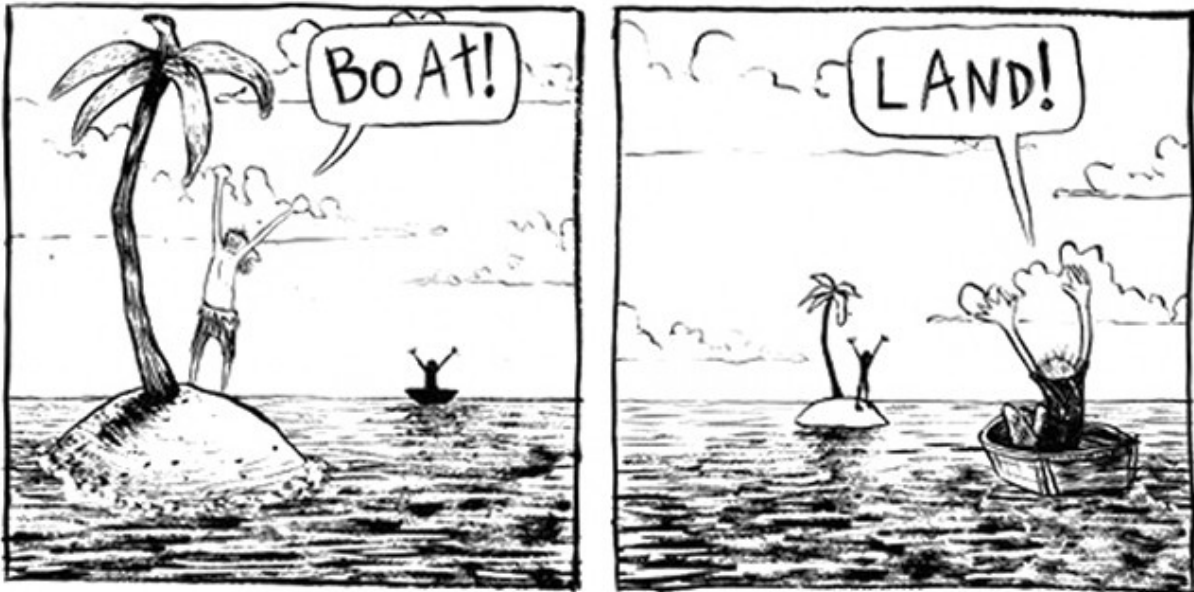


# Een dataperspectief, Update op Gegevens of Data



Perspective...

Published on December 30, 2019



Jef Bergsma

Enterprise Architect with a focus on data and information

## ***IT Een dataperspectief, Update op Gegevens of Data***

*Eerder definieerde ik 'data' als een digitale representatie van gegevens opgebouwd uit elementaire eenheden (zie <https://succesvol-met-it.nl/wp-content/uploads/2021/07/Gegevens-of-Data.pdf>). Dit in tegenstelling tot **gegevens** die ik definieerde als vastgelegde feitelikheden zoals feiten, meningen, metingen en registraties, ongeacht de verschijningsvorm. Uit de verschillende gesprekken die ik sinds de publicatie hierover heb gevoerd, blijkt dat er behoefte is aan een verduidelijking.*

Met deze **werkdefinities** heb ik geprobeerd duidelijk te maken dat bij de digitalisering van een bestaand proces de voordelen van datagericht werken niet vanzelfsprekend zijn. Het scannen van een formulier maakt nog niet dat je bruikbare data-elementen hebt, daar is

meer voor nodig. Vandaar dat ik data niet als digitaal definieer en daarbij heb aangegeven dat het uit **elementaire eenheden** moet bestaan.

Vanuit het **perspectief** van een softwareontwikkelaar of een data-scientist zijn waardes die in een computerbestand zitten gewoon data. Ook als dat bestand een video stream is, een foto in .jpg formaat of een pdf document. In lijn met mijn eerdere redenering kan deze data worden omgezet en gevisualiseerd als informatie voor een gebruiker. Is een computerbestand een elementaire eenheid en valt het als zodanig dan ook onder de werkdefinitie?

De 'waardevolle consequenties' die ik in de vorige **publicatie** aan de werkdefinitie heb gekoppeld, maken duidelijk dat er nog enige verduidelijking nodig is. De eerste consequentie die ik aangaf is: 'We gebruiken de term data alleen voor digitale gegevens, een formulier kan worden omgezet in data maar is op zichzelf geen data'. Daarmee wordt impliciet gezegd dat een bestand geen data is. En dat lijkt toch erg in tegenspraak met het werkveld van een softwareontwikkelaar of een data-scientist.

In de vorige publicatie is omwille van de eenvoud een aspect weggelaten, namelijk het **toepassingsperspectief**. De consequenties in de vorige publicatie zijn aangegeven vanuit het perspectief van de (eind)gebruiker in een digitaliseringstraject. Vanuit dat perspectief is de elementaire eenheid een enkelvoudige waarde van een gegeven, een waarde waar je bijvoorbeeld een eenheid, een syntax of een lijst met waarden aan kunt koppelen. Vanuit gebruikersperspectief zijn er voor datagericht werken twee soorten bestanden:

- bestanden die bestaan uit een **verzameling dataelementen** zoals een .csv bestand
- bestanden die als geheel een **object representeren** zoals een afbeelding, video, of document.

De eerste soort kan worden ingelezen als elementaire data-eenheden zoals ze in het bestand zitten. Het bestand is eigenlijk een soort transportmiddel, de envelop waarin de elementen zitten. De tweede soort bestaat uit bestanden die voor de gebruiker op zichzelf een element vormen dat als geheel wordt gebruikt.

Vanuit het perspectief van de softwareontwikkelaar of data-scientist ziet de wereld er iets anders uit. Aan hen de taak om uit de bestanden die onderdelen te selecteren die nodig zijn voor analyse en verwerking. Daarbij is een hexadecimale waarde, een bitpatronen in bestanden met afbeeldingen, video of formulieren zomaar een relevante elementaire eenheid waar ze mee werken en voldoet dit aan de werkdefinitie van data. Beide soorten bestanden bevatten een **digitale representatie van gegevens** en zijn opgebouwd uit (voor het toepassingsperspectief relevante) **elementaire eenheden**.

De definitie van '**data**' als een digitale representatie van gegevens opgebouwd uit elementaire eenheden moet misschien wel worden aangevuld tot: *een digitale representatie van gegevens opgebouwd uit elementaire eenheden die relevant zijn voor het toepassingsperspectief*.