

Save the Data

.docx im Jahre 2133

Milagros Wernicke
Staatsarchiv Bern



Archives de l'Etat de Berne, FN Jost N 873

Die Zukunft der Daten in der kantonalen Verwaltung des Kantons Bern

110 Jahre ist die längste Schutzfrist für Unterlagen mit Personendaten. Erst danach (oder mit bewilligten Einsichtsgesuch) stehen sie der Öffentlichkeit zur Einsichtnahme zur Verfügung (Kantonales Gesetz über die Archivierung [ArchG], Art. 18).

Ablieferung im Jahre 2023 Öffentlich zugänglich im Jahre 2133

Dateiformate?

Was?

Aktueller Stand

Viele der in der Kantonsverwaltung erzeugten Daten wurden im .docx-Format erstellt.

.pdf ist der große Konkurrent, allerdings ist der Großteil der Datenproduktion in der Kantonsverwaltung im OOXML-Format (.docx).

Obwohl .docx das am häufigsten verwendete Format ist, **wird es nicht als archivtaugliches Format zugelassen**. Die verschiedenen Word-Versionen, die es gibt, und gewisse Funktionen (z. B. automatisch aktualisierte Felder) können zu veränderbarem Inhalt führen.



Was?

Eines der Prinzipien der Langzeitarchivierung ist die Unveränderbarkeit.

Das unbestrittenste Format für die Archivierung ist .pdf/A. Wenn es sich um .docx handelt, wird **die Konvertierung in .pdf/A** empfohlen.

Herausforderung

Konvertierungen führen immer zu Informationsverlust und müssen möglichst verhindert werden.

Der Inhalt muss vollständig erhalten bleiben (Unveränderbarkeit)
z.B. folgendes können wir nicht gebrauchen

- falsches Seitenlayout
- abgeschnittene Tabellen
- geänderte/fehlende Schriftart
- schwarze Fläche anstelle Grafik

Challenge

Analyse von .docx-Dokumenten

- ❖ Analysieren und beschreiben von kritischen Elementen im Bezug auf Veränderbarkeit in .docx-Dokumenten

Entwickeln einer Software, welche folgendes berücksichtigt:

- ❖ Beschreiben der Versionen und technischen Details jedes .docx (Size, Last Modified, Format-Version, PUID/Identifiers, z.B. Dynamische Felder)
- ❖ Erstellen eines neuen, archivtauglichen Word-Dokumentes

Erwartungen

Was soll am Ende der Hackdays erreicht werden?

Entwicklung einer Anwendung, ähnlich wie PRONOM oder DROID:

- Sie analysiert .docx-Dokumente und beschreibt die kritischen Elemente bezüglich Veränderbarkeit in einer Zusammenfassung.
- Sie erstellt ein neues, archivtaugliches Word-Dokument (z.B. ohne dynamische Felder).



Ressourcen

Dataset mit verschiedenen Dokumenten im Word-Format.

Softwarebeispiele für die Validierung von Archivformaten:

.droid (Software zum Herunterladen)

PRONOM (Applikation Web)

Quelle: The National Archives UK - Preservation digital records

Nach den Hackdays?

Die Entwicklung der Software würde es ermöglichen:

- .docx-Dokumente mit minimalsten Anpassungen als archivtaugliche Formate zu speichern – die Unveränderbarkeit des Dokuments sicherzustellen.
- Bereitstellung als Open-Source-Software für andere Institutionen, die sich der Archivierung widmen.
- Aufbewahrung von .docx-Dokumenten im Staatsarchiv Bern.



Staatsarchiv Bern (StAB)

Kontakt:

Staatsarchiv Bern
Stefan Ryter
Stv. Staatsarchivar

Staatsarchiv
Falkenplatz 4, 3001 Bern
T +41 31 633 51 17
stefan.ryter@be.ch

Staatsarchiv Bern
Milagros Wernicke
e-Archivarin

Staatsarchiv
Falkenplatz 4, 3001 Bern
T +41 31 633 51 01
milagros.wernicke@be.ch