

Fun with File Formats in FRDR

@Risk North 3 - November 21, 2024

Amanda Tomé, Preservation Coordinator, Digital Research Alliance of Canada



**Digital Research
Alliance of Canada**

**Alliance de recherche
numérique du Canada**

FRDR
Federated **Research**
Data Repository



DFDR
Dépôt fédéré de
données de recherche

Background

- Federated Research Data Repository (FRDR) is a general purpose, bilingual repository
- Accept datasets from all research domains
- Wide range of file formats
- ~400 file formats unknown or misidentified



File Format Work

- Development of file format signatures
- File format signatures can be external or internal

00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	10	11	12	13	14	15	16	17	18	19	1A	1B	1C	1D	1E	
01	00	01	00	EE	3D	00	00	2F	00	00	00	05	00	00	00	72	61	64	61	72	2E	72	65	76	69	73	69	6F	6E	2E	. = / radar.revision.
6D	61	6A	6F	72	00	01	01	72	61	64	61	72	2E	72	65	76	69	73	69	6F	6E	2E	6D	69	6E	6F	72	00	01	11	major radar.revision.minor
6F	72	69	67	69	6E	2E	63	6F	64	65	00	01	00	6F	72	69	67	69	6E	2E	74	69	6D	65	00	09	4D	6F	6E	20	origin.code origin.time Mon
4A	75	6E	20	20	31	20	30	30	3A	30	31	3A	30	32	20	32	30	32	30	00	6F	72	69	67	69	6E	2E	63	6F	6D	Jun 1 00:01:02 2020 origin.com

Hex editor example for rawacf format



Community



- Submission of format signatures to PRONOM
- Reach wider digital preservation community
- Reach a variety of tools and software
- PRONOM Drop-Ins increase collaboration

Details for: MATLAB Script File

 Save as... [XML](#) | [CSV](#)  Print

Go to: [Summary](#) | [Documentation](#) > | [Signatures](#) > | [Compression](#) > | [Character encoding](#) > | [Rights](#) > | [Reference files](#) > | [Properties](#) >

Summary

Name	MATLAB Script File
Version	
Other names	
Identifiers	PUID: fmt/1678
Family	
Classification	Text (Structured)
Disclosure	
Description	MATLAB (an abbreviation of MATrix LABoratory) software and language enables the user to undertake numerical computation and visualization, programme and analyze data, develop algorithms, and create models. Files with extension m are the source code that executes MATLAB commands. A single MATLAB simulation can contain multiple .m files, each one performing a different function.
Orientation	
Byte order	
Related file formats	None.
Technical Environment	
Released	
Supported until	
Format Risk	
Developed by	None.
Supported by	 MathWorks / MathWorks
Source	 Technische Informationsbibliothek (TIB) Hannover / Technische Informationsbibliothek (TIB) Hannover
Source date	28 Apr 2022
Source description	
Last updated	28 Apr 2022
Note	

[PRONOM Interface](#)

Next Steps

- Create documentation about unknown formats
- Share documentation with the Alliance's Network of Experts
- Share with larger digital preservation community





Digital Research
Alliance of Canada

Alliance de recherche
numérique du Canada



alliancecan.ca



[@Alliance_Can](https://twitter.com/Alliance_Can)



[/AllianceCan](https://www.linkedin.com/company/AllianceCan)



amanda.tome@alliancecan.ca