

Eisen door de verschillende publicatiekanalen/uitgevers gesteld aan wetenschappers op het gebied van RDM

Er is gekeken naar de eisen die op de website van uitgevers van wetenschappelijke publicaties worden vermeld ten aanzien van een Data Availability policy. Dan gaat het vooral om wetenschappelijke tijdschriften, voor boeken zal er over het algemeen nog geen beleid zijn en in het Promotiereglement van de UvA staat geen bepaling over het delen van data. De voorwaarden die door funding bodies worden gesteld zijn buiten beschouwing gelaten.

In de periode 2000 - 2013 werd door medewerkers van de FNWI bij ongeveer 150 uitgevers gepubliceerd. Ongeveer 95 % van de publicaties werd uitgegeven door de 28 uitgevers in Tabel 1. Van deze uitgevers hebben er 8 (29 %) een duidelijk omschreven Data Availability policy op hun website. American Physical Society en Oxford University Press hebben wel een uitleg over het toevoegen van supplementary material, maar hieronder wordt aanvullend materiaal verstaan (afbeeldingen, geluid, filmpjes). In die zin lijkt het minder van belang voor RDM-beleid.

In het geval dat een uitgever tijdschriften van een wetenschappelijke society in zijn pakket aanbiedt, kan het zijn dat de Society wel eisen stelt op het gebied van het beschikbaar stellen van data. Dat is onder andere bij een aantal titels van Wiley en University of Chicago Press het geval. Op de hoofdwedite was er niets over te vinden, maar op de website van het tijdschrift wel.

Tabel 1: Uitgevers met data availability policy en hun percentuele aandeel in het totale aantal FNWI publicaties in 2000 - 2013¹.

Uitgever	Data availability policy	Beschikbaar stellen van data	Aandeel in publicaties FNWI in 2000 - 2013	Cumulatief
Elsevier	Ja	Aanmoediging	25,04%	25%
Springer	Deels, Springer Open	Aanmoediging, sommige titels verplicht	18,32%	43%
John Wiley And Sons	Nee	Sommige titels wel	15,45%	59%
IOP	Nee	-	6,46%	65%
American Chemical Society	Nee	-	5,71%	71%
American Physical Society	Nee, wel sup. mat.		5,03%	76%
Royal Society Of Chemistry	Ja, Royal Society	Voorwaarde	2,22%	78%

¹ Het aandeel in publicaties door FNWI medewerkers is berekend op grond van cijfers in UvA-View, <http://uva-view.nl/>

American Institute Of Physics	Nee	-	1,94%	80%
Oxford Journals	Nee, wel sup. Mat.	-	1,76%	82%
Taylor & Francis	Nee	-	1,51%	83%
IEEE	Nee	-	1,48%	85%
Nature Publishing Group	Ja	Op aanvraag	1,37%	86%
Cambridge University Press	Nee	-	1,16%	87%
Public Library of Science	Ja	Verplichting	0,85%	88%
American Society For Microbiology	Nee	-	0,78%	89%
ACM	Nee	-	0,62%	90%
American Association For The Advancement Of Science	Ja	In principe wel, in ieder geval op aanvraag	0,61%	90%
Biomed Central	Deels, gelijk aan SpringerOpen	Aanmoediging, sommige titels verplicht	0,49%	91%
National Academy Of Sciences	Nee	-	0,44%	91%
American Society For Biochemistry And Molecular Biology	Nee	-	0,42%	92%
The University Of Chicago Press	Nee	Sommige titels wel	0,42%	92%
Ecological Society Of America	Nee	-	0,41%	93%
Royal Society	Ja	Voorwaarde	0,40%	93%
World Scientific Publishing Company	Nee	-	0,39%	93%
American Society Of Plant Biologists	Nee	-	0,37%	94%
Sage Publications	Nee	-	0,32%	94%
Society For Neuroscience	Nee	-	0,29%	94%
Inderscience Publishers	Nee	-	0,29%	95%

De uitgever met het strengste beleid is Public Library of Science (PLOS)². Sinds maart 2014 moeten alle data die ten grondslag liggen aan de in het artikel beschreven bevindingen volledig en zonder beperkingen openbaar beschikbaar zijn. Een Data Availability Statement moet worden bijgevoegd waarin de auteurs verklaren dat dit inderdaad het geval is.

² PLOS: Data Availability. <http://journals.plos.org/plosone/s/data-availability>

Volgens van Noorden (2014) voldeden eind 2014 slechts 8 (40%) van de 20 gecheckte papers uit PLOS ONE volledig aan het Open data mandaat van PLOS ONE, de overige papers bevatten onvoldoende gegevens om de bevindingen te repliceren.³

Het aandeel van artikelen in PLOS tijdschriften in het totale aantal FNWI publicaties bedroeg in 2014 nog geen procent.

Nature stelt dat de gegevens beschikbaar moeten zijn voor editor en reviewers en op aanvraag. Science stelt dat alle data die nodig zijn beschikbaar moeten zijn voor elke lezer van Science.

Van de drie uitgevers met bij elkaar een aandeel van bijna 60 % in het aantal FNWI publicaties wordt alleen voor enkele titels van SpringerOpen en wat Society tijdschriften gevraagd dat alle data vrij beschikbaar zijn. Elsevier en Springer moedigen het plaatsen van gegevens in openbare repositories wel aan, maar verplichten dit niet voor het merendeel van hun titels niet.

Voor bepaalde types data is plaatsing in een specifieke repository wel verplicht of wordt sterk aanbevolen. Hiervoor zijn er al vele erkende vakspecifieke repositories beschikbaar. (Zie tabel 2.)

Voor data die niet in een vak specifiek repository kunnen worden ondergebracht zijn generieke repositories beschikbaar. Bovendien bieden veel uitgevers de mogelijkheid om data als supplementair materiaal op hun eigen website te hosten.

Bij alle uitgevers die het beschikbaar stellen van onderzoeksdata verplichten kunnen de data gedeponeerd te worden in een openbare repository zoals Figshare of Dryad.

Het UvA-HvA-beleid met betrekking tot het publiceren van onderzoeksdata is dat onderzoeksdata gepubliceerd moet worden, tenzij er gegronde redenen zijn om dat niet te doen (privacy, staatsveiligheid etc.). Dit beleid en de data availability policy's die zijn onderzocht, zitten dus - met financiers als NWO - op dezelfde lijn.

Tabel 2: Repositories per data type zoals vermeld op website van uitgevers.

Data type	Nature	PLOS	Royal Society	Science	SpringerOpen
If no specific repository available	Dryad	Dryad	Dryad (10 GB wordt vergoed)		LabArchives, LLC (gratis voor SpringerOpen auteurs)
	Figshare	Figshare	Figshare		
	Supplementary Information file		Supplementary material	Supplementary material	
Protein sequences	Uniprot		GenBank	GenBank	SwissProt
			EMBL	EMBL	
			DDBJ	DDBJ	
			Protein Information Resource (PIR)	Protein Databank of Japan	PIR
DNA and RNA	Genbank	GenBank	GenBank	GenBank	GenBank

³ van Noorden, Richard. Confusion over publisher's pioneering open-data rules. *Nature*, 2014, 515.7528: 478-478.

sequences					
	DNA DataBank of Japan (DDBJ)	DDBJ	DDBJ	DDBJ	DDBJ
	EMBL Nucleotide Sequence Database (ENA)	EMBL	EMBL	EMBL	EMBL
DNA and RNA sequencing data	NCBI Trace Archive				
	NCBI Sequence Read Archive (SRA)				
Genetic polymorphisms	dbSNP				
	dbVar				
	European Variation Archive (EVA)				
Linked genotype and phenotype data	dbGAP				
	The European Genome-phenome Archive (EGA)				
Macromolecular structure	Worldwide Protein Data Bank (wwPDB)			Worldwide Protein Data Bank	Worldwide Protein data Bank
	Biological Magnetic Resonance Data Bank (BMRB)			Electron Microscopy Data Bank (EMDB)	
	Electron Microscopy Data Bank (EMDB)			Macromolecular Structure Database MSD EMBL-EBI	
				BioMag Res Bank	
Microarray data (must be MIAME compliant)	Gene Expression Omnibus (GEO)	Gene Expression Omnibus (GEO)	Gene Expression Omnibus (GEO)	Gene Expression Omnibus (GEO)	Gene Expression Omnibus (GEO)
	ArrayExpress	ArrayExpress	ArrayExpress	ArrayExpress	ArrayExpress
					CIBEX
Biological materials	Jackson Laboratory				
	European Mouse Mutant Archive (EMMA)				
	European Conditional Mouse Mutagenesis Program (EUCOMM)				
	Knockout Mouse Project (KOMP)				
	RIKEN Bioresource Centre				
	Mutant Mouse Regional				

	Resource Centers				
	American Type Culture Collection				
Plasmids	Addgene				Addgene
					PlasmID
Ecological data		Dryad		Dryad	
Phylogenetic data			TreeBase		
Proteomic data			EBI PRIDE		
Chemical data			PubChem		
Crystallographic data for small molecules	Cambridge Structural Database			Cambridge Structural Database	Cambridge Crystallographic Data Centre
Nucleic Acids					Nucleic Acids Database
Mass spectrometry data					ProteomeExchange
Chemical structures					PubChem Substance
					PubChem BioAssay
Computational modelling					BioModels database
Climate data				NOAA climate repository	

Opmerkingen per Uitgever

Public Library of Science PLOS⁴

Alle data en gerelateerde metadata moeten worden gedeponeerd in een geschikte openbare repository, tenzij ze al deel uitmaken van het ingediende artikel. Het Data Availability statement dient de namen van de repositories, DOI's en/of accession numbers voor de relevante datasets te specificeren.

Kleinere datasets kunnen worden geüpload als Supporting Information files. Deze moeten worden aangeleverd in een formaat waaruit de data eenvoudig kunnen worden geëxtraheerd. Dus bijvoorbeeld liever een spreadsheet dan een pdf, als het gaat om data in tabellen.

Wanneer ethische of wettelijke redenen bestaan om de data niet in een repository te deponeren of bij te voegen, dan is het mogelijk aan te geven dat ze op verzoek beschikbaar worden gesteld aan alle geïnteresseerde onderzoekers.

Persoonlijke belangen, zoals patenten of potentieel toekomstig onderzoek gelden niet als reden voor een uitzondering op het beleid.

PLOS vereist dat auteurs voldoen aan vakspecifieke standaarden voor het prepareren en registreren van data. Te kiezen repositories dienen bij voorkeur te voldoen aan geaccepteerde standaarden, zoals de criteria van het Centre for Research Libraries of het Data Seal of Approval. Licenties moeten niet restrictiever zijn dan CC-BY.

⁴ Data Availability. <http://journals.plos.org/plosone/s/data-availability>

Gegevens uit onderzoek met menselijke proefpersonen moeten zo worden behandeld dat ze de privacy van de deelnemers niet in gevaar brengen.

Nature Publishing Group

Voor publicaties in Nature tijdschriften⁵ zijn auteurs verplicht zijn om alle materiaal, gegevens, code, en de bijbehorende protocollen op verzoek onmiddellijk beschikbaar te stellen aan de lezers. Eventuele beperkingen van de beschikbaarheid moet worden opgegeven aan de redactie. Elke beperking moet worden vermeld in het artikel.

Sinds mei 2013 moeten voor life sciences papers details over de experimentele en analytische opzet worden bijgevoegd. Hiervoor wordt het invullen van een checklist gevraagd.

Ondersteunende data moeten beschikbaar zijn voor editors en peer reviewers. Data dienen bij voorkeur gedeponeerd te worden in een openbare repository. Scientific Data (ISSN 2052-4463), een zuster tijdschrift van Nature houdt een lijst bij vinden van goedgekeurde en aanbevolen repositories, georganiseerd naar discipline. (<http://www.nature.com/sdata/data-policies/repositories>).

Een minder wenselijk alternatief is het beschikbaar maken van data sets als Supplementary Information files, die na publicatie via Nature.com vrij toegankelijk zijn.

Als het technisch onmogelijk is om een databestand aan te leveren, moeten de auteurs een URL of andere unieke identifier aanleveren.

Auteurs worden aangemoedigd om een Data Descriptor te publiceren in Scientific Data (Open Access zuster tijdschrift van Nature.com; <http://www.nature.com/sdata/>; ISSN 2052-4463)

Cellijnen dienen bij voorkeur gedeponeerd te worden in repositories met een certificaat van authenticiteit. Op de website zijn links naar aanvullende bronnen te vinden voor het checken van de identificatie van cellijnen.

Computer Code moet op verzoek beschikbaar moet worden gesteld aan editors en reviewers.

Aanbevolen wordt experimentele protocollen te delen via Protocol Exchange (ISSN 2043-0116 <http://www.nature.com/protocolexchange/>).

Voor Clinical trials bestaan aparte richtlijnen.

Science

Alle gegevens die nodig zijn om de conclusies die in het manuscript worden getrokken te begrijpen, te evalueren, en uit te breiden moeten beschikbaar zijn voor de lezers van Science⁶. Datasets moeten gedeponeerd in goedgekeurde repositories en accession numbers moeten in het

⁵ Nature: Availability of data, material and methods
<http://www.nature.com/authors/policies/availability.html>

⁶ Science: Data and materials availability
http://www.sciencemag.org/site/feature/contribinfo/prep/gen_info.xhtml

gepubliceerde artikel worden opgenomen. Het naleven van MIBBI richtlijnen (Minimum Information for Biological and Biomedical Investigations) wordt aangemoedigd.

Dit geldt ook voor alle computer code die gebruikt is bij het creëren of analyseren van de data.

Na de publicatie moet aan alle redelijke verzoeken om gegevens en materialen te verstrekken, worden voldaan. Beperkingen hierop moeten bekend worden gemaakt aan de editors. Fossielen of andere zeldzame exemplaren moeten worden gedeponneerd in een openbaar museum of repository en beschikbaar voor onderzoek.

Grote datasets waarvoor geen goedgekeurde repository beschikbaar is, kunnen ondergebracht worden als supplementary materials bij Science, of als dat niet mogelijk is, op een institutionele web site mits een kopie van de gegevens in depot wordt gehouden bij Science om beschikbaarheid voor de lezers te garanderen.

Elsevier

Elsevier⁷ ondersteunt het principe dat onderzoek gegevens vrij beschikbaar worden gesteld voor alle onderzoekers. Auteurs worden aangemoedigd hun datasets te plaatsen in relevante data repositories of andere beschikbare kanalen. In gevallen waarin er geen relevante data repository voorhanden is, stelt Elsevier auteurs in staat om aanvullende gegevens bij het gepubliceerde artikel op te slaan als supplementair bestand. Auteurs behouden het copyright van dergelijke aanvullende gegevens.

Het gebruiken van datacitatie praktijken zoals neergelegd in de Joint Declaration of Data Citation Principles (<https://www.force11.org/datacitation>) wordt aangemoedigd.

Springer Open / BioMedCentral

Sommige Springer Open/BioMed Central tijdschriften ondersteunen of verplichten een link naar gegevens in een open access repository of naar additionele bestanden die met het artikel gepubliceerd worden. In het geval van een link moet een unieke persistent identifier worden gegeven en een hyperlink naar de gegevens.

SpringerOpen / BioMedCentral werkt samen met o.a. DataCite aan een lijst met data repositories. Deze lijst kan gevonden worden op de website van DataCite.

Door een overeenkomst met LabArchives, LLC, kunnen auteurs die manuscripten aanbieden aan SpringerOpen/BioMed tijdschriften een gratis abonnement krijgen op LabArchives. Dat is een Electronic Laboratory Notebook met 100 MB opslagruimte. Daarmee kunnen bestanden gedeeld en gepubliceerd worden; van het artikel kan naar deze bestanden worden gelinkt. Aan bestanden die via LabArchives worden gepubliceerd wordt een digital object identifier (DOI) toegekend.

Royal Society Publishing

Het is een voorwaarde dat auteurs de data en onderzoeksmaterialen die de resultaten in hun artikel ondersteunen beschikbaar maken. Als er geen geschikte repository beschikbaar is kunnen de gegevens geüpload worden als supplementary material. Een aantal tijdschriften hebben het

⁷ <http://www.elsevier.com/about/open-science/research-data>

deponeren van data in Dryad geïntegreerd in hun submission system en vergoeden de kosten voor het deponeren van 10 GB data.

Uitzonderingen kunnen gemaakt worden voor gevoelige gegevens zoals afkomstig van onderzoek aan menselijke onderwerpen en vindplaatsen van bedreigde soorten.

Alle supplementary material wordt ter beschikking gesteld onder een CC-BY licentie.

John Wiley & Sons

Op de website van John Wiley & Sons⁸ zijn geen eisen te vinden. Wiley werkt samen met organisaties als CODATA, het World Data System, de Research Data Alliance, DataCite en NISO teneinde initiatieven te ontwikkelen die het uiteindelijk mogelijk maken onderzoeksgegevens te gebruiken, hergebruiken, citeren en op te slaan.

Titels die worden gepubliceerd voor een wetenschappelijke Society, zoals bijvoorbeeld de British Ecological Society, stellen soms wel als voorwaarden dat alle gegevens, wiskundige en computermodellen die de resultaten ondersteunen worden gearchiveerd in een geschikte openbaar repository (Dryade, TreeBase, NERC Data Centre, GenBank, Figshare) of een andere repository naar keuze.

Ecological Society Of America

ESA onderhoudt een repository met gegevens van datasets behorend bij artikelen in ESA tijdschriften: ESA Data Registry (<http://data.esa.org/esa/style/skins/esa/index.jsp>) Dit zijn alleen beschrijvingen, de datasets zelf worden elders gedeponeed.

University of Chicago Press

Sommige tijdschriften bevatten de Data policy met eisen voor het leveren van computerprogramma's, configuratiebestanden of scripts gebruikt om het experiment uit te voeren en / of om de gegevens te analyseren.

American Physical Society

Aanvullend materiaal dat niet essentieel is voor het begrijpen van de hoofdresultaten van het artikel kan in elektronische vorm gedeponeed worden als Supplemental Material (SM).

Oxford University Press

Ondersteunend materiaal dat niet kan worden opgenomen, en dat niet essentieel is voor opname in de volledige tekst van het manuscript, maar desondanks de lezer ten goede zou komen, kan worden opgenomen als Supplementary data.

George Meerburg

Juni 2015

⁸ http://exchanges.wiley.com/data-curation_251.html