaler Forschungsdaten: Besondere Herausforderungen

## 3. Motivation und Beispiele

## 4. Ihre Fragen!

# Warum dLZA?

A photograph of a vintage 3.5-inch floppy disk resting on top of a stack of old photo albums. The disk has handwritten text on its label: "To: The future Marina" and "From: The present/past Marina". Below the disk, the words "Millennium Disk" are written. The background shows the edges of several photo albums, one of which has a visible price tag of "\$10".

[e: <https://twitter.com/marinarachael/status/615703954224230400/photo/1>](https://twitter.com/marinarachael/status/615703954224230400/photo/1)

Microsoft Office 2007 rendering of a Microsoft works 4.0 file with a sentence added and different word count.  
Bild von Euan Cochrane (cc-by) <https://www.flickr.com/photos/20440494@N03/8097117620/>

# Warum dLZA?

-61.78	138.10	3	-0.7	104	Oct-1991
-59.63	139.17	3	-0.2	32	Oct-1991
-58.13	139.83	3	1.5	38	Oct-1991
-55.90	140.85	3	1.2	32	Oct-1991
-54.12	141.67	3	1.8	39	Oct-1991
-54.12	141.67	3	1.8	130	Oct-1991
-53.85	140.67	3	3.2	70	Oct-1991
-53.53	141.98	3	1.4	53	Oct-1991
-53.13	142.25	3	2.5	39	Oct-1991
-51.85	142.58	3	3.9	50	Oct-1991
-51.45	142.58	3	4.3	64	Oct-1991
-51.25	142.90	3	7.0	126	Oct-1991
-50.67	143.25	3	5.6	57	Oct-1991

Quelle: Röhleman, Carsten (2006): Compilation of temperature and UK'37 alkenone unsaturation index from filter samples. doi:10.1594/PANGAEA.547958

# Digitale Langzeitarchivierung

- „mehr als nur die dauerhafte Speicherung digitaler Informationen auf einem Datenträger. Vielmehr schließt [sie] die Erhaltung der dauerhaften Verfügbarkeit und damit eine Nachnutzung und Interpretierbarkeit der digitalen Ressourcen mit ein“. (nestor 2010, Kap.1:3)
- Umfasst Strategien und Maßnahmen, um Nutzbarkeit über technischen und kulturellen Wandel hinweg zu erhalten.

# Ziele und Strategien der dLZA

Objektebene	Erhaltungsziel	Risiken	Strategien
Bitstream	Unveränderte Bitsequenz, Substanzerhaltung	Verlust / Verfälschung von Bits durch Manipulation, Unfälle, technische Obsoleszenz	Datensicherung, adäquater Einsatz von Speichermedien
Datei	Lesbarkeit durch Software	Obsolete Formate, fehlende Software	Migration, Emulation
Informations-objekt	Interpretierbarkeit durch den Menschen	Verlust der Verständlichkeit aufgrund fehlender Informationen oder kulturellem Wandel	Anreicherung mit Kontextinformationen (Metadaten, Dokumentation)

# dLZA von Forschungsdaten

## Besondere Herausforderungen

Definitorisch

Was sind überhaupt  
Forschungsdaten?

Rechtlich

Urheberrecht,  
Datenschutz

Semantisch

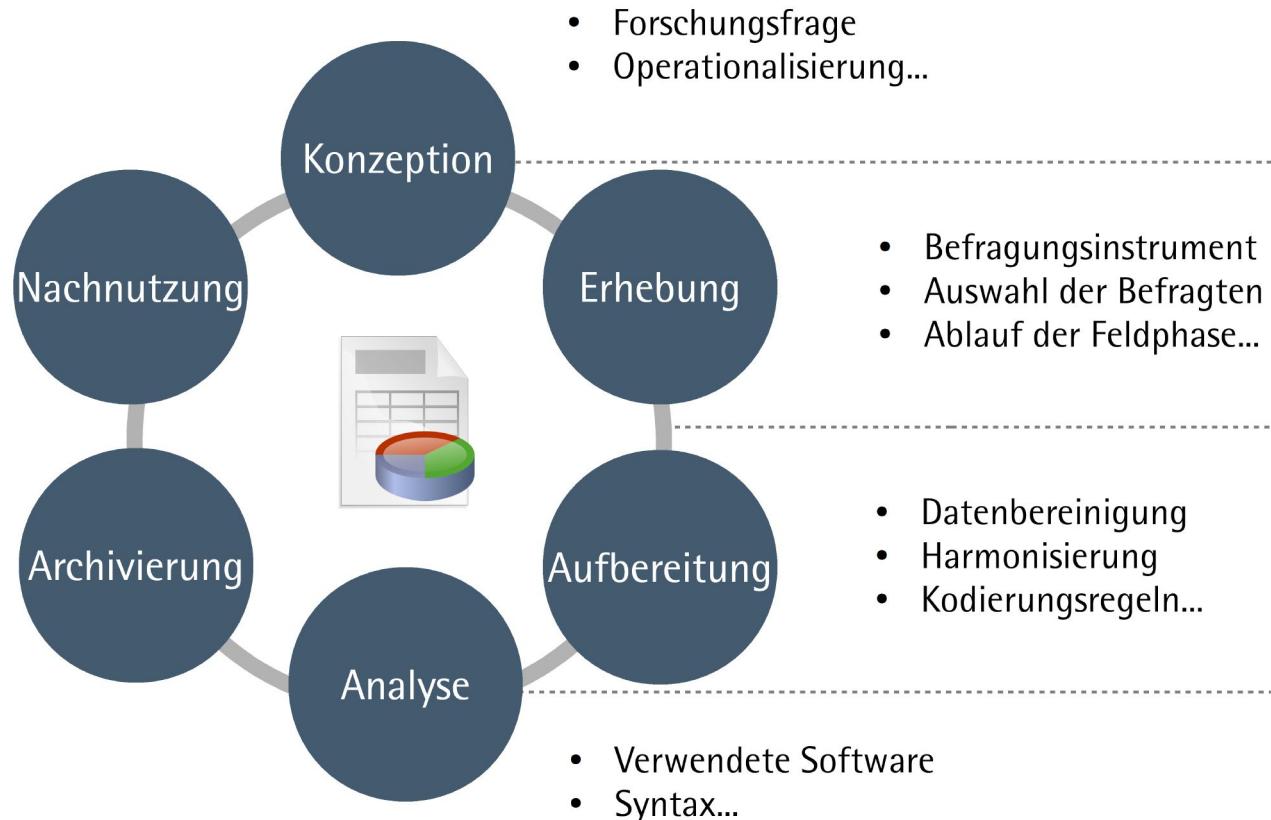
Wichtigkeit von  
Kontextinformation,  
z.B. über den  
Entstehungsprozess,  
für das Verständnis



Bild: ceo

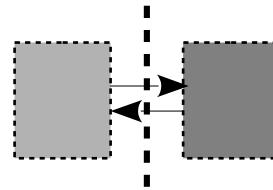
# Kontextinformationen

## Beispiele aus der empirischen Sozialforschung



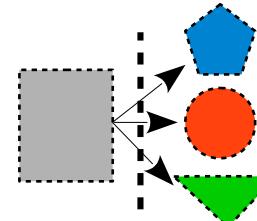
# Gründe für die Datenarchivierung

## Replikation



Forschungsergebnisse  
nachvollziehbar und überprüfbar  
machen

## Nachnutzung



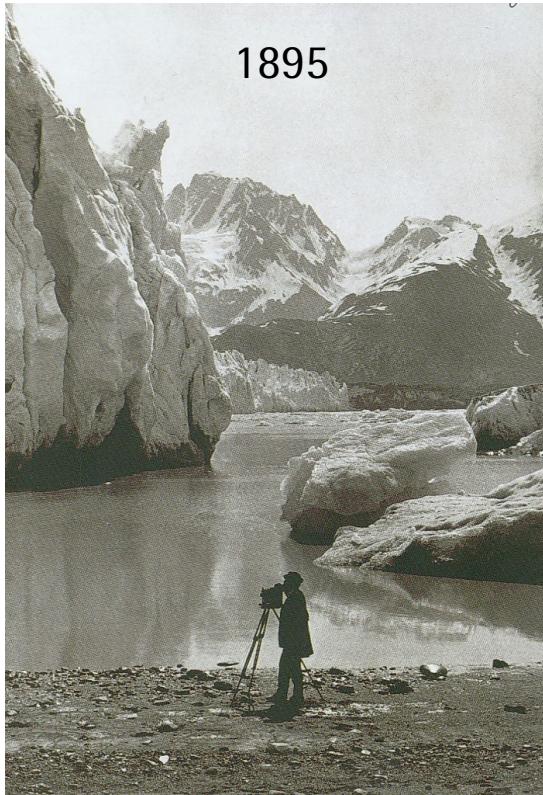
Nutzung von Daten zur  
Beantwortung neuer  
Forschungsfragen ermöglichen

## Förderbedingungen

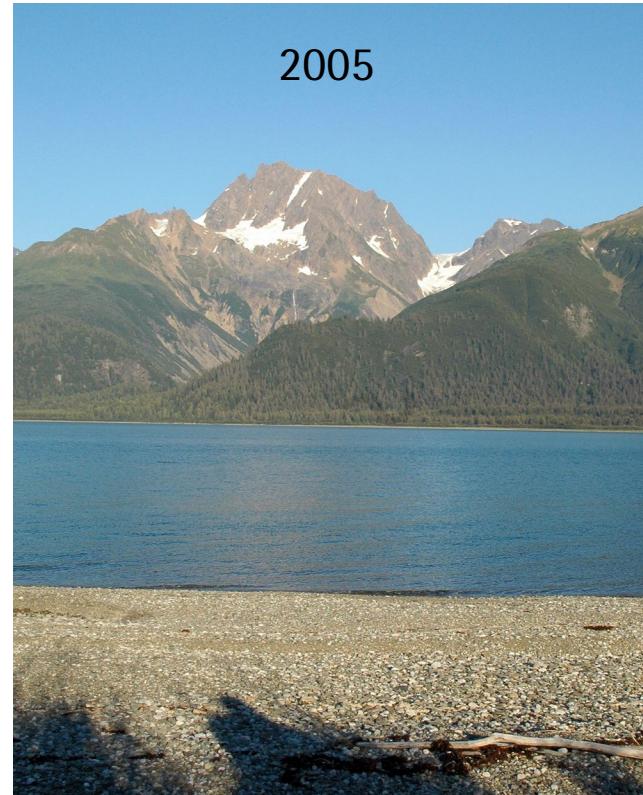
z.B. Open Data Pilot in Horizon 2020

# Beispiel Nachnutzung

## Repeat Photography



1895



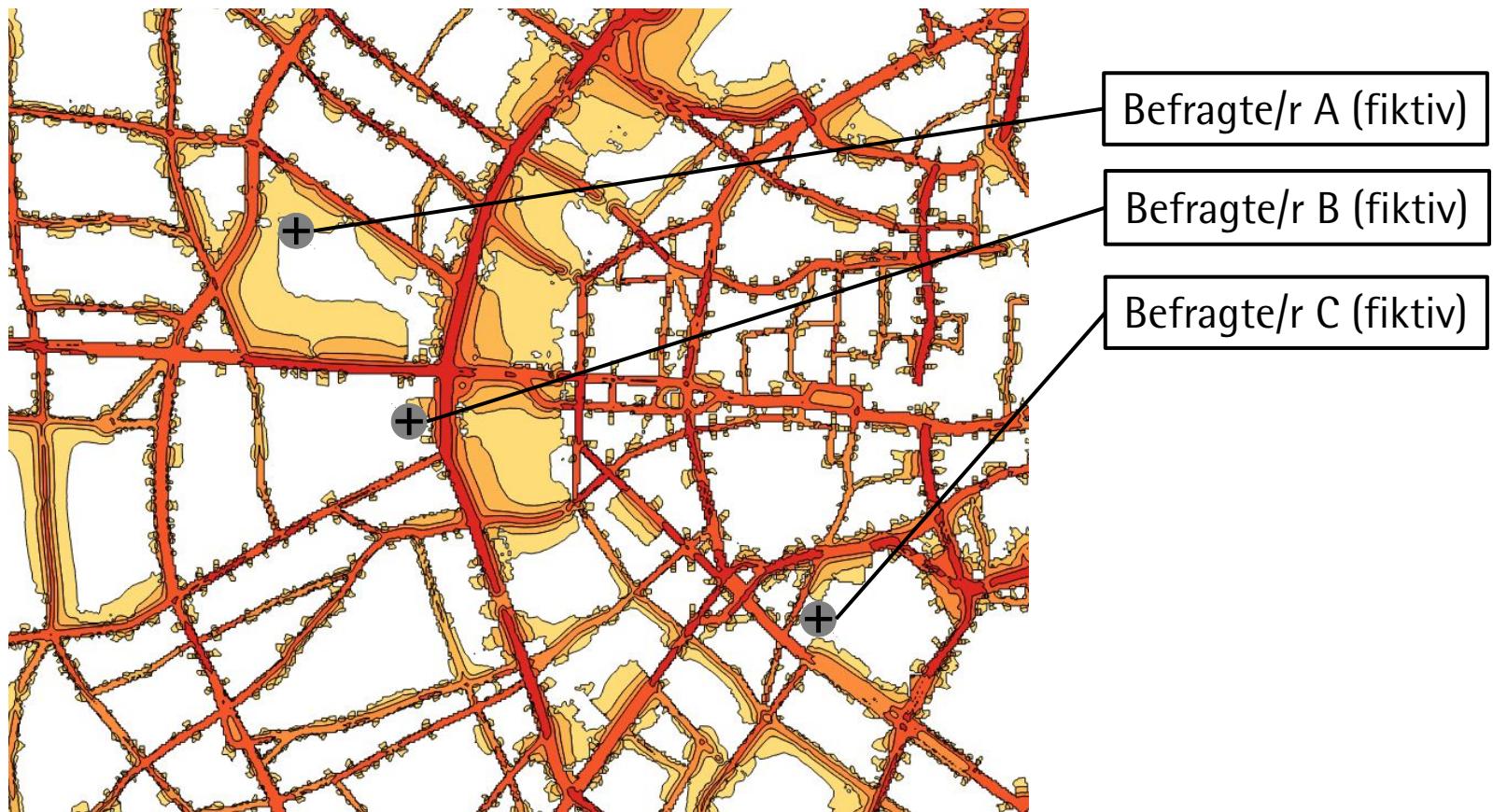
2005

Muir Glacier, Alaska

Quelle: U.S. Geological Survey [http://www.usgs.gov/climate\\_landuse/glaciers/repeat\\_photography.asp](http://www.usgs.gov/climate_landuse/glaciers/repeat_photography.asp)

# Beispiel Nachnutzung

## Verknüpfung von Geodaten und Umfragedaten



# Literatur

nestor (2010). *nestor-Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung.* Version 2.3. Hrsg. H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann, K. Huth.  
<urn:nbn:de:0008-2010071949>

Thibodeau, K. (2002). Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years. In: Council on Library and Information Resources (Hrsg.), *The State of Digital Preservation: An International Perspective* (S. 4-31).  
<http://www.clir.org/pubs/abstract//reports/pub107>