



---

**Planung des Forschungsdaten-Managements:**

# **Speicherung, Erhaltung und Nachnutzung**

---

Torsten Rathmann  
Deutsches Klimarechenzentrum  
rathmann@dkrz.de

6. Dezember 2011, Göttingen

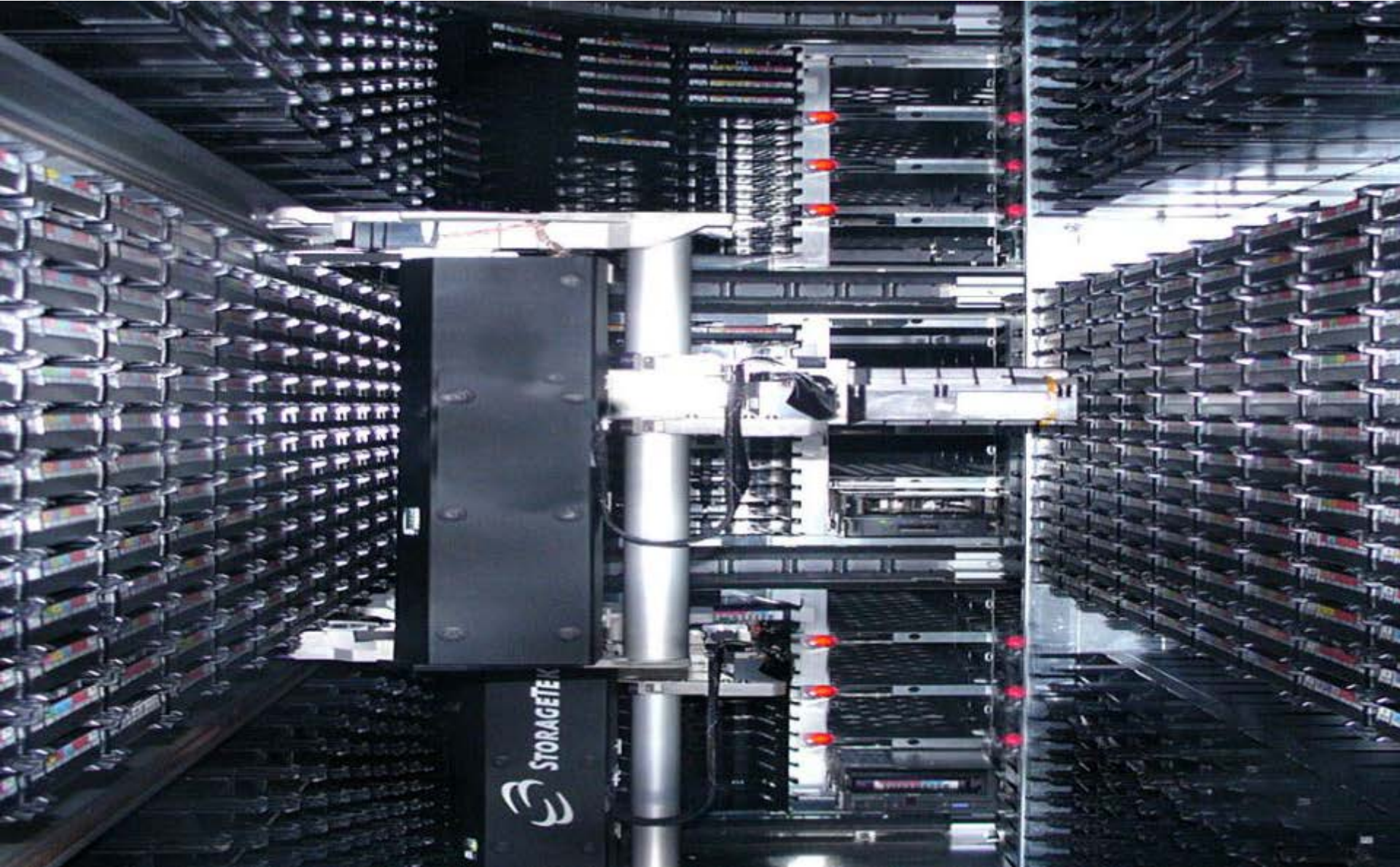
GEFÖRDERT VOM

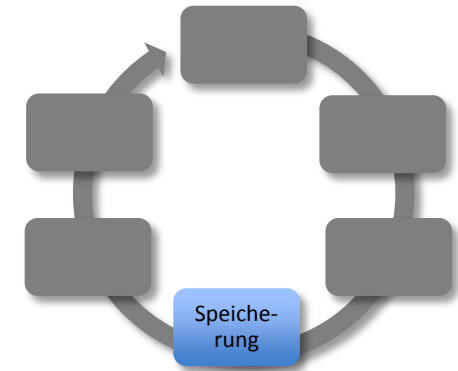


Bundesministerium  
für Bildung  
und Forschung

# Speicherung und Infrastruktur

Das Bild zeigt das Innere eines Storage-Containers und stammt von <http://www.wdc-mare.org/service/>





RAID-Plattenschrank

Mit welchen Technologien wird gespeichert?

Magnetband, Platte oder optisches Medium?

High Performance Storage System (HPSS): Temporär auf Platte, automatische Verlagerung auf Band

Welche Datenmenge? Wie häufig wird zugegriffen?

Werden Sicherungskopien erstellt?

An welchen Orten werden die Daten gespeichert?

- Im selben Gebäude

Brandabschnitt 1

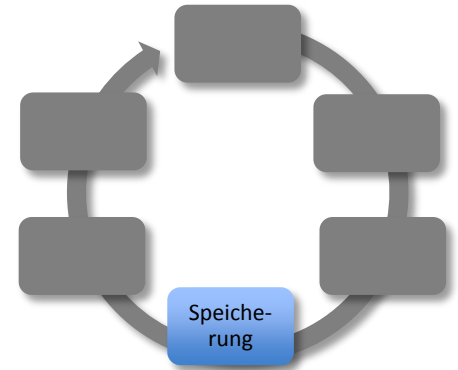


Brandabschnitt 2

- Weit voneinander entfernt

Werden Sicherungskopien überprüft?

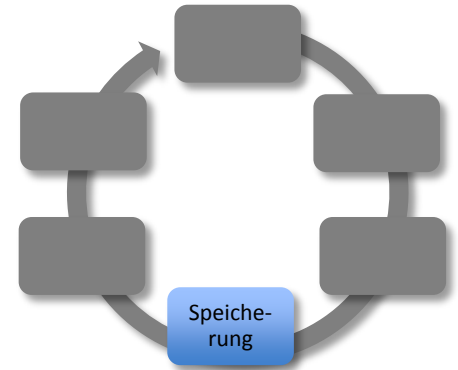
- Prüfsummen (Blockebene bis zu End-to-End)



# Speicherung und Infrastruktur



Datenmanagement-  
Server am DKRZ



Gibt es spezielle Anforderungen, z.B.

- Grafikkarten zur Visualisierung
- Rechenkapazität für Postprocessing?

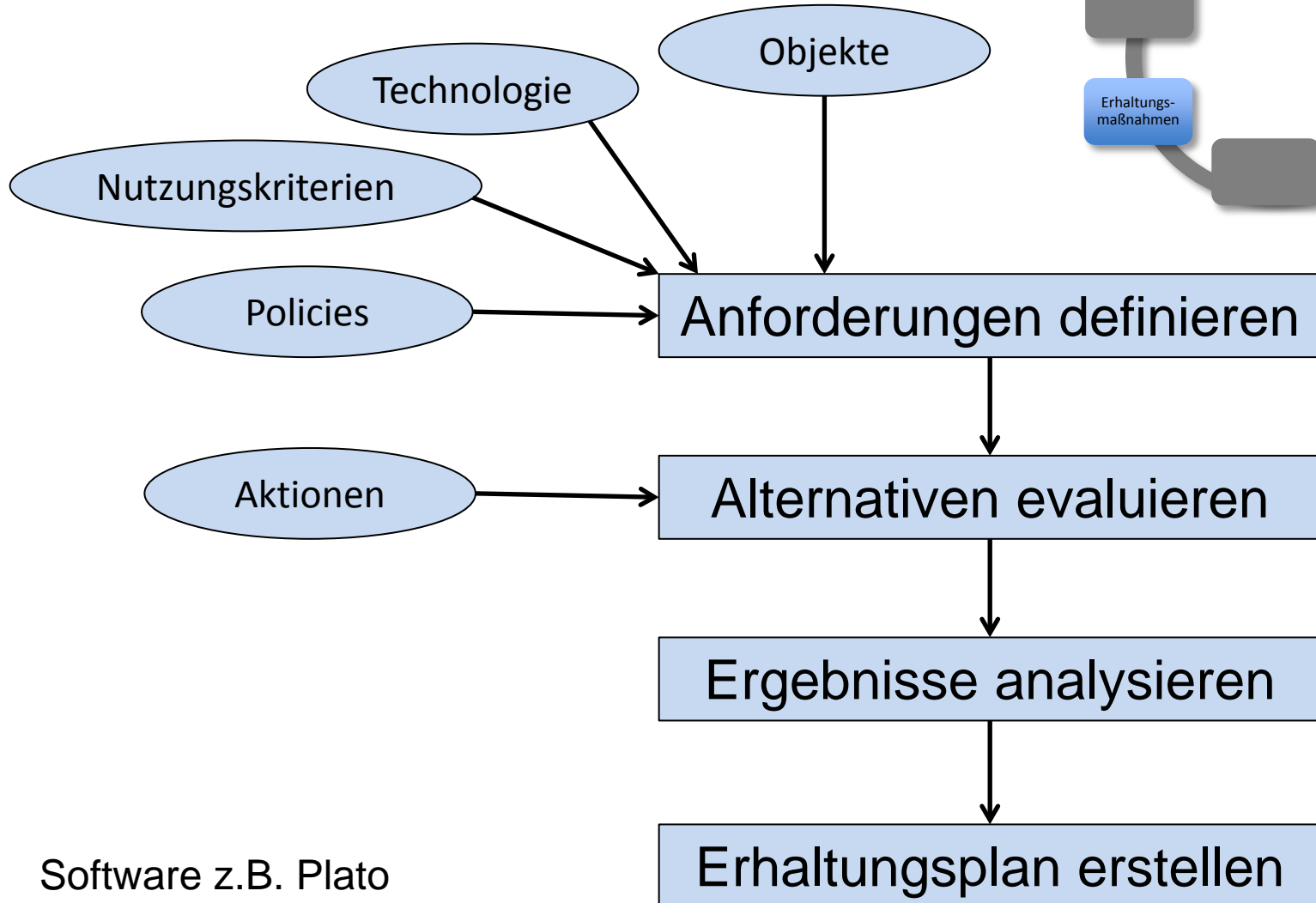
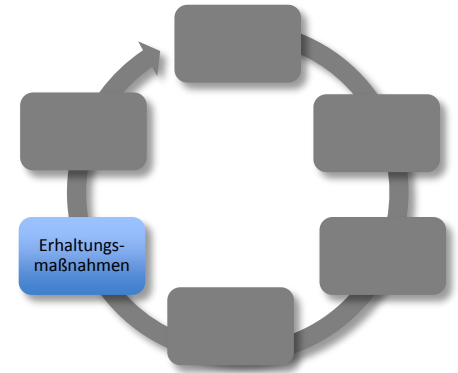
Wer ist für die Speicherung verantwortlich?

- Während des Projekts
- Nach Projektende



Das Bild zeigt einen C64 mit Diskettenlaufwerk und 5¼"-Diskette und ist [www.wikipedia.org](http://www.wikipedia.org) entnommen. Alte Hardware dauerhaft funktionsfähig zu halten ist kaum machbar, weil elektronische Bauteile altern. Beispielsweise können Batterien auslaufen und Kondensatoren platzen.

# Erhaltungsmaßnahmen



Software z.B. Plato



# Erhaltungsmaßnahmen

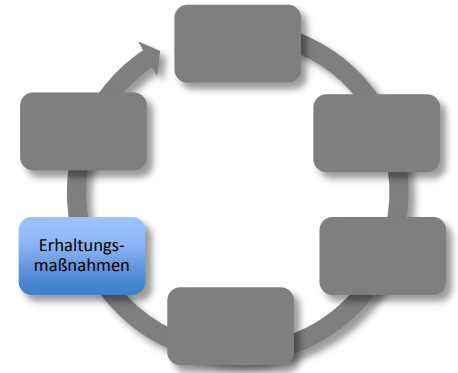
Sind

- Technologien
- Abhängigkeiten
- Nutzungszielgruppe
- Anforderungen an die Nutzung der Daten dokumentiert?

Haben sich Veränderungen ergeben?

Wird Aufbewahrungswürdigkeit neu bewertet?

Gibt es Nachfolgeregelung?







# Zugriff und Nutzung

---

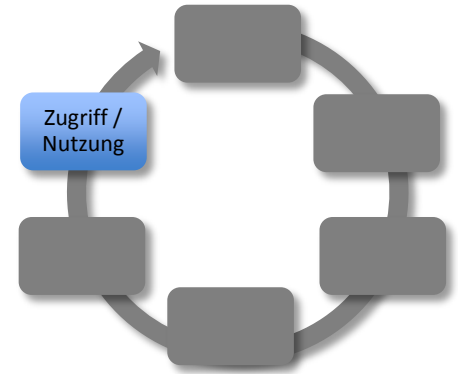
Discover Our Service  
click tabs to navigate



# Nachnutzung und Suchbarkeit

Wie bekomme ich technisch Zugang?

- Web-Portal (HTTP)
- Andere Protokolle: JDBC (Java Data Base Connectivity), ...



Sind die Daten suchbar?

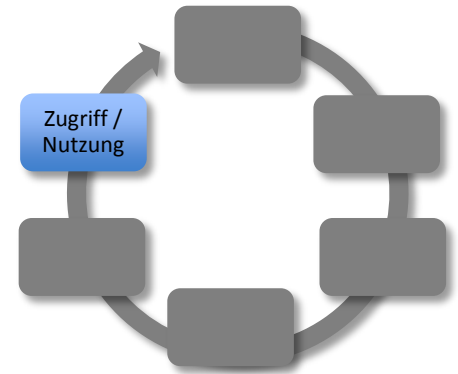
- Suche in Metadaten
- Suche in Daten (setzt schnellen Zugriff voraus)
- Gleichzeitige Suche in mehreren Archiven mit OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting)
- Registries für Suche nach Archiven und Diensten



# Offener Zugang versus Zugriffsbeschränkung

## Offener Zugang

- Erzeugung von Forschungsdaten mit hohen Kosten verbunden
- Mehrfacherzeugung und –bezahlung muss vermieden werden
- Berliner Erklärung: Selbstverpflichtung aller wichtigen deutschen Wissenschaftsorganisationen, den freien Zugang zu unterstützen
- Mittelvergabe immer häufiger an frei zugängliche Forschungsdaten geknüpft

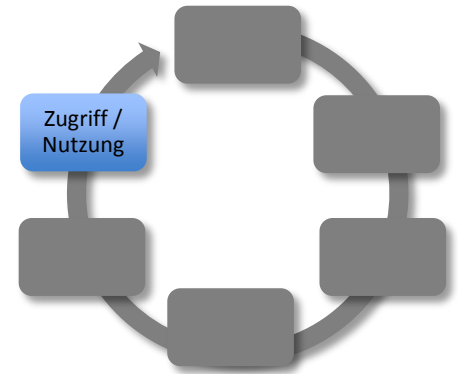




# Offener Zugang versus Zugriffsbeschränkung

Gründe, Daten nicht freizugeben

- Datenschutz
- Geheimhaltungspflicht
  - Bundesstatistikgesetz
  - Betriebsgeheimnisse
- Lizenzbedingungen
- Unfertigkeit (Arbeitskopie)



Gibt es ein Recht auf Erstnutzung? Gibt es Sperrfrist?



# Offener Zugang versus Zugriffsbeschränkung

Wie werden Zugriffsbedingungen durchgesetzt?

- Authentifizierung
  - Benutzername + Passwort
  - Zertifikat
  - OpenID
- Einverständniserklärung dass Nutzungsbedingungen beachtet werden
- Einschränkung des Zugangs auf bestimmte Personen
- Erlaubnis des Eigentümers
- Gesicherter Zugang, über den Daten nur analysiert, nicht aber kopiert werden können

