

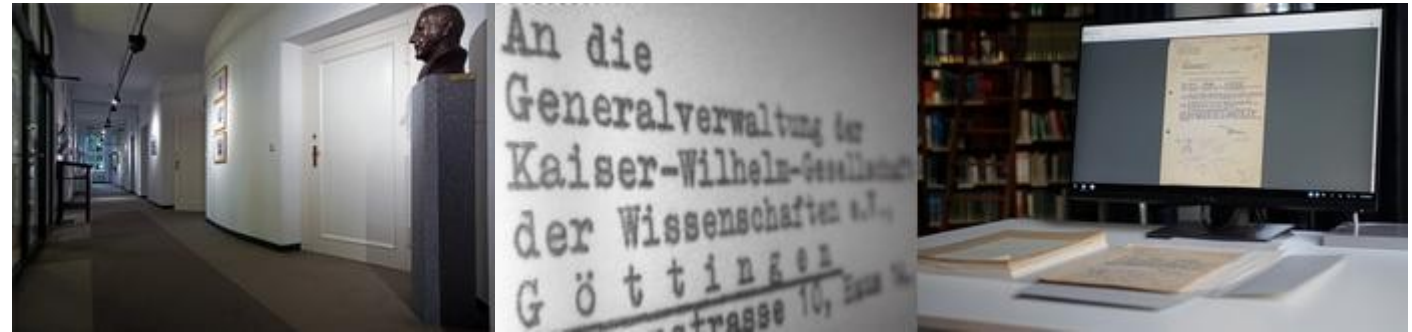


MAX-PLANCK-GESELLSCHAFT

Aspekte der Langzeitarchivierung von Forschungsdaten

Georg Pflanz

Archiv der Max-Planck-Gesellschaft



Quelle: AMPG



- 1975 gegründet
- **Status und Zweck:**
 - „Das Archiv der Max-Planck-Gesellschaft dient als **öffentliches Archiv** der Forschung der Geschichte der Kaiser-Wilhelm-Gesellschaft sowie der Max-Planck-Gesellschaft, ihrer Verwaltung und wissenschaftlichen Arbeit sowie der allgemeinen historischen Forschung und sachlichen Information“ (§1, Archivstatut)
- **Geltungsbereich:**
 - „Dieses Statut gilt grundsätzlich **für die Archivierung von Unterlagen sämtlicher Organe der Max-Planck-Gesellschaft, ihrer Institute und sonstigen Einrichtungen.** Ebenso gilt es für die Archivierung der Unterlagen, die das Archiv in dokumentarischer Ergänzung seiner Bestände übernimmt, betreut und dauerhaft verwahrt.“ (§2, Archivstatut)

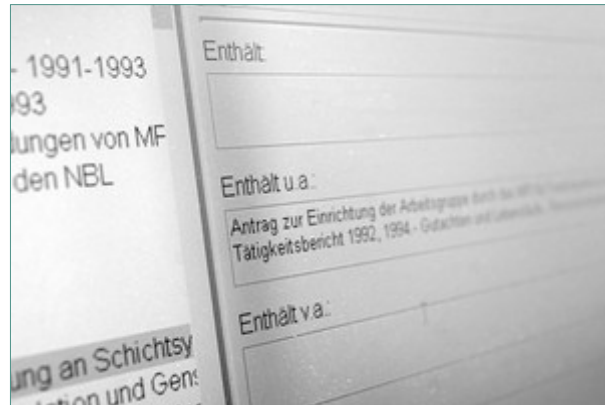


- **Sammlungsschwerpunkte:**
 - Verwaltungsschriftgut der GV und Institute von KWG/MPG
 - ca. 340 Vor- und Nachlässe
 - Fotografien, Ton- und Filmaufnahmen
 - Karten, Urkunden, Pläne, Sammlungen
- → **Überlieferungsziel:**
 - Bildung einer aussagekräftigen Überlieferung zur MPG und seiner Vorgängergesellschaft (Kaiser-Wilhelm-Gesellschaft [KWG]) und deren Zugänglichmachung im Rahmen der rechtlichen Möglichkeiten
- Aktuell, ca. 4.000 lfm Schriftgut
- Seit Frühjahr 2024 Aufbauphase eine digitalen Langzeitarchivs (dLZA)

Was verstehen wir unter „Archivieren“



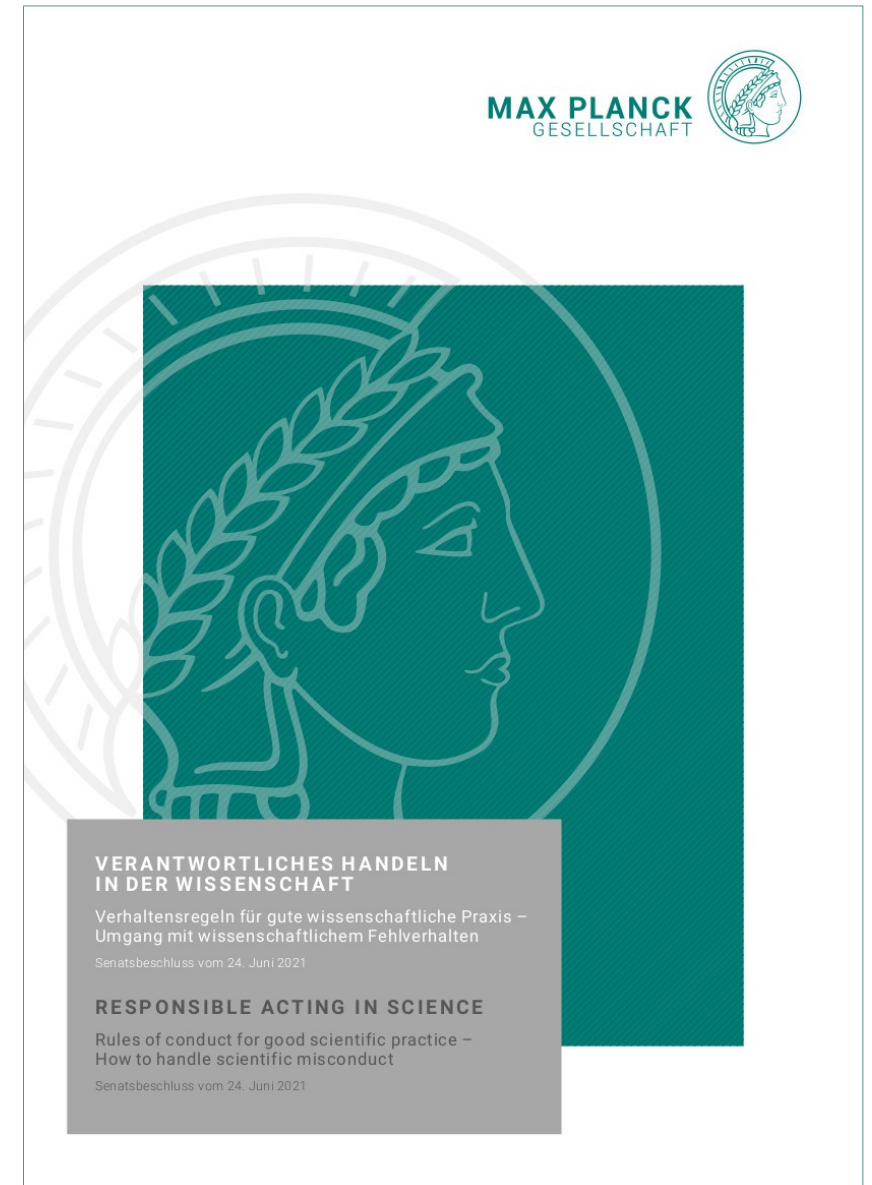
- **Archivierung meint das**
 - **Sammeln** (durch Bewerten und Übernehmen)
 - **Erschließen** (~ Datenbankgestützte erfassen und beschreiben)
 - **Bestandserhaltung**
 - **Zugänglich machen** (im Rahmen der rechtlichen Möglichkeiten)
- von aussagekräftigen Unterlagen und deren **Bewahrung über einen sehr langen Zeitraum**





- **Regulatorische Grundlage: Regeln zum verantwortlichen Handeln in der Wissenschaft**
 - „Die Langzeitarchivierung von Forschungsdaten zielt insbesondere auf die mögliche Nachnutzung der Daten durch unterschiedliche Disziplinen über einen sehr langen Zeitraum, der über grundlegende technologische und soziokulturelle Wandlungsprozesse hinausreicht.

Die Langzeitarchivierung beginnt mit Ablauf der definierten Aufbewahrungsfristen und beschränkt sich auf einen aussagekräftigen Teil der Daten. Die entsprechende Auswahl treffen die Forschungsgruppenleiter*innen zusammen mit der Institutsleitung und dem Archiv der Max-Planck-Gesellschaft.“
 - „Dem Max-Planck-Archiv ist bezüglich der Archivierung von Forschungsdaten sowie den zugrundeliegenden zentralen Materialien und gegebenenfalls der eingesetzten Forschungssoftware in jedem Fall eine koordinierende Rolle zuzuschreiben.“
(<https://www.mpg.de/199493/regelnWissPraxis.pdf>)





- **Forschungsdaten:** „[...] sind (digitale) Daten, die während wissenschaftlicher Tätigkeit (z. B. durch Messungen, Befragungen, Quellenarbeit) entstehen. Sie bilden eine Grundlage wissenschaftlicher Arbeit und dokumentieren deren Ergebnisse.“
(<https://forschungsdaten.info/praxis-kompakt/glossar/#c269824>)
- **potenzielle Nutzergruppen:** historische Geisteswissenschaften, Naturwissenschaften in interdisziplinären Kontexten
- **Archivierung von Forschungsdaten:** Erfassung, Aufbewahrung und Zugänglichmachung von (digitalen) Daten über einen sehr langen Zeitraum in dem sowohl die
 - Authentizität
 - Integrität und
 - Verständlichkeit
 - Zugänglichkeitdieser Informationen sichergestellt werden



- Klärung von Rechtsfragen (Datenschutz, Urheberrecht)
- Unterstützung im Auswahlprozess
- Definition der potenziellen Nutzer (Historiker) und Formulierung ihrer Bedürfnisse
- Informationen über langzeitarchivtaugliche Dateiformate
- Im "Notfall": Archivierung ausgewählter Forschungsdaten für einen sehr langen Zeitraum archivieren (wenn es keine geeignetere Institution gibt)
- Gemeinsam mit den Wissenschaftlerinnen und Wissenschaftlern Kriterien entwickeln als Grundlage für eine Entscheidung über die Langzeitarchivierung von Forschungsdaten; wir haben eine lange Erfahrung im Stellen von Fragen



- **Formelle Aspekte:**

- **rechtliche Klärung** des Zugangs zu den Daten ("Wem gehören die Daten?", "Unter welchen Voraussetzungen dürfen die Daten nach Projektende durch andere Forschende genutzt werden?")
 - → Urheberrecht
 - → Datenschutz
- Verwendung **archivtauglicher Formate** (nicht-proprietär, menschenlesbar, langzeitstabil, Metadaten vollständig enthalten) → KOST (https://kost-ceco.ch/cms/kad_main_de.html)
- **strukturierte Ablage** und **ausreichende Dokumentation**

The screenshot shows the KOST website interface. At the top, it says 'KOST Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen' with language options 'DE | FR | IT' and a search bar. A left sidebar contains navigation buttons: 'Aktuell', 'Standards', 'Dateiformate (KaD)', 'Projekte', 'Publikationen', and 'Tools & Services'. The main content area is titled 'Dateiformate (KaD) Katalog archivischer Dateiformate'. It features an 'Einleitung' section with a paragraph about the catalog describing 52 important and widespread date formats. Below this are several category buttons: 'Text', 'Audio', 'Tabellenkalkulation', 'Hypertext', 'CAD/CAM', and 'Bewertung'. On the right side, there is an 'Empfehlung' section with a paragraph about the detailed analyses and recommendations provided by KOST. Below this are more category buttons: 'Bild', 'Video', 'Datenbanken', 'GIS', 'Weitere Kategorien', and 'Hintergrund'. The 'Hintergrund' section at the bottom right explains that the analysis results are also summarized in a matrix.



- **Inhaltliche Aspekte → welche Inhalte kommen für eine LZA in Frage?**
 - Können die Daten in Zukunft repliziert werden oder wäre dies nur - wenn überhaupt - mit großem Ressourcenaufwand möglich?
 - Ist meine Forschungsfrage oder mein Forschungsansatz bahnbrechend für die Forschung?
 - Benötige ich die Forschungsdaten in der Zukunft für weitere Forschung?
 - Welche weiteren Forschungsfragen wären in Bezug auf die eigene, aber auch andere Disziplinen auf der Basis der erhobenen Daten möglich?
 - Inwieweit würde der Verlust der Daten zukünftige Forschung in diesem oder anderen Forschungsfeldern behindern/verhindern?
 - Sind die Daten nach den FAIR-Prinzipien so aufbereitet und dokumentiert, dass sie von potenziellen gefunden und weiterverwendet werden können?



- Eine mögliche Langzeitarchivierung von Forschungsdaten muss von den Daten-produzierenden Stellen **von Beginn** an mitgedacht werden → **Berücksichtigung im DMP!**
- rechtliche Klärung des Zugangs zu den Daten ("Wem gehören die Daten?", "Unter welchen Voraussetzungen dürfen die Daten nach Projektende durch andere Forschende genutzt werden?")
- Verwendung archivtauglicher Formate (nicht-proprietär, menschenlesbar, langzeitstabil, Metadaten vollständig enthalten)
- Benennung der zu erhaltenden **Inhalte** und **Funktionalitäten**
- strukturierte Ablage und ausreichende Dokumentation
- Benennung eines Ansprechpartners
- Etc.



Vielen Dank für Ihre Aufmerksamkeit!

Kontakt: pflanz@archiv-berlin.mpg.de