



Een Visie op Datamanagement

Ron Dekker



Persoonlijke noot

Eigen ervaringen:

- goede kennis van de data nodig om alle valkuilen te omzeilen.
- het is veel werk van een ruwe dataset naar een onderzoeksbestand.
 - om de data te cleanen
 - verantwoordend middels codeboeken en prints van je programmatuur
 - een goede databeschrijving te geven voor nieuwe gebruikers
- het kost veel tijd om een vreemde dataset goed te leren kennen.
- het delen van kennis en van data leidt wederzijds tot meerwaarde.
- soms ontstaat er paradigm shift → heel anders gaan denken
 - bv. niet meer in termen van surveys denken, maar vanuit registraties
 - bv. van mens-machine interactie naar machine-machine interactie

- 
- The background of the slide is a photograph of a pond. In the foreground, there are several large, green lily pads floating on the water. In the middle ground, there are two large, cylindrical, metal water treatment tanks. The water is slightly murky. The sky is not visible.
- Het geeft weer hoe complex dataverzameling, verantwoording en hergebruik is.

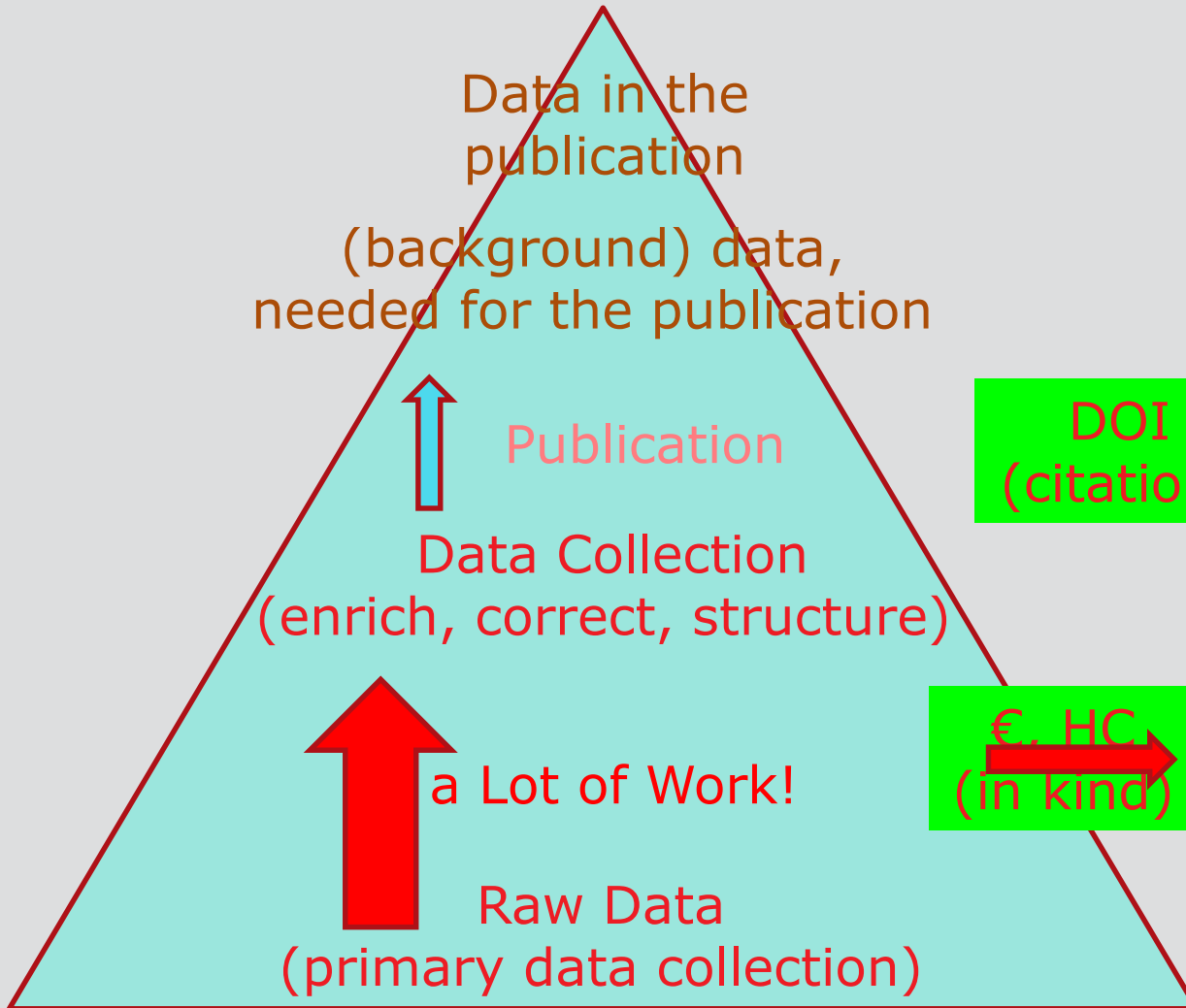


Over NWO

NWO

- Nationale wetenschapsfinancier
- Tweede geldstroom voor onderzoek
- Wetenschappelijk onderzoek laten excelleren
 - Competitie en selectie via *peer review*
 - Financiering van ruim 5.500 wetenschappers
 - Wetenschappelijke infrastructuur
- Kennisbenutting

Data Pyramid



Codes of Conduct, Integrity, Contracts

Publications by Others

Misuse

Data ready for re-use

Archive Catalogue

Metadata, access, software, apps, ...

Facilitate



Uitgangspunten

Datamanagement

- NWO beschouwt het met elkaar delen van onderzoeksdata als een gebruikelijke gang van zaken in het wetenschappelijk onderzoek;
- hergebruik van data wordt bevorderd door
 - goede openbare informatie over de data,
 - toepassing van gestandaardiseerde coderingssystemen en
 - koppeling met andere dataverzamelingen

Datamanagement

- NWO ziet
 - het verzamelen van data,
 - het bewerken en
 - het cureren van data tot en met
 - het deponerendaarvan als activiteiten die behoren tot het reguliere, wetenschappelijke onderzoek
- NWO is bereid die werkzaamheden te subsidiëren.



Implementatie

Datamanagement

- De NWO-regeling subsidies wordt explicieter gemaakt wat betreft omstandigheden voor 'restricted access' en embargo-periodes van dataopslag;
- In programmabrochures en subsidieaanvraagformulieren wordt een dataparaagraaf opgenomen die is aangepast aan het betreffende subsidie-instrument en de discipline;
- voor verdere uitwerking kan een datamanagementplan van toepassing zijn;

Datamanagement

- Datamanagement wordt meegenomen in het beoordelingstraject van subsidieaanvragen;
- Expertise op bureauniveau van datamanagement wordt bevorderd;
- Op alle beleidsniveaus wordt Open Access van data bevorderd;



Noblesse oblige

Datamanagement

- NWO eist maximale toegang tot wetenschappelijke data uit projecten die door NWO (mede) zijn gefinancierd, voor hergebruik door onderzoekers;
- NWO eist dat de door NWO gefinancierde onderzoekers omgaan met data conform de reeds bestaande gedragscodes;

Datamanagement

- NWO streeft er naar dat de data duurzaam worden opgeslagen, in een 'digital repository' met een internationaal aanvaard kwaliteitskeurmerk;
- Voor NWO geldt: Open Access tot data waar dat kan, 'restricted access' waar dat nodig is;

Datamanagement

- NWO stimuleert de ontwikkeling van een (inter-)nationaal netwerk van *trusted digital repositories*;
- NWO beschouwt SURF als meest aangewezen partij voor het leveren van de basisinfrastructuur voor dataopslag.
- NWO streeft naar een taakverdeling
 - NWO en instellingen als 'front office' in de rol van onderzoeksfinancier; systeemtoezicht
 - Intermediairs: 'ontzorgen', professionaliseren van de keten; Communities





En nu?

Datamanagement

- informatie over de databestanden
 - gebruikers- en toegangsrechten
 - bestandsformaten
 - toevoegen van databestanden
 - versiebeheer
 - onderzoeksfases
 - workflows
 - **beschrijvingen van de datasets**
 - intellectueel eigendom
 - samenwerking
 - gebruikte tools

Datamanagement

- overzetten naar een betrouwbare opslag
 - **terugvinden van datasets**
 - samenhang met publicaties
 - meetbaarheid van (her)gebruik
 - digitale duurzaamheid

Datamanagement

- toegang tot en het delen van de data
 - **gebruik van bestanden**
 - mogelijke openbare toegang
 - zoekmogelijkheden



Persoonlijke oproep

Kennis effectief inzetten

- Lokaal (inhoud) en
- Nationaal (data cycle)

Samenwerken

- Federatief
(vervolg op Wageningen 2013)

Standaarden

- Op zo een hoog mogelijk niveau
- Over disciplines en instellingen heen

Paradigm shift?

- Waarde van data
→ omgaan met data

