

## Empfohlene Dateiformate für OPARU

Um eine langfristige Nutzung der gespeicherten Daten gewährleisten zu können, empfehlen wir für die Veröffentlichung der Daten *weit verbreitete* Dateiformate zu verwenden, die *offenen (normierten) Standards* folgen und *nicht proprietär* sind. Proprietäre Dateiformate sollten, wenn möglich, in entsprechende Standardarchivformate konvertiert werden.

Prinzipiell gilt, dass zur Veröffentlichung Formate gewählt werden sollten, deren Verwendbarkeit auch in Zukunft als wahrscheinlich gilt. Falls Sie in Ihrem Fach mit nicht alltäglichen Formaten arbeiten, haben Sie verschiedene Optionen um eine spätere Nachnutzung zu vereinfachen:

- Die Dateien können zusätzlich in ein offenes Archivformat konvertiert werden. Die Originaldateien sollten dabei jedoch erhalten bleiben.
- Falls dies nicht möglich ist, sollte die mit den Daten veröffentlichte readme-Datei zusätzlich eine Beschreibung zur Nutzung des Formats enthalten.
- Die Dateiergung und das genutzte Format sollten konsistent sein.

**Für Textpublikationen wie Dissertationen, Artikel, sollte stets das PDF-Format (PDF/A-1b nach ISO 19005-1:2005) verwendet werden.**

Falls Sie sich unschlüssig sind, in welchem Format die Forschungsdaten veröffentlicht werden sollen, kann folgende nicht abschließende Liste gängiger Dateiformate als Hilfestellung dienen. *Archivformat* zeigt an, ob das gewünschte Format bestimmten Kriterien zur Archivierung von Dateien entspricht (die Erfüllung ist dabei kein Muss zur Veröffentlichung). *Applikation* zeigt an, ob das Format eine Anwendung zur Ausführung/Nutzung der Daten benötigt:

<i>Dateiart</i>	<i>Formatname</i>	<i>Datei- endung</i>	<i>Archiv- format</i>	<i>Appli- kation</i>	<i>Eigenschaften/Hinweise</i>
Text	Adobe PDF/A-1b	.pdf	X	X	- Standardformat für Textveröffentlichungen - Achten Sie auf eingebettete Fonts und auf eine korrekte visuelle Darstellung des Dokuments nach der Erstellung!
	Plain Text (mit Standardkodierungen wie z.B. ASCII, UTF-8)	.txt, .ans, .asc	X		- Das Textformat kann als stabilstes Format der Informatik gelten
	Hypertext Markup Language (HTML)	.htm, .html	X		
	Extensible Markup Language	.xml, .sgml, .xhtml, .xsl	X		
	Rich Text Format	.rtf			
	LaTeX, TeX	.tex, .latex		X	
	Microsoft Word XML (OOXML)	.docx	X	X	

	Microsoft Powerpoint XML (OOXML)	.pptx	X	X	
	Open Document Text (ODF)	.fodt, .odt	X	X	
	Open Document presentation (ODF)	.fodp, .odp	X	X	
	Postscript	.ps		X	
<i>Tabelle/ Datenbank</i>	Comma-separated values	.csv, .txt	X		- geeignetes Archivformat zur Abbildung von Tabellen
	Tab-separated values	.tsv, .tab	X		- geeignetes Archivformat zur Abbildung von Tabellen
	Microsoft Excel XML (OOXML)	.xlsx		X	- offenes Containerformat für Tabellen - mögliches Archivformat, wenn die Funktionalität der Tabellenkalkulation erhalten werden soll
	Open Document spreadsheet (ODF)	.fods, .ods		X	- offenes Containerformat für Tabellen - mögliches Archivformat, wenn die Funktionalität der Tabellenkalkulation erhalten werden soll
	Data File (Textfiles)	.rdata, .mat, .sdd	X		- Dateien sollten im HDF5-Standard abgespeichert werden
<i>Bild (Rastergrafik)</i>	Tag Image File Format	.tiff, .tif	X		- De-Facto-Standard - Proprietäres Format von Adobe
	Portable Network Graphics	.png	X		- Format für Bilddateien - komprimiert verlustfrei
	JPEG 2000 (verlustfreie Komprimierung)	.jp2, .jpg2	X		- Format für Bilddateien - komprimiert verlustfrei
	JPEG/JPEG2000 (komprimiert)	.jpg, .jp2, .jpg2, .jpxml, .jp3d, .jpf, .jpm, .jpx			- Format für Bilddateien - komprimiert verlustbehaftet
	Graphics Interchange Format	.gif			- Format für Bilddateien - komprimiert verlustfrei
<i>Bild (Vektorgrafik)</i>	Scalable Vector Graphics (SVG)	.svg	X		- Standardformat für Vektorgrafiken
	InDesign	.indd		X	- Format sollte in ein geeignetes Format zur Darstellung der Inhalte formatiert werden, z.B. PDF/A, SVG, o.ä.
	Illustrator	.ai		X	- Format sollte in ein für den Inhalt geeignetes Format zur Darstellung der Inhalte formatiert werden, z.B. PDF/A, SVG, o.ä.
<i>Audio</i>	Waveform Audio File Format (WAVE)	.wav	X		- Standard Containerformat für Audiodateien - sehr speicherintensiv
	MPEG Audio Layer III	.mp3	X		- sehr verbreitet - komprimiert verlustbehaftet

	Free Lossless Audio Codec (FLAC)	.flac			- Offenes Containerformat für Audiodateien - komprimiert verlustfrei
	Audio Interchange File Format (AIFF)	.aiff, .aif, .aifc			- Standard Audioformat für Macintosh - unkomprimiert - hoher Speicherplatzverbrauch
	Advanced Audio Coding (AAC)	.aac, .3gp, .mp4, .m4a, .m4b, .mpg, .mpeg			- komprimiert verlustbehaftet
<i>Video</i>	Motion JPEG2000 (unkomprimiert)	.mj2, .mjp2	X		- komprimiert verlustfrei - Datenmenge kann schnell groß werden
	MP4 (MPEG-4 Part 14, H.264/MPEG-4 AVC)	.m4v, .mp4v, .mp4, .avi, .mov			- Video-Containerformat - komprimiert verlustfrei - geringer Speicherplatzverbrauch
	Motion JPEG2000 (komprimiert)	.mj2, .mjp2			- geringe Speicherdichte - komprimiert verlustbehaftet
<i>Andere</i>	SAS (Statistical Analysis Software)	.sas, .sps		X	
	Java Servlet	.do			X
	SQR Output File	.lis			X
	Compressed Archive File	.zip	X		X
	GZIP Compressed Archive File	.gz			X
	...				

### Nicht gelistete Formate

In OPARU können prinzipiell alle Dateiformate gespeichert werden. Diese werden in der Form erhalten, in der Sie vom Nutzer hochgeladen werden. Um eine Nachnutzung auch in Zukunft möglichst wahrscheinlich zu halten gibt es diverse Aspekte auf die man bei der Wahl des Dateiformats, neben der Beschreibung zur Nachnutzung, achten kann. Bewertungskriterien<sup>1</sup> sind: Offenheit, Verbreitung, Komplexität, Schutzmechanismen, Selbstdokumentation, Robustheit, Abhängigkeiten.

<sup>1</sup> nestor-Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Kapitel 7 Formate. Version 2.3 - 2010. Ed. Karsten Huth. <[urn:nbn:de:0008-2010062462](http://nbn-res.de:urn:nbn:de:0008-2010062462)>.