

Bestandsformaten

Maaïke Verburg en Michelle van den Berk (DANS)

Met contributie van Valentijn Gilissen (Team Lead Data Processing)

14 Oktober 2024

Inhoud

- Introductie
- Bestandsformaten in de praktijk
- Tools en hulpmiddelen
- Opdracht
- Toepassing op je eigen dataset



Kennis maken met het onderwerp

Gebruik de QR code of de URL om deel te nemen aan de quiz



partici.fi/81240621



Bestandsformaten en de FAIR principes

R1.3. (meta)data meet domain-relevant community standards

- Het kiezen van bestandsformaten is gerelateerd aan het 'Reusability' principe
- Bestandsformaat met lange termijn stabiliteit
- Bestandsformaat die collega's in dezelfde discipline ook gebruiken
- Voorkeur voor open formaten t.o.v. gepatenteerde



Data Sharing and Management Snafu in 3 Short Acts - NYU Health Sciences Library;
https://www.youtube.com/watch_popup?v=66oNv_DJuPc

Bestandsformaten in de praktijk

- Open of meest gebruikt?
- 'Lossy' en kleiner of
- 'lossless' en groter?!



UK Data Service n.d.; photo: [pexels.com/Shiva Smyth](https://www.pexels.com/photo/stacked-stones-on-a-beach/)

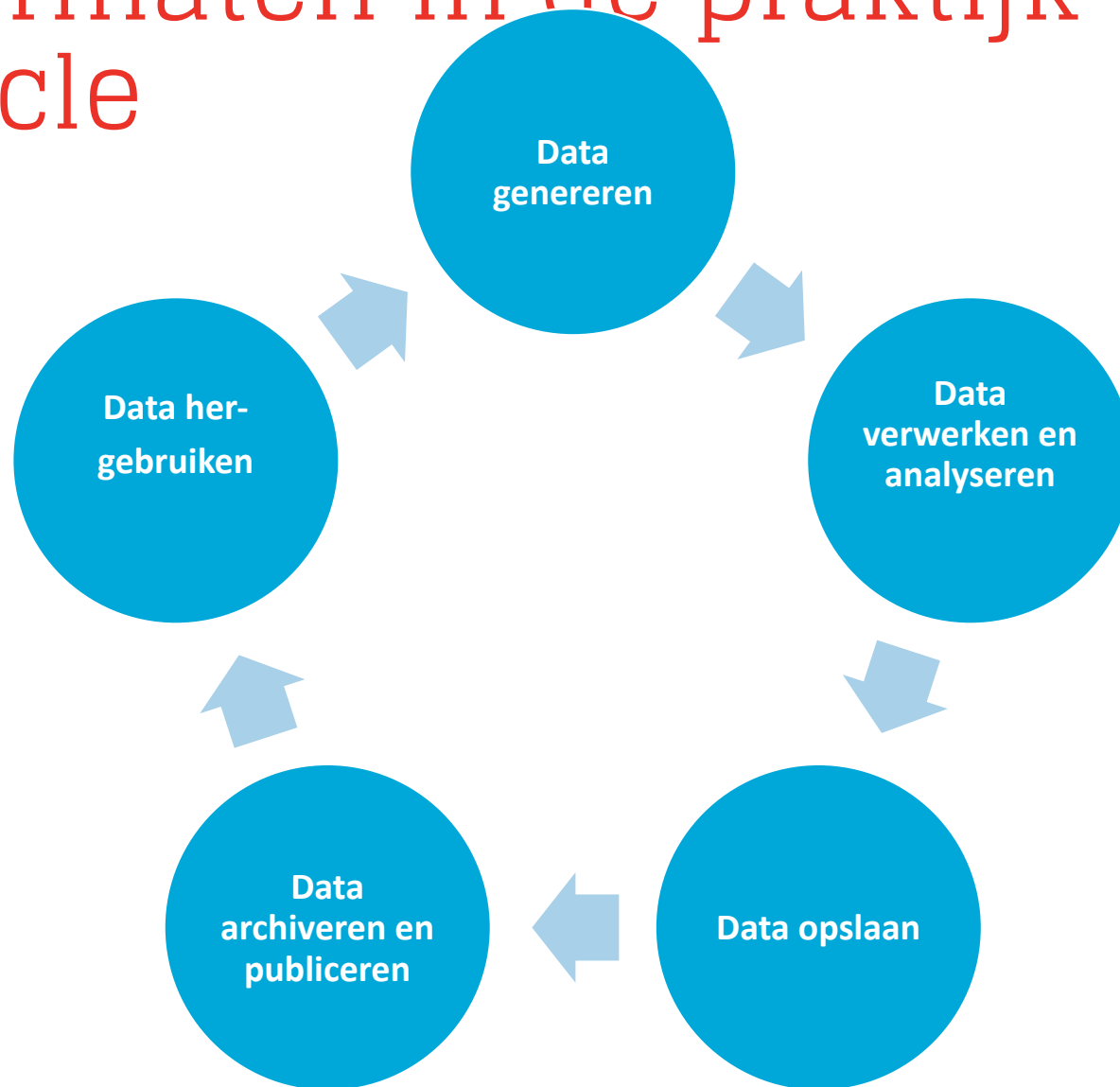
Bestandsformaten in de praktijk

Gebruiksformaten en conserveringsformaten:
korte termijn vs lange termijn



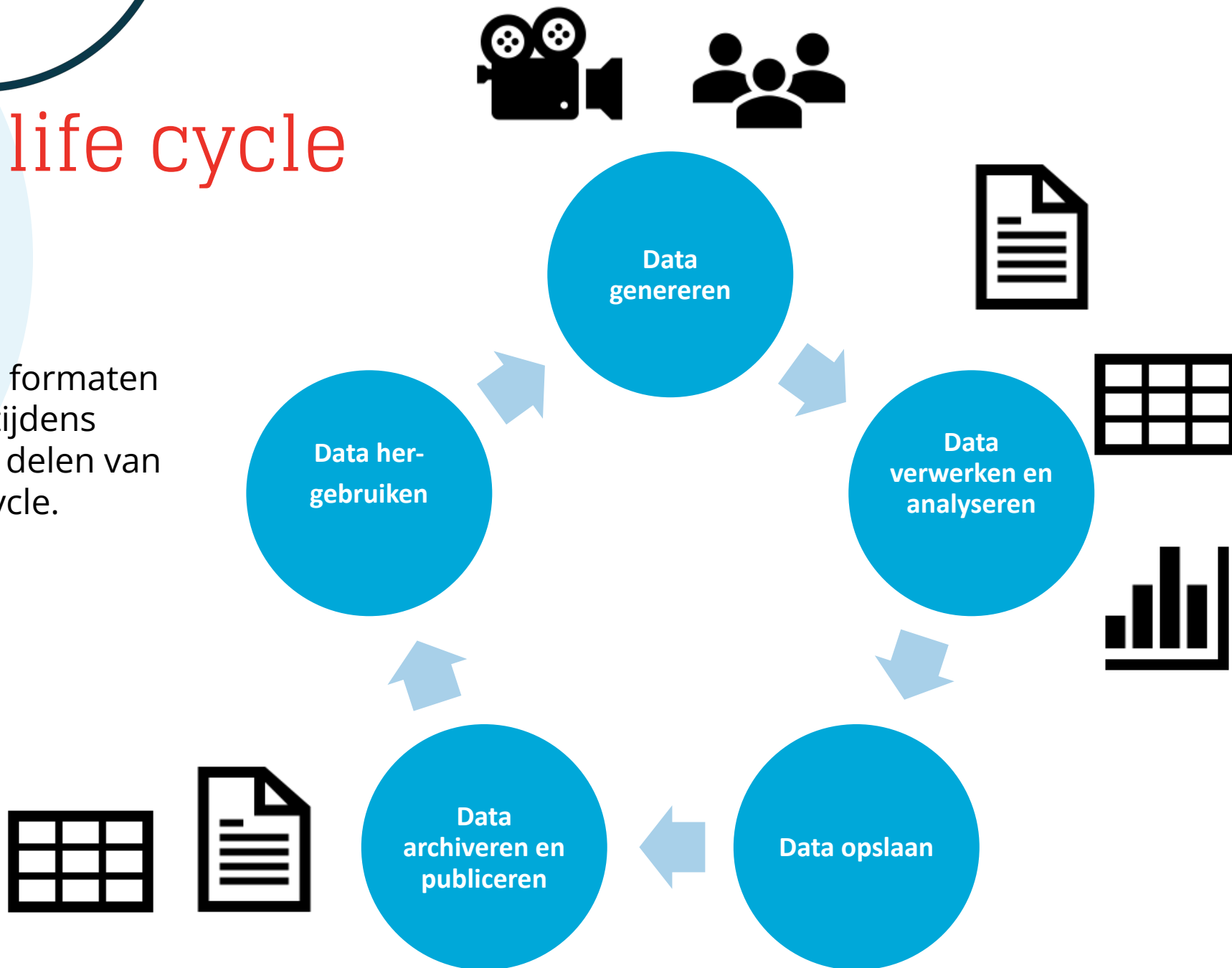
- Welk formaat is het beste voor het genereren van data?
- Welk formaat is het beste voor het analyseren van data?
- Welk formaat is het beste voor archiveren op de lange termijn?

Bestandsformaten in de praktijk – data life cycle



Data life cycle

Verschillende formaten zijn relevant tijdens verschillende delen van de data life cycle.



Data life cycle



Het 'beste' formaat van hetzelfde soort data is niet noodzakelijk hetzelfde voor de verschillende stadia van de data life cycle.



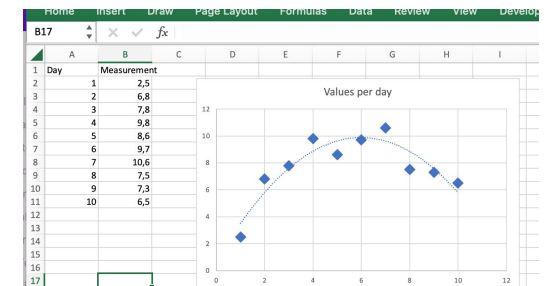
Data her-gebruiken

Data genereren

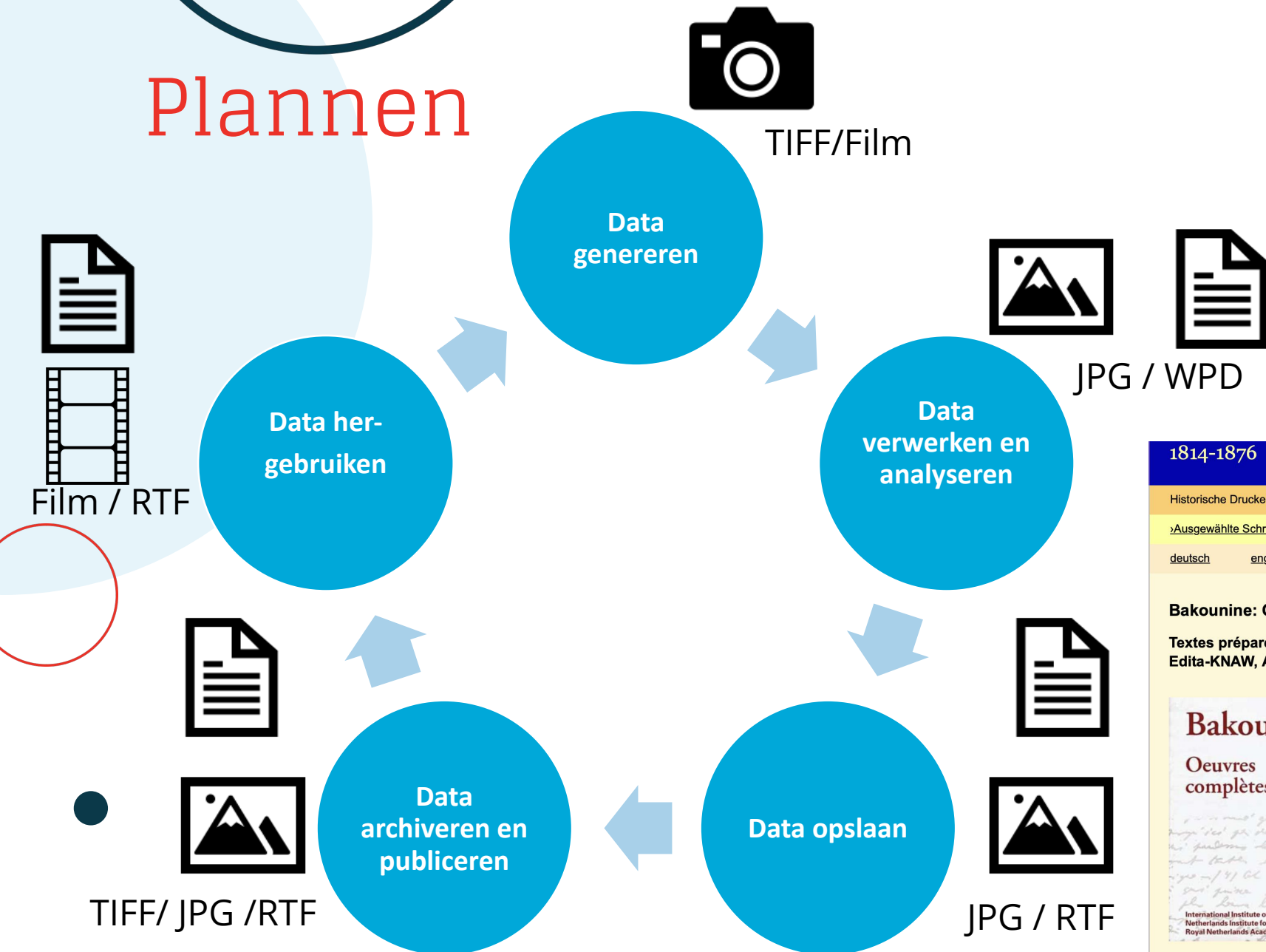
Data verwerken en analyseren

Data archiveren en publiceren

Data opslaan



Plannen



1814-1876

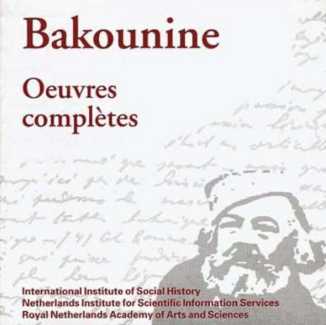
Historische Drucke Editions-Projekt Literatur von/über B. **Forschung** Links Kontakt

[»Ausgewählte Schriften«](#) **Weitere Primärliteratur** [Sekundärliteratur](#)

[deutsch](#) [englisch](#) [französisch](#) [italienisch](#) [russisch](#) [spanisch](#)

Bakounine: Oeuvres complètes [CD-ROM]

Textes préparés à l'Institut international d'Histoire sociale.
Edita-KNAW, Amsterdam 2000.



Der erste Editionsversuch von Bakunins Gesamtwerk. Erarbeitet von einer Arbeitsgruppe am »Internationaal Instituut voor Sociale Geschiedenis« (Amsterdam) in den Jahren 1992-2000 auf Grundlage der Editionsarbeiten für die »Archives Bakounine«. Alle Texte sind in der jeweiligen Originalsprache und ggf. einer französischen Übersetzung wiedergegeben.

Trotz konzeptioneller Unzulänglichkeiten – dem Verzicht auf die Wiedergabe der Aussagen Bakunins vor den sächsischen und österreichischen Untersuchungskommissionen 1849/1850, dem Fehlen jeglicher Anmerkungen, dem Verzicht auf die Publikation von Briefen an Bakunin – muß die *Bakounine CD-ROM* dennoch als unverzichtbares Arbeitsinstrument für die Bakuninforschung gelten. Zu den größten Pluspunkten zählen die technischen Möglichkeiten der CD-ROM (Suche im französischen Gesamttext, Anzeige fast aller erhaltenen Manuskripte als Bilddatei etc.), vor allem aber die beeindruckende Anzahl von bisher schwer erreichbaren oder sogar völlig neuen Dokumenten in den

Inhalt der CD-ROM

- » [Schriften](#)
- » [Briefe](#)
- » [Diverses](#)

Projektbeschreibung

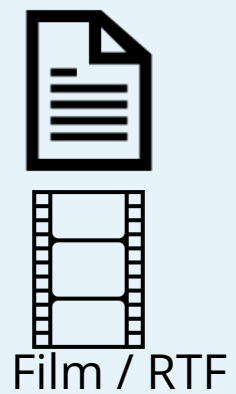
Rezensionen und Artikel

Korrekturen

Ergänzungen

- » [Schriften](#)
- » [Briefe](#)
- » [Diverses](#)

Plannen



Film / RTF



Data genereren



Data verwerken en analyseren



Data opslaan

Data archiveren en publiceren



Data hergebruiken



Search bar with 'Find' button and 'Advanced search' link.

New Search / Search Results / Michail Aleksandrovič Bakunin... / Collection Summary

Cite this Email this Export Record Add to Favorites Share Add to Selected Items

Michail Aleksandrovič Bakunin Papers

Collection Summary Content List Content and Context Access and Use Subjects

Request Order reproduction

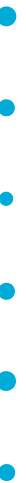
Creator Bakunin, Michail Aleksandrovič

Abstract Collection consisting of part of his papers and other

Similar Items

Group portrait J. Monchal, Charles Perron, Michail Aleksandrovič Bakunin ... by: Chevalier, A. Published: (1869, February-March.)

Adolf Reichel Papers by: Reichel, Adolf



Multiple export formats

Converteren

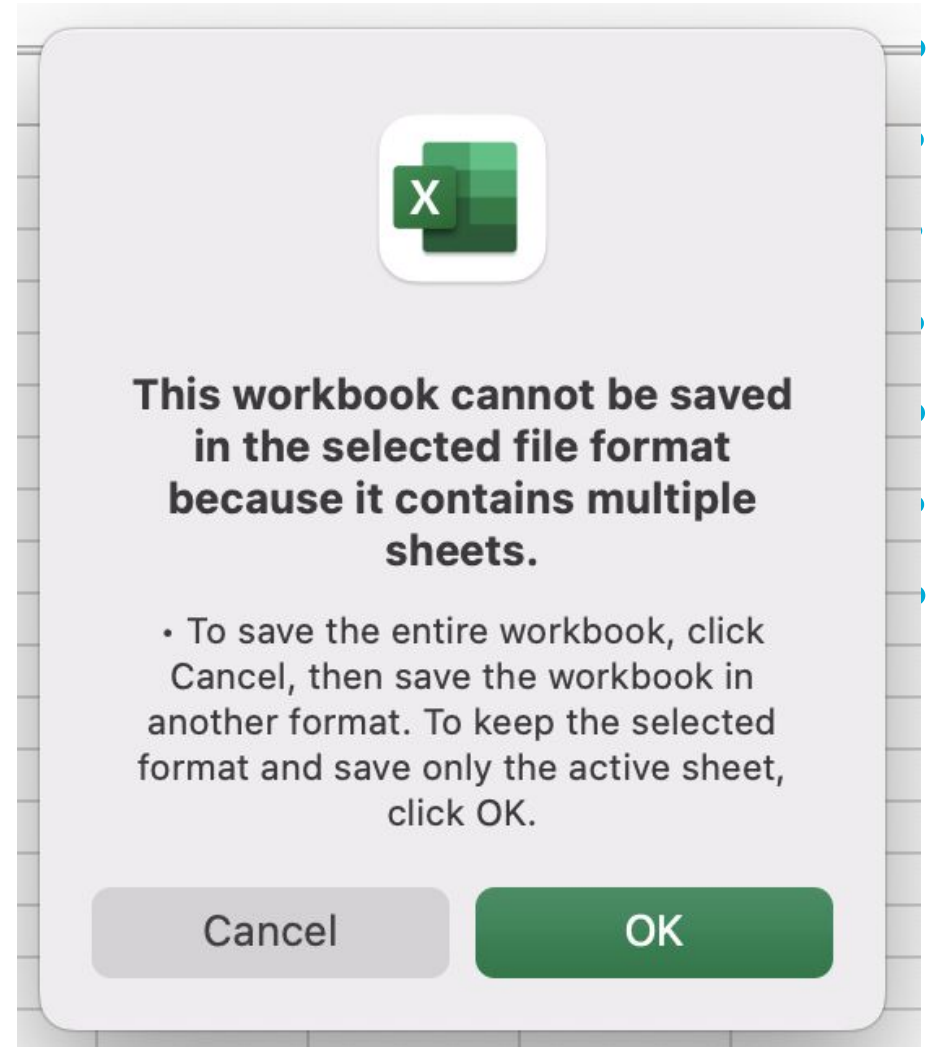
Wie gaat het doen?

	A	B	C	D	E	F	G	H
1	Year	Jan	Feb	Mar	Apr	May	Jun	Jul
2	1901	0,4755666	-0,5263906	-0,689437	-0,6897283	0,2307089	0,2424614	0,226212
3	1902	-0,6482666	-0,9141324	0,1987364	0,2386418	0,3166548	0,2626176	0,237507
4	1903	-1,034767	0,4392903	0,4172272	0,3802334	0,3017741	0,2937942	0,279854
5	1904	0,9107613	0,9516645	1,397697	2,080346	2,36191	2,300385	2,06439
6	1905	4,335482	3,524991	3,580668	3,177127	2,721378	2,445801	2,23483
7	1906	-0,1082141	-0,2556525	-0,5803101	-0,8518403	-0,7654676	-0,6742815	-0,549688
8	1907	0,3854368	0,3531811	0,360834	0,1093097	5,25E-02	4,27E-02	6,93E-0
9	1908	0,9454606	0,5973831	0,6079255	0,7120826	0,7413627	0,7520733	0,705833
10	1909	0,8784969	1,2309	-0,7513102	-0,8917843	-0,9909972	-0,8607174	-0,791246
11	1910	-0,3681316	-0,3494938	-0,6850204	0,2973971	0,4806909	0,4841326	0,474878
12	1911	-1,698672	0,655858	0,7790136	1,106845	1,687875	1,574059	-1,71E-0
13	1912	-0,6433261	2,11E-02	0,27929	0,3792913	0,650485	0,6101905	0,626000
14	1913	1,085119	1,197526	-0,339393	-0,3491245	0,1930234	0,1433113	0,159122
15	1914	1,117158	-0,5998158	-0,7205119	0,1054635	0,2886972	0,4102866	0,426193
16	1915	1,072052	-0,3948369	-0,4191109	0,543312	0,6772007	0,8199425	0,834285
17	1916	-0,9876643	-0,9189602	-1,215199	-1,489355	-1,128029	-1,01634	-0,945540
18	1917	-0,9802691	-1,082403	-1,144901	-1,1413	-1,103474	-1,074134	-1,03169
19	1918	-1,436377	-1,386652	-1,646967	6,55E-02	0,2082452	0,3245574	0,304352
20	1919	0,1046181	6,65E-02	0,1905844	0,2248848	0,3198789	0,3528494	0,359459
21	1920	0,222246	0,2507312	0,4226207	0,5076058	0,446622	0,222246	0,222246
		T	PDSI CRU	SPEI 12m	SPEI 1m	overview	+	
		↓	↓	↓	↓			
		CSV	CSV	CSV	CSV			



Converteren en mogelijk verlies van informatie

- Deel van de dataset, macros, wijzigingen
- Opmaak
- Kwaliteit, resolutie
- Metadata
- Andere schriften



Tools en hulpmiddelen

Lijsten van voorkeursformaten

- DANS - <https://dans.knaw.nl/nl/bestandsformaten/>
- UKDS - <https://ukdataservice.ac.uk/learning-hub/research-data-management/#format-your-data>

Type	Preferred format(s)	Non-preferred format(s)
Tekstdocumenten	<ul style="list-style-type: none">• PDF/A (.pdf)• ODT (.odt)	<ul style="list-style-type: none">• Microsoft Word (.doc)• Office Open XML (.docx)• Rich Text File (.rtf)• PDF anders dan PDF/A (.pdf)
Platte tekst	<ul style="list-style-type: none">• Unicode text (.txt)	<ul style="list-style-type: none">• Non-Unicode text (.txt)
Opmaaktaal	<ul style="list-style-type: none">• XML (.xml)• HTML (.html)• gerelateerde bestanden: .css, .xslt, .js, .es	<ul style="list-style-type: none">• SGML (.sgml)• Markdown (.md)
Programmeertaal	<ul style="list-style-type: none">• MATLAB• NetCDF• Text-Fabric	
Spreadsheets	<ul style="list-style-type: none">• ODS (.ods)• CSV (.csv)	<ul style="list-style-type: none">• Microsoft Excel (.xls)• Office Open XML Workbook (.xlsx)• PDF/A (.pdf)

Type of data	Recommended formats	Other acceptable formats
Quantitative tabular data with extensive metadata. A dataset with variable labels, code labels, and defined missing values, in addition to the matrix of data.	Proprietary formats of statistical packages e.g. SPSS (.sav), Stata (.dta), .sas7bdat. Delimited text and command ('setup') file (SPSS, Stata, SAS, etc.) containing metadata information. Some structured text or mark-up file containing metadata information, e.g. DDI XML file.	SPSS portable format (.por). MS Access (.mdb/.accdb).
Quantitative tabular data with minimal metadata. A matrix of data with or without column headings or variable names, but no other metadata or labeling.	Comma-separated values (CSV) file (.csv). Tab-delimited file (.tab). Including delimited text of given character set with SQL data definition statements where appropriate.	Delimited text of given character set – only characters not present in the data may be used as delimiters (.txt). Widely-used formats: MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb), OpenDocument Spreadsheet (.ods).
Geospatial data.	ESRI Shapefile (essential – .shp, .shx, .dbf, optional – .prj, .sbx, .sbn). Geo-referenced TIFF (.tif, .tiff).	ESRI Geodatabase format (.mdb). MapInfo Interchange Format (.mif) for vector data.

Tools en hulpmiddelen - vervolg

- Wegwijzer Voorkeursformaten
<https://www.wegwijzervoorkeursformaten.nl/index.php/Hoofdpagina>
- Preservation Watch - Netwerk Digitaal Erfgoed
<https://kiacommunity.nl/groups/47-preservation-watch/welcome>
- Home Computer Museum - Data herstel dienst
<https://www.homecomputermuseum.nl/diensten/data-herstel-service/>

Wegwijzer voorkeursformaten

[Gebruikersaccount](#)

[Zoeken](#)

[Stappenplan](#)

[Register](#)

[Instellingspagina](#)

[Over de Wegwijzer](#)



Wegwijzer Voorkeursformaten

Wegwijzer Voorkeursformaten

De Wegwijzer Voorkeursformaten is een website die instellingen helpt met beleid op bestandsformaten. Welke bestandsformaten zijn geschikt als archiefformaat? Kunnen we PDF's, jpg's en zip-files over 100 jaar nog uitlezen? Welke technische eigenschappen zijn kenmerkend voor het gekozen formaat? Hoe maak je beleid dat past bij jouw eigen archief? Hoe doen collega-instellingen dat?

Via een interactief [Stappenplan](#) kunnen instellingen bestandsformaten onderzoeken en een beleidskeuze maken per bestandsformaat. Keuzes van de instellingen worden in de Wegwijzer opgeslagen. Zo ontstaat een open kennisbank over het gebruik van bestandsformaten in Nederland. Hiermee bevordert de Wegwijzer dat instellingen elkaar helpen en bevragen.

De Wegwijzer biedt daarnaast een [Register](#). Het Register is een naslagwerk voor onderzoek naar bestandsformaten. Het Register combineert gegevens afkomstig uit verschillende internationale bronnen en biedt daarom een goede basis aan informatie.



Overzicht instelling

Instelling		L code
Archief De Domijnen		
Archief Eemland		
DANS	Den Haag	
EYE Filmmuseum	Amsterdam	
Erfgoed Leiden en Omstreken		
Expertisegroep 3		
Fries Film en Audio Archief		
Gemeente 's-Hertogenbosch		
Gemeentearchief Zaanstad		
Haags Gemeentearchief		
Het Flevolands Archief		
Het Nieuwe Instituut		
Historisch Centrum Limburg	Maastricht & Heerlen	NL-MtRHCL & NL-HrIRi
Koninklijke Bibliotheek		
Nationaal Archief		
Nederlands Instituut voor Beeld en Geluid	Hilversum	
Noardeast-Fryslan		
Noord-Hollands Archief		
OKV Den Ham TEST2	Den Ham	
Regionaal Archief Alkmaar		
... meer resultaten		



Gegevens bekijken: Formaten

Formaten
Klik op de filters hieronder om het aantal resultaten te verkleinen. Tussen haakjes staat het aantal resultaten dat aan de filtercriteria voldoet, en dat overblijft als u het filter selecteert. Het is mogelijk om meerdere filters te combineren. Geselecteerde filters verschijnen hierboven, en kunnen weer uitgeschakeld worden door op het kruisje achter het filter te klikken.

▼ Toepassingsgebied:

AV Wrapper (27) · Afbeelding (302) · Algemeen (45) · Audio (71) · Audio sequencing (23) · Berichtuitwisseling (1) · Bewegend beeld (22) · CAD/CAM en 3D (241) · Container (31) · Database (57) · Digitale handtekening (7) · Disk image (14) · Exacte wetenschappen (57) · Geo-informatie (65) · Interactieve media (52) · Lettertype (14) · Muzieknotatie (6) · Productiviteit (492) · Publishing (76) · Software en code (130) · Systeem ondersteunende software (1) · Tekst (86) · Webarchivering (8) · Wetenschappelijke toepassingen (62)

Formaatfamilie:

Zoekresultaten

▼ Toepassingsgebied PRONOM:

Aggregate (51) · Audio (98) · Database (120) · Dataset (116) · Email (5) · Font (7) · GIS (34) · Image (Raster) (211) · Image (Vector) (190) · Model (118) · Page Description (39) · Presentation (127) · Spreadsheet (66) · Text (Mark-up) (70) · Text (Structured) (98) · Text (Unstructured) (12) · Video (39) · Word Processor (159)

MIME-type:

Zoekresultaten

Verantwoordelijke organisatie:

Zoekresultaten

▼ Openheid:

Overzicht voorkeursformaten

Deze pagina biedt een overzicht van alle bestandsformaten die door een of meerdere instellingen zijn gekozen.

Formaatfamilie	Formaat	Instelling(en) met voorkeursbeleid t.a.v. formaat
PDF	Acrobat PDF/A - Portable Document Format 1a	DANS, Expertisegroep 3, Historisch Centrum Limburg, Stadsarchief Amsterdam
PDF	Acrobat PDF/A - Portable Document Format 1b	DANS, Expertisegroep 3, Historisch Centrum Limburg
PDF	Acrobat PDF/A - Portable Document Format 2a	DANS, Expertisegroep 3, Historisch Centrum Limburg, Stadsarchief Amsterdam
PDF	Acrobat PDF/A - Portable Document Format 2b	DANS
PDF	Acrobat PDF/A - Portable Document Format 2u	DANS, Historisch Centrum Limburg
PDF	Acrobat PDF/A - Portable Document Format 3a	DANS
PDF	Acrobat PDF/A - Portable Document Format 3b	DANS
PDF	Acrobat PDF/A - Portable Document Format 3u	DANS
PNG	Animated Portable Network Graphics	Stadsarchief Amsterdam
AutoCAD	AutoCAD Drawing R11/12	DANS
Blender	Blender 3D 32 bit	DANS
Blender	Blender 3D 64 bit	DANS
WAVE	Broadcast WAVE 0 Generic	Nederlands Instituut voor Beeld en Geluid
WAVE	Broadcast WAVE 1 Generic	Nederlands Instituut voor Beeld en Geluid, Stadsarchief Amsterdam
WAVE	Broadcast WAVE 2 Generic	Historisch Centrum Limburg, Nederlands Instituut voor Beeld en Geluid
Comma Separated Values formaatfamilie	Comma Separated Values	DANS, Stadsarchief Amsterdam
Digital Imaging and Communications in Medicine		

Preservation Watch

- Preservation Watch': Toezicht houden op ontwikkelingen op het gebied van data preservering (inspelen op nieuwe kansen, risico's tijdig aan zien komen en aanpakken, etc.)
- Netwerk Digitaal Erfgoed heeft een groep experts samengesteld om deze functie uit te voeren.
- Onder andere: monitoren van bestandsformaten

The screenshot shows the Preservation Watch website. On the left is a dark sidebar with navigation links: Entree, Startpagina, Tijdlijn, Groepen, Wie is wie?, Agenda, and Deel. The main header is pink and contains the KIA logo, the title 'Preservation Watch', and links to 'Entree' and 'Leden'. A search bar and a 'Word lid' button are also present. The main content area has a green callout box with a question about monitoring technological developments. Below this is a section for 'Nieuwste Tech Watch Blogs' with three green boxes labeled 'Groene software', 'Een leeslijst over Chat-', and 'Monitoring van'. A 'Blogreeksen' section follows, displaying three columns of blog entries with titles like 'Monitoring bestandsformaten', 'Green IT', and 'Bias in Metadata'.

Entree

Startpagina

Tijdlijn

Groepen

Wie is wie?

Agenda

Deel

KIA

Preservation Watch Entree Leden

Snelmenu

Word lid

Erfgoedcollecties zijn in toenemende mate digitaal en online beschikbaar. De digitale wereld is extreem dynamisch. Sommige ontwikkelingen bieden kansen om erfgoed nog beter houdbaar, bruikbaar en zichtbaar te maken. Andere ontwikkelingen zijn juist bedreigend of risicovol zoals de klimaatimpact van digitaal erfgoed. Of beiden, zoals AI

Om hier op in te spelen, is het belangrijk om zicht te hebben op die kansen en risico's. Dat vergt de inrichting van een functie die preservation watch genoemd wordt. En dat vraagt veel van een individuele erfgoedorganisatie of archiefdienst.

Preservation watch overstijgt het belang en de mogelijkheden van afzonderlijke instellingen. Daarom hebben enkele deelnemers in het [Netwerk Digitaal Erfgoed](#) zich verenigt in een netwerkgroep. De netwerkgroep Preservation Watch richt zich op het signaleren van met name technologische ontwikkelingen.

Belangrijke signalen en blogs door techwatchers worden hier gepubliceerd.

Heb je een vraag of wil je graag de discussie aangaan over de monitoring van technologische ontwikkelingen? Plaats het op het kennisplatform Preservation.

Nieuwste Tech Watch Blogs

Groene software

Een leeslijst over Chat-

Monitoring van

Blogreeksen

Rein van 't Veer: Monitoring bestandsformaten

Tineke van Heijst: Green IT

Ryan Brate: Bias in Metadata

Monitoring van

Groene software: Meten om te weten en te verbeteren

Bias in metadata 3: Adding (polyvocal) context to semantic

Home Computer Museum

Een grote collectie oude computers en softwares en om oude data in te lezen en te converteren naar nieuwere bestanden.



Data herstel en Digitaal Erfgoed

Deze pagina is in bewerking

It happens to everyone, accidentally deleting important data from a storage medium. Or you found a diskette or other medium that you can no longer read? The HomeComputerMuseum offers a solution in the way that only the HomeComputerMuseum can.

"Finally being able to read my lost data again!"

Data deleted?

No cure, no pay

Accidentally deleted data from your SD card, USB drive or hard disk? We have specialized data recovery software and can even temporarily revive failed hard drives to secure data.

Your data is offered to you via your own new medium or shared via a cloud service. As soon as we have shared this data and you have it safely back in your hands, we will irrevocably delete the data again in accordance with the GDPR standards.

We do not calculate per MB, but what it actually costs us in terms of time. Starting price is € 25 and only needs to be paid if we have been able to restore the data.

Unreadable medium?

For companies, individuals and institutions

In collaboration with the Network Digital Heritage, we offer companies, private individuals and (cultural) institutions the opportunity to read old media (such as diskettes in all formats, tape drives, ZIP disks, SyQuest disks, JAZ disks and even cassette tapes).

You read it correctly, the reading CAN be done by us, but the data may be too special that you would prefer to do this yourself. No problem, we have the self-processing street for that. Under the guidance of an employee, we will show you how to retrieve data from the media, after which you can carry it out yourself, while enjoying a cup of coffee or tea.

Of course we can also read this data for you and, just like with deletion, send it to you after which it will be deleted.

For smaller cultural institutions this is free, for individuals and companies it is charged at an hourly rate.

Opdracht

Gebruik de QR code of de URL om naar Particify te gaan.



partici.fi/81240621



Any questions of feedback?

-
-
-
-
-
-
-

More information

Visit our website **www.dans.knaw.nl**

And follow us online



Mastodon @DANS_knaw_nwo



LinkedIn @DANS



X @DANS_knaw_nwo

