

Dateiformate

Überlegungen zu Dateiformaten sollten bereits möglichst zu Beginn eines Forschungsvorhabens berücksichtigt werden, da die nachhaltige Aufbewahrung und Nachnutzbarkeit von Datensätzen entscheidend von der Verwendung geeigneter Dateiformate abhängt.

Dateiformate

Dateiformate legen die Syntax (Regeln zur Anordnung von Zeichen) und Semantik (Bedeutung von Zeichen) einer Datei fest. Damit ist die Kenntnis des vorliegenden Dateiformats essentiell, um die in einer Datei gespeicherten Informationen richtig interpretieren zu können.

Dateiformate werden meist für einen bestimmten Zweck entwickelt, zum Beispiel für die Speicherung von Text- oder Bildinformationen oder von ausführbaren Anwendungen.

Dateiformate können urheberrechtlich geschützt ("proprietär") sein und dadurch Abhängigkeiten von bestimmten Programmen bzw. Softwareherstellern befördern. Im Gegensatz dazu werden bei offenen Formaten die Spezifikationen veröffentlicht. Offene Formate können von freier und Open-Source-Software implementiert werden und reduzieren dadurch die Bindung an bestimmte Programme oder Softwarehersteller.

Formatwahl

Mit der Wahl eines geeigneten Dateiformats können Sie die nachhaltige Aufbewahrung und Nachnutzbarkeit Ihrer Datensätze sicherstellen. Idealerweise wählen Sie ein Format, das den Ansprüchen langfristiger Aufbewahrung (z.B. gute Dokumentation, offener Formatstandard) und optimaler Nachnutzbarkeit (z.B. hoher Verbreitungsgrad der erforderlichen Software, gute Editierbarkeit der Datensätze) entspricht.

Generell gilt, dass die Formatierung den nationalen und internationalen Vorgaben der jeweiligen Fachdisziplin entsprechen sollte.

Mit RADAR-BB können Sie Forschungsdaten in beliebigen Dateiformaten archivieren und publizieren. Um möglichst gute Voraussetzungen für eine langfristige Aufbewahrung und Nachnutzbarkeit Ihrer Daten zu gewährleisten, empfehlen wir Ihnen jedoch die Verwendung geeigneter Formate.

Datentyp	Empfohlene Formate	Weitere geeignete Formate	Ungeeignete Formate
Textdokument	XML-basierte Formate wie Microsoft Office XML (.docx) und Open Office XML (.sxw), Open Document Format (.odt) und Strukturierter Text/Markup (.xml, .sgml, .html, .dtd, .xsd und weitere)	Portable Document Format PDF, PDF/A-1, PDF/A-2 (.pdf), Rich Text Format (.rtf) Reiner Text, Plain Text (.txt)	Microsoft Word (.doc)
Tabelle	Comma Separated Values (.csv) sowie XML-basierte Formate wie Microsoft Office XML (.xlsx)	Portable Document Format PDF, PDF/A-1, PDF/A-2 (.pdf)	Microsoft Excel (.xls)
Datenbank	ANSI SQL (.sql) Comma Separated Values (.csv)		
Statistische Angaben	SPSS Portable (.por) SAS transport (.sas) STATA (.dta)		
Rastergrafik	TIFF v. 6, unkomprimiert (.tif, .tiff) GeoTIFF (.geotiff) für georeferenzierte Graphik) Adobe Digital Negative (.dng; für Kamerarohdaten)	Portable Network Graphics (.png) Joint Photographic Expert Group (.jpeg, .jpg) Graphics Interchange Format (.gif) Bit-Mapped Graphics Format (Microsoft) (.bmp) Photoshop (Adobe) (.psd) CorelPaint (.cpt) JPEG2000 (.jp2, .jpx) RAW image format (.nef, .crw und weitere)	
Vektorgrafik	Scalable Vector Graphics (.svg)	Portable Document Format PDF, PDF/A-1, PDF/A-2 (.pdf)	
Video	Lossless AVI (.avi) MPEG-1, MPEG-2 (.mpg, .mpeg) MPEG-4, H264 (.mp4) FLV (.flv)		
Audio	WAVE (.wav) AIFF (.aiff)		
Computer Aided Design (CAD)	AutoCAD DWG (Version 2000), DXF (.dfx, Release 12/14)		

Virtual Reality, 3D	X3D (.x3d; speichert Animationen sowie DG,P,F,I,B,M,V,L,T,G) OBJ (.obj; speichert DG,P,F,I,B,M,G) COLLADA (.dae, XML-basiert; speichert Animationen & DG,P,B-Rep,F,I,B,M,V,L,T,G) PLY (.ply; speichert DG,F,I,B,M)	Virtual Reality Modeling Language (.vrml und .avi, .mpg, .jpeg) Universal 3D Format (.u3d und .avi, .mpg, .jpeg) STL (.stl und .jpeg) DXF (.dfx und .jpeg)	
Nuclear Magnetic Resonance (NMR)		Gerätespezifische Datenformate (Varian/Agilent, Bruker, Jeol)	Oft werden fachspezifische/proprietäre Formate genutzt. Diese sind z.B. durch eine disziplinspezifische Richtlinie/Hersteller vorgegeben. Somit ist Ihre Verwendung im betreffenden Fachbereich notwendig.
Difference Gel Electrophoresis (2D-DIGE)		TIFF (.tif, .tiff) Gel Image (.gel)	Oft werden fachspezifische/proprietäre Formate genutzt. Diese sind z.B. durch eine disziplinspezifische Richtlinie/Hersteller vorgegeben. Somit ist Ihre Verwendung im betreffenden Fachbereich notwendig.

Die Empfehlungen orientieren sich an internationalen Standards wie dem Kriterienkatalog zur besonderen Eignung von **Dateiformaten** zur Langzeitarchivierung der Library of Congress und aktuellen Entwicklungen im Bereich des Forschungsdatenmanagements und werden regelmäßig überprüft und aktualisiert.

Eine weitergehende Übersicht über geeignete Dateiformate im PDF-Format finden Sie z.B. auf den Seiten der Informationsplattform forschungsdaten.info sowie auf der Website des FIZ Karlsruhe unter <https://radar.products.fiz-karlsruhe.de/de/radarabout/dateiformate>.

Die Empfehlung zu Dateiformaten wurde erstellt mit Materialien von:



FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur

FIZ Karlsruhe – Leibniz-Institut für
Informationsinfrastruktur
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen
Tel. +49 7247 808-841

info@radar-service.eu
www.radar-service.eu



Lizenziert unter CC-BY 4.0 | <https://creativecommons.org/licenses/by/4.0/>