



Regeling Datamanagement Universiteit Leiden

Aandachtspunten en toelichtingen bij de verschillende artikelen zijn schuin gedrukt.

A. Definities en uitgangspunten

- 1. Deze regeling is bedoeld als universitair raamwerk voor een datamanagementbeleid en vergt nadere uitwerking op disciplineniveau. Het is aan elk faculteitsbestuur om te bepalen hoe de regeling binnen de faculteit nader uitgewerkt wordt en welke rol hierbij is weggelegd voor de instituten.**
- 2. Onder onderzoeksdata wordt verstaan: alle gegevens die worden verzameld en gegenereerd tijdens wetenschappelijk onderzoek.**
Hieronder vallen ook de gegevens die worden verkregen door bewerking en analyse van (ruwe) onderzoeksdata. We volgen hiernmee NWO, dat stelt “wij verstaan onder ‘data’ zowel verzamelde, onbewerkte data alsook geanalyseerde, gegenereerde data”.
Dit uitgangspunt betekent niet dat alle data ook moeten worden opgeslagen. De beslissing over welke data al dan niet op te slaan wordt neergelegd op het niveau van de discipline (zie artikel 16).
- 3. Deze regeling is van toepassing op alle digitale en niet-digitale onderzoeksdata.**
Hiermee volgen we het beleid van NWO. Bij niet-digitale data kan bijvoorbeeld gedacht worden aan opgeslagen biomateriaal, geluidsopnames, papieren vragenlijsten en papieren labjournaals. NB: deze bepaling betekent niet dat alle niet-digitale data per se gedigitaliseerd moeten worden, maar wel dat voor niet-digitale data dezelfde eisen gelden m.b.t. vindbaarheid, herbruikbaarheid, duurzame opslag, etc.

Voorafgaand aan het onderzoek

- 4. Voor de start van de dataverzameling van een onderzoeksproject moet een datamanagementplan (DMP) worden opgesteld. In het DMP wordt het datamanagementbeleid van de faculteit/het instituut specifiek uitgewerkt voor het desbetreffende onderzoeksproject.**

Tijdens het onderzoek

- 5. Gedurende het onderzoek moeten onderzoeksdata veilig worden opgeslagen, hetgeen betekent dat de integriteit, beschikbaarheid en – indien vereist - vertrouwelijkheid van de data gegarandeerd moeten blijven.**
Let op: na afloop van het onderzoek moeten de onderzoeksdata duurzaam worden opgeslagen in combinatie met de metadata, software en overige documentatie die nodig zijn voor hergebruik (artikel 7). Dit moet ook tijdens het onderzoek in het oog gehouden worden, bijvoorbeeld omdat het veelal onmogelijk zal zijn om na het onderzoek alle data op een goede manier te voorzien van metadata.



Na afloop van het onderzoek

6. **Onderzoeksdata moeten op een zodanige manier worden beheerd dat zij ten laatste**
- **op het moment van de laatste publicatie voortkomend uit het onderzoek, of**
- **op het moment van het formeel afsluiten van het onderzoeksproject**
ten minste vindbaar, toegankelijk, begrijpelijk, herbruikbaar en duurzaam opgeslagen zijn.

We noemen hier twee momenten, omdat niet elk onderzoek wordt afgesloten met een publicatie. Toch is het ook dan van belang dat de onderzoeksdata worden opgeslagen en terugvindbaar zijn. Extern gefinancierde onderzoeksprojecten zullen in de regel formeel worden afgerond bij de financieel-administratieve afsluiting.

Een uitleg over de gebruikte criteria:

- Vindbaar

De data moeten vindbaar zijn voor andere onderzoekers en betrokkenen. De vindbaarheid wordt bevorderd door de data te deponeren in een (vakspecifiek) data-archief of repository. De gegevens over de data (metadata) zijn dan op een gestandaardiseerde manier geregistreerd en vindbaar, ook voor zoekmachines. De vindbaarheid in de toekomst wordt gegarandeerd door aan de data een persistent identifieerder toe te kennen. Een veel gebruikte identifieerder voor data is DOI (digital object identifier). Een DOI wordt toegekend door een data-archief of -uitgever en is een uniek nummer van een digitaal object, in dit geval een dataset. Een toegekende DOI blijft hetzelfde, ook als de plaats van de dataset (URL) verandert. Een DOI of andere persistent identifieerder wordt gebruikt om de dataset te citeren of om er naar te verwijzen.

- Toegankelijk:

Toegankelijkheid van data betekent niet per definitie dat de data direct volledig openbaar gemaakt moeten worden. In sommige gevallen zullen de data bijvoorbeeld voor verificatie toegankelijk gemaakt moeten worden voor een subsidieverstrekker of tijdschrift, maar blijven ze voor de brede buitenwereld ontoegankelijk. Bij privacygevoelige data zal volledige openbaarheid nooit een optie zijn. Onderzoeksdata kunnen onder embargo worden opgeslagen: alleen degene die de data heeft gedeponerd heeft toegang. De duur van het embargo wordt in overleg met het data-archief vastgelegd. Een andere mogelijkheid is om alleen toegang te verlenen als daartoe een verzoek wordt ingediend aan de onderzoeker. Deze weet wie zijn data raadpleegt en maakt afspraken over het (her)gebruik ervan. NB: Als ervoor wordt gekozen om de toegang tot de data te beperken, is het wel noodzakelijk dat de metadata (omschrijving van de data) vindbaar zijn.

- Begrijpelijk

De data moeten op een zodanige manier beschreven zijn met behulp van metadata en eventueel aanvullende documentatie dat ook andere onderzoekers de data kunnen begrijpen en gebruiken.

- Herbruikbaar

Dit criterium wordt uitgewerkt onder artikel 7.

- Duurzaam opgeslagen

Dit criterium wordt uitgewerkt onder artikel 8.



7. **Onderzoeksdata moeten worden opgeslagen in combinatie met de metadata, overige documentatie en eventueel de software(versie) die noodzakelijk zijn voor potentieel hergebruik van de data.**

Onderzoeksdata moeten op een zodanige manier worden opgeslagen dat ze onafhankelijk zijn van de onderliggende apparatuur/hardware, zoals microscopen, scanners of opnameapparatuur. Waar mogelijk wordt gekozen voor duurzame bestandsformaten die door data-archieven worden ondersteund. In sommige gevallen kan opslag van hardware overwogen worden; denk bijvoorbeeld aan software die alleen compatibel is met een verouderd computer operating system. Hierbij moet wel de kosten/baten verhouding in de gaten gehouden worden.

8. **Na afloop van het onderzoek moeten onderzoeksdata veilig en duurzaam worden opgeslagen, hetgeen betekent dat de integriteit, beschikbaarheid en – indien vereist - vertrouwelijkheid van de data gegarandeerd moeten blijven. De dataopslag dient plaats te vinden volgens internationaal geldende richtlijnen. Dit kan voor digitale data worden gerealiseerd door de data op te slaan in een *Trusted Digital Repository* (een ‘online depot’). Niet-digitale onderzoeksdata moeten worden opgeslagen volgens de gebruikelijke standaarden in de discipline.**

Een Trusted Digital Repository beschikt over apparatuur en programmatuur (hardware en software) voor de duurzame en betrouwbare bewaring van digitale informatie. Deze voorziening moet zijn ingebed in een organisatie die betrouwbaar is en bijvoorbeeld beschikt over een missie voor de lange termijn en lange termijnfinanciering.

Er zijn drie internationaal gangbare standaarden voor Trusted Digital Repositories en de manier waarop zij digitale data duurzaam opslaan. In volgorde van oplopende eisen aan de betrouwbaarheid zijn dat: het Data Seal of Approval¹, het Nestor Seal² (ook bekend als DIN 31644) en de RAC-standaarden³ (ook bekend als ISO 16363). Op dit moment is dataopslag bij een repository met een Data Seal of Approval (de laagste categorie) afdoende voor NWO en Horizon 2020.

Het verdient aanbeveling om niet-digitale data zo veel mogelijk te digitaliseren voor duurzame opslag.

9. **De minimale bewaartermijn van onderzoeksdata is tien jaar.**

Hiermee volgen we de Nederlandse Gedragscode Wetenschapsbeoefening van de VSNU (versie 2014).

B. Verantwoordelijkheden

10. **Het faculteitsbestuur is verantwoordelijk voor de nadere uitwerking van deze regeling en bepaalt op welk niveau deze uitwerking plaatsvindt: op facultair niveau en/of instituutsniveau.**

Deze regeling is een kader waarbinnen de faculteiten hun eigen beleid kunnen uitwerken. Soms zal het voor de hand liggen om dit alleen te doen op het niveau van de faculteit, soms alleen op het niveau van het instituut en soms zal een combinatie van beide het meest geëigend zijn. Denk bij deze laatste optie bijvoorbeeld aan een algemene facultaire regeling met daarbij een appendix specifiek voor elk instituut.

¹ <http://www.datasealofapproval.org>

² http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel_node.html

³ <http://www.iso16363.org/>



11. De wetenschappelijk directeur van het instituut is verantwoordelijk voor de juiste uitvoering van het beleid.

In de datamanagementplannen van de onderzoeksprojecten binnen het instituut wordt beschreven hoe het beleid in de praktijk wordt gebracht (zie hieronder).

12. In het DMP van een onderzoeksproject moet worden vastgelegd hoe de verantwoordelijkheden m.b.t. datamanagement belegd zijn tussen de verschillende medewerkers die betrokken zijn bij het project.

Bij veel onderzoeksprojecten zijn meerdere onderzoekers betrokken, bijvoorbeeld promovendi, postdocs en leden van de vaste staf. In het DMP wordt vastgelegd wie welke verantwoordelijkheid draagt. Speciale aandacht verdienen projecten waarbij onderzoekers van meerdere instituten en/of externe partijen betrokken zijn. Het is belangrijk om ook vast te leggen wat er met de data en de verantwoordelijkheid hiervoor zal gebeuren als een van de betrokken onderzoekers vertrekt.

13. Alle DMPs moeten centraal binnen de faculteit of het instituut worden opgeslagen, minimaal gedurende 20 jaar.

Let op: na het vernietigen van het DMP kan de onderzoeker of het instituut niet meer aantonen hoe bepaalde beslissingen m.b.t. het vernietigen van data tot stand zijn gekomen (uit het oogpunt van wetenschappelijke integriteit). De bewaartermijn van een DMP is langer dan die van de data, omdat dan in voorkomende gevallen met behulp van het DMP kan worden aangetoond dat het vernietigen van een bepaalde dataset (bijvoorbeeld na de standaard tien jaar) in lijn is met het DMP.

14. Op universitair niveau wordt een *data protection officer* aangewezen, die verantwoordelijk is voor de naleving van het beleid m.b.t. veiligheid van onderzoeksgegevens.

In het kader van Horizon 2020 kan bij honorering van onderzoeksvoorstellen worden gevraagd om een verklaring van de data protection officer dat binnen het project de 'technical data protection procedures' nageleefd worden. De data protection officer is daarmee verantwoordelijk voor de veiligheid van de onderzoeksgegevens gedurende het onderzoek.

15. Het College van Bestuur is verantwoordelijk voor het bieden van adequate voorzieningen en ondersteuning om verantwoord databeheer mogelijk te maken.

Dit houdt niet in dat alle voorzieningen en ondersteuning per se binnen de universiteit aangeboden moeten worden; waar nodig en zinvol kunnen afspraken worden gemaakt met (inter)nationale partijen die voorzieningen en/of ondersteuning op het gebied van datamanagement aanbieden.

C. Uitwerking

16. Bij de uitwerking van deze universitaire regeling dienen ten minste onderstaande aspecten geadresseerd te worden. Ook afwijkingen van deze regeling worden beargumenteerd vastgelegd. Bij de uitwerking van de regeling wordt aangesloten bij de normen die op dat moment gangbaar zijn in de discipline.

- Hoe de verantwoordelijkheden met betrekking tot datamanagement binnen het instituut zijn belegd;
- Wat de relevante data zijn om op te slaan, gezien het type onderzoek binnen het instituut;
- Wat het beleid is ten aanzien van data voortkomend uit bachelor-, master- en research master-onderwijs (in het bijzonder scripties);
- Hoe tijdens het onderzoek tegemoet gekomen wordt aan de eis dat de integriteit, beschikbaarheid en vertrouwelijkheid van de data gegarandeerd moeten blijven;



- Hoe na het onderzoek tegemoet gekomen wordt aan de eis dat de data op het moment van publicatie van het onderzoek vindbaar, toegankelijk, begrijpelijk, herbruikbaar en duurzaam opgeslagen zijn.
- Een eventuele maximale bewaartermijn van data en DMPs.

17. In het DMP van een onderzoeksproject komen minimaal de volgende aspecten aan de orde:

- Hoe de verantwoordelijkheden met betrekking tot datamanagement binnen het project zijn belegd en hoe wordt gehandeld bij het vertrek van (een van de) onderzoeker(s);
- Type(n) data die worden gegenereerd en verzameld (met format en omvang);
- Methode(n) van verzamelen of oorsprong van de data (met hard- en software);
- Standaarden en metadata voor de documentatie van de data (disciplinegericht en/of in lijn met de standaarden die zijn opgesteld door het repository waar de data worden opgeslagen);
- Waar de data tijdens het onderzoek worden opgeslagen en hoe de veiligheid en toegang zijn georganiseerd;
- Welke maatregelen worden genomen voor de duurzame opslag van de data en de toegang op de lange termijn;
- Wie er op welk moment toegang heeft tot de data;
- De omgang met privacygevoelige of anderszins vertrouwelijke data.

Voor het opstellen van een DMP kan de onderzoeker gebruik maken van de templates die subsidieverstrekkingen als NWO en de Europese Commissie hanteren. Ook binnen de universiteit is een dergelijk template ontwikkeld. In het DMP kan de onderzoeker deels verwijzen naar of putten uit het datamanagementbeleid van de eigen faculteit of instituut.

18. Alle hierboven genoemde artikelen kennen een implementatietermijn van drie jaar vanaf het moment van vaststelling, met uitzondering van de aanstelling van de data protection officer.

Het voorgestelde beleid heeft dermate veel impact dat het niet reëel is om dit op korte termijn verplicht in te voeren. Door een implementatieperiode te kiezen maken we ook ruimte voor faseverschillen tussen faculteiten en instituten. Dit laat onverlet dat met name onderzoekers die subsidieaanvragen willen indienen waarschijnlijk al op korte termijn aan een groot deel van dit beleid zullen moeten voldoen.

19. Deze regeling geldt niet met terugwerkende kracht.

Data verzameld in reeds afgeronde onderzoeksprojecten vallen dus in principe niet onder de regeling. Er is hier een grijs gebied als een nieuw project voortbouwt op eerder verzamelde data; er zal per geval moeten worden bekeken of het haalbaar is om ook de oude data duurzaam op te slaan.

20. Als daar aanleiding toe is, zal deze regeling gedurende bovengenoemde implementatiefase worden herzien door het College van Bestuur.

Bij deze herziening zullen ten minste het Onderzoeksberaad, de Regiegroep ICT en Onderzoek en het Platform Informatiemanagers worden geraadpleegd.

De Regeling Datamanagement is vastgesteld door het College van Bestuur op 19 april 2016.