

# Digitale Langzeitarchivierung an der Universitätsbibliothek Bern

Formatpolicy



---

*b*  
**UNIVERSITÄT  
BERN**

Dokument erstellt: Mai 2020

Dokument erstellt von: Marion Prudlo

Version: 1.0

Status: Veröffentlicht; wird mit weiteren Formaten kontinuierlich ausgebaut

Regelmässige Überprüfung: Jährlich

Bei Fragen zu Formaten, die nicht in diesem Dokument aufgeführt sind,  
kontaktieren Sie bitte: [berda@ub.unibe.ch](mailto:berda@ub.unibe.ch)

## **Inhalt**

1.	Einleitung .....	3
2.	Erhaltungsniveaus .....	3
2.1	Niveau 1.....	3
2.2	Niveau 2.....	4
2.3	Niveau 3.....	4
3.	Formate und Erhaltungsniveau .....	5
3.1	Text und Dokument Formate .....	5
3.2	Bildformate .....	6
4.	Quellen.....	6

# 1. Einleitung

Die Policy der Universitätsbibliothek Bern basiert auf nationalen und internationalen Standards. Da es nur wenige Dateiformate gibt, die vollarchivtauglich sind, muss vom Archiv definiert werden, welche Formate in welcher Form unterstützt werden. Formate, die nicht oder nicht mehr archivtauglich sind, werden normalisiert indem sie in ein neues nachhaltiges Format migriert werden. Dies beinhaltet allerdings auch Risiken, da neue Formate nicht generell auch die Funktionalitäten der alten unterstützen. Bei der Migration entsteht eine neue Version des digitalen Objekts, das Original wird beibehalten damit der ursprüngliche Bitstream immer vorhanden ist.

Das Ziel dieser Policy ist, die Anzahl der Formate im Archiv so klein wie möglich zu halten und potentielle Datenlieferant\*innen zu informieren, so dass sie bereits beim Erstellen von digitalen Daten, Entscheide fällen können, die später Auswirkungen auf die Archivierung haben werden.

Die gängigen Formate und das Vorgehen des Archivs bezüglich der Erhaltung dieser Formate werden beschrieben. Die Policy wird regelmässig überprüft und neue Formate werden je nach technologischer Entwicklung und bei Bedarf kontinuierlich aufgenommen.

## 2. Erhaltungsniveaus

### 2.1 Niveau 1<sup>1</sup>

- Bei diesem Niveau handelt es sich um den höchsten Standard. Das Archiv investiert dafür am meisten mit dem Ziel, die Struktur und die Funktionalität des Objekts zu erhalten.
- Dieses Niveau ist nur für bevorzugte Formate möglich, die dokumentiert und weit verbreitet sind, sodass man davon ausgehen kann, dass das Risiko der Obsoleszenz in nächster Zeit gering ist. Eine originalgetreue Reproduktion sollte langfristig möglich sein. Bei einem Medien- oder Formatwechsel ist das in der Regel nicht mehr ohne Veränderungen möglich.

Kriterien für Niveau 1 sind:

- öffentlich dokumentiert
- weit verbreitet und damit lange relevant
- kann von verschiedenen Softwareprodukten dargestellt werden
- nicht-proprietär, lizenzfrei
- geringe Komplexität
- keine Abhängigkeit von anderen Dateiformaten
- verlustfreie Kompression möglich
- ohne Zugriffsschutzmechanismen (Kopierschutz, Verschlüsselung, DRM)

---

<sup>1</sup> Es ist nicht zwingend nötig, dass alle Kriterien erfüllt sind, damit ein Format unter ein bestimmtes Niveau fällt

## 2.2 Niveau 2

- Bei diesem Niveau handelt es sich um einen mittleren Standard. Die Formate, die hierunter fallen haben ein niedriges Risiko in nächster Zeit obsolet zu werden.
- Die Daten werden nach Möglichkeit normalisiert (dafür werden Kopien angelegt, die Originaldaten werden immer behalten).

Kriterien für Niveau 2 sind:

- öffentlich dokumentiert
- weit verbreitet und damit lange relevant
- kann von verschiedenen Softwareprodukten dargestellt werden
- verlustfreie Kompression eventuell nicht möglich
- ohne Zugriffsschutzmechanismen (Kopierschutz, Verschlüsselung, DRM)

## 2.3 Niveau 3

- Bei diesem Niveau handelt es sich um Formate, die nicht für die Langzeitarchivierung geeignet sind. Diese Formate entsprechen nicht den unter Niveau 2 aufgeführten Kriterien und es besteht daher die Gefahr, dass sie in absehbarer Zeit nicht mehr gelesen werden können.
- Für Formate in diesem Niveau kann von BerDA keine Lesbarkeit sichergestellt werden. Die Daten sollten daher vor der Übergabe an das Archiv vom Datenbesitzer/von der Datenbesitzerin in ein geeignetes Format migriert werden. Ist dies nicht der Fall, weil kein geeignetes Format zur Verfügung steht wird der Bitstream archiviert.

Kriterien für Niveau 3 sind:

- es handelt sich um ein Format zu dem es kaum öffentliche Informationen gibt
- wenig verbreitet
- verlustbehaftete Kompression
- wird nur von einer oder wenigen Software Plattformen unterstützt

## 3. Formate und Erhaltungsniveau

### 3.1 Text und Dokument Formate

Format	Niveau	Eigenschaften	Preservation Planning
PDF/A-1	1	<ul style="list-style-type: none"> <li>• Offener Standard</li> <li>• Weltweit unterstützt und weit verbreiteter ISO Standard</li> <li>• Basiert auf PDF Version 1.4</li> <li>• Alle Bilder, Graphiken, Schriften müssen eingebettet sein.</li> <li>• Farben müssen gerätunabhängig sein</li> <li>• Transparente Elemente, JavaScript (ausführbarer Code), Multimedia sind nicht erlaubt</li> <li>• Darf nicht passwortgeschützt oder verschlüsselt sein</li> <li>• Darf keine Links nach aussen enthalten</li> <li>• Darf keine audio, video oder 3D Daten enthalten.</li> <li>• XMP für Metadaten (Autor*in, Thema, Inhalt, Erstellungsdatum, etc.)</li> <li>• PDF/A-1a: Entspricht vollumfänglich dem PDF/A-1b Standard und hat Merkmale, die für die barrierefreie Zugänglichkeit von Inhalten und ihre Darstellung auf Mobilgeräten wichtig sind (sog. «tagged PDF»).</li> <li>• PDF/A-1b: Gewährleistet die langfristige Erhaltung des Erscheinungsbilds eines Dokuments</li> </ul>	<ul style="list-style-type: none"> <li>• Wird als Abgabeformat empfohlen</li> </ul>
PDF/A-2	1	<ul style="list-style-type: none"> <li>• Offener Standard</li> <li>• Weltweit unterstützt und weit verbreiteter ISO Standard</li> <li>• Basiert auf PDF Version 1.7 und baut PDF/A1 weiter aus</li> <li>• Erlaubt eingebettete Dateien sofern sie PDF/A1 oder PDF/A2 entsprechen</li> <li>• Kann nicht ohne Informationsverlust ausgedruckt werden</li> <li>• Erlaubt Transparenz (z.B. für Wasserzeichen), JPEG2000 Kompression</li> <li>• Grösse über 10 GB möglich</li> <li>• PDF/A-2a, PDF/A-2b (wie PDF/A-1)</li> <li>• PDF/A-2u: u steht für Unicode, d.h. die Fonts im Text müssen eine Unicodeentsprechung haben die den Einsatz von Bildschirmlesegeräten erlauben</li> </ul>	<ul style="list-style-type: none"> <li>• Wird als Abgabeformat empfohlen</li> </ul>
PDF 1.3 und niedriger	2		Werden nach PDF/A-1 migriert
PDF 1.4 und höher	2		Werden nach PDF/A-1 migriert

## 3.2 Bildformate

Format	Niveau	Eigenschaften	Preservation Planning
TIFF (Tagged Image File Format)	1	<ul style="list-style-type: none"><li>• Proprietäres Format, das Adobe gehört aber die Spezifikation ist frei erhältlich</li><li>• Weit verbreitet für qualitativ hochwertige Bilder</li><li>• Bisher lizenzfrei</li><li>• Verlustfreie Kompression möglich</li></ul>	<ul style="list-style-type: none"><li>• Wird als Abgabeformat empfohlen, da ein de facto Standard</li></ul>
JPG 2000	1	<ul style="list-style-type: none"><li>• Verlustfreie und verlustbehaftete Kompression möglich</li></ul>	<ul style="list-style-type: none"><li>• Wird als Abgabeformat akzeptiert, wenn kein TIFF vorhanden</li></ul>

## 4. Quellen

KOST, Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen  
[https://kost-ceco.ch/cms/kad\\_pdfa2\\_de.html](https://kost-ceco.ch/cms/kad_pdfa2_de.html) (Stand, Juli 2019)  
<https://kost-ceco.ch/cms/pdf-a-1.html?highlight=pdf> (Stand, Juli 2019)  
[https://kost-ceco.ch/cms/kad\\_tiff\\_de.html](https://kost-ceco.ch/cms/kad_tiff_de.html) (Stand, Juli 2019)

Library of Congress. Sustainability of Digital Formats : Planning for Library of Congress Collections, <https://www.loc.gov/preservation/digital/formats/fdd/fdd000318.shtml> (Stand, May 2019)