

Kwaliteit van denken en de relatie met kwaliteit van data

door Danny Greefhorst

Er is een toenemende stroom aan data, informatie en kennis. De vraag is wat dit allemaal betekent. Sommigen zien het als ‘information overload’, terwijl anderen het juist zien als een vorm van zingeving – het stromen van data is een hoger liggend doel naar een internet van alle dingen.

Praktisch zullen we ons moeten aanpassen aan deze nieuwe werkelijkheid. Maximaal meebewegen met deze toenemende stroom heeft implicaties voor zowel onszelf als voor de datastroom die we genereren. Als we niet uitkijken verdrinken we in de stroom. Er is behoefte aan een hogere kwaliteit van denken alsook een hogere kwaliteit van data. In deze bijdrage verken ik deze kwaliteiten, hun relatie en het belang van context en ethiek.

De betekenis van de toenemende stroom

Als mens zullen we ons moeten aanpassen aan de toenemende beschikbaarheid en uitwisseling van data, informatie en kennis. Als we niets doen, worden we simpelweg overstroomd. Onze hoofden zullen in toenemende mate worden gevuld met allerlei indrukken, feiten, gevoelens en gedachten. Het zal steeds lastiger worden om dit soort zaken van elkaar te scheiden, gegeven de toename in omvang en snelheid van de stroom. Om

overeind te blijven zullen we ons vastgrijpen aan wat we snel kunnen bedenken en wat anderen ons vertellen. Of dat ook het meest verstandig is en aansluit bij wat we willen en hoe we ons tot anderen willen verhouden is maar de vraag. De kans is groot dat we suboptimale keuzes maken en gewoonweg meebewegen met anderen. De social media laten ons leven in filterbubbels, waarbij we alleen informatie te zien krijgen die is geselecteerd door het ingebouwde algoritme. We krijgen geen informatie te zien die ons eigen standpunt tegenspreekt. Populisme ligt op de loer en de moraliteit komt in het geding. Voor we het weten begeven we ons onbewust aan de verkeerde kant van de lijn en zorgen we voor grotere verschillen en doen we een verdere afbreuk aan onze fysieke leefomgeving (*zie ook* [Harari, 2017]).

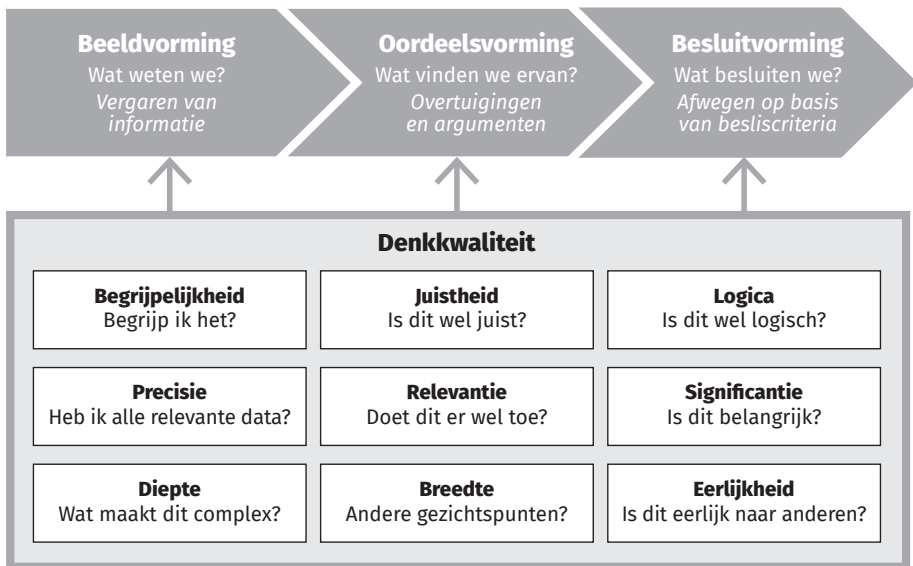
Onze basishouding zou daarom vooral kritisch moeten zijn. We moeten ons continu afvragen wat de stroom aan data en informatie precies betekent, feiten scheiden van gedachten en bewuste keuzes maken op basis van een logische redeneerlijn. Kritisch denken is een vaardigheid die handvatten biedt om de natuurlijke aanleg om zomaar iets te geloven te beteugelen en om verborgen aannames, onlogische redeneringen en denkfouten te herkennen en te voorkomen. Kritisch denken is waarschijnlijk al zo oud als de mensheid. Boeddha en Socrates gebruikten al kritische vragen om te bepalen of op autoriteit gebaseerde kennis rationeel, helder en logisch consistent kan worden gerechtvaardigd. Kritisch denken verduidelijkt doelen, onderzoekt aannames, onderscheidt verborgen waarden, evalueert het bewijs, volbrengt acties en beoordeelt de conclusies. Het vraagt bereidheid om je eigen redeneringen kritisch te evalueren en je overtuigingen te herzien wanneer de hiervoor noodzakelijke argumenten zich aandienen.

Kwaliteit van denken

Een kritische houding leidt tot een hogere kwaliteit van denken. Het zorgt ervoor dat we betere en meer transparante keuzes maken. Praktisch betekent het vooral dat je meer vragen moet stellen. Je moet vanuit allerlei perspectieven, aspecten en belangen kijken en reflecteren over wat mensen zeggen. Alles wat onduidelijk is, zal vervolgens leiden tot vragen. Het

stellen van vragen doe je als individu, maar daarnaast zal het ook vaak in groepsverband zijn, bijvoorbeeld in vergaderingen, bijeenkomsten of gesprekken. Het soort vragen is daarbij dan ook afhankelijk van de fase waarin een groepsproces verkeert. In algemene zin bestaat een groepsproces uit drie fasen: beeldvorming, oordeelsvorming en besluitvorming. De kernvragen die bij deze fasen horen zijn ‘wat weten we?’, ‘wat vinden we ervan?’ en ‘wat besluiten we?’.

De genoemde kernvragen kun je verder ontleden in de meer specifieke vragen zoals weergegeven in figuur 1. De vragen zijn in meer fasen relevant, maar in de figuur globaal gepositioneerd onder de meest relevante fase. Ze zijn een vertaling en samenvatting van de vragen zoals beschreven door Linda Elder en Richard Paul in hun boek over kritisch denken [Elder & Paul, 2013]. De auteurs hebben de vragen geclusterd naar kwaliteitsdimensies, die ze zelf ‘intellectuele standaarden’ noemen. Door het gestructureerd gebruik van deze kwaliteitsdimensies en het stellen van de juiste vragen ontstaat een bepaalde kwaliteit van denken. Een ander deel van de toolkit zoals beschreven door Elder en Paul bestaat uit een



.....
 Figuur 1: Denkkwaliteiten en hun rol in groepsprocessen

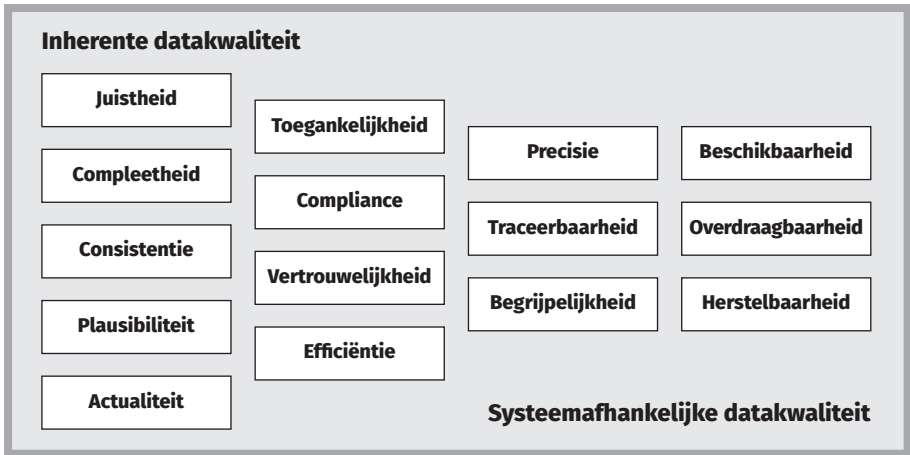
denkmodel dat verschillende aspecten van een vraagstuk belicht. In dat model maken de auteurs een onderscheid tussen doel, kernvraag, informatie, interpretatie, concepten, aannames, implicaties en gezichtspunt. Het expliciet onderscheiden van deze aspecten zorgt ervoor dat er een logische redeneerlijn ontstaat.

Kwaliteit van data

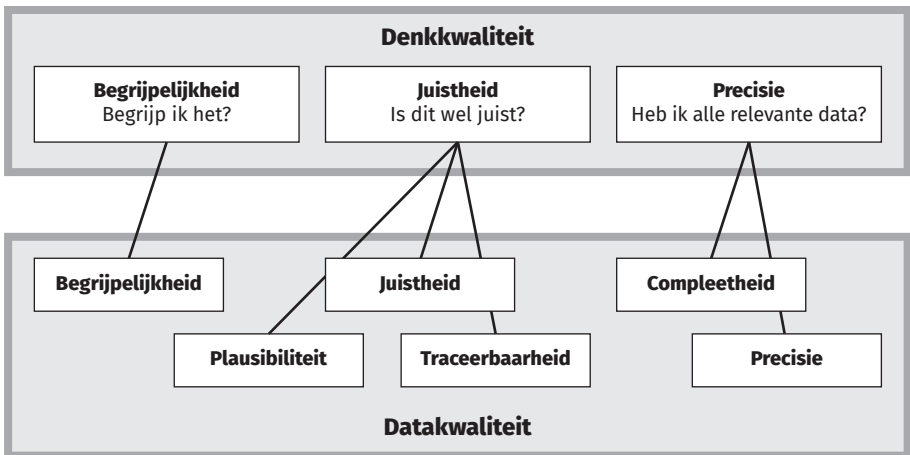
De mens heeft een eigen verantwoordelijk voor het omgaan met de toenevende stroom aan data en informatie, door het borgen van een bepaald niveau van denken. De hoeveelheden data en informatie zijn echter te groot om dit als mens allemaal te verwerken. Je kunt het als mens ook niet allemaal overzien. Het is daarom dan ook interessant om te zien dat denkkwaliteit een directe relatie heeft met datakwaliteit. Datakwaliteit is namelijk iets dat deels geautomatiseerd kan worden gecontroleerd, waardoor data met een lage kwaliteit deels automatisch kan worden uitgefilterd.

Datakwaliteit is de mate waarin data voldoen aan de behoeften vanuit het gebruik. Het gaat over dimensies zoals juistheid, compleetheid, consistentie, actualiteit, plausibiliteit, precisie en traceerbaarheid. Er zijn standaarden die je een lijst van kwaliteitsdimensies bieden, zoals de ISO/IEC 25012 [ISO, 2008] en 25024 standaarden (zie figuur 2). Zij geven ook aan hoe je datakwaliteit kunt meten door indicatoren te beschrijven. Indicatoren zijn concreet meetbare eigenschappen, waarmee je kwaliteit bijvoorbeeld ook in een dashboard of rapportage kunt laten zien. De standaarden maken verder nog een onderscheid tussen kwaliteitsdimensies die vooral over de data zelf gaan (inherente datakwaliteit) en kwaliteitsdimensies die meer gerelateerd zijn aan het systeem dat de data aanbiedt (systeem-afhankelijke datakwaliteit).

Het is interessant om te zien dat een aantal van de kwaliteitsdimensies voor data (grotendeels) overeenkomt met kwaliteitsdimensies voor denken (zie figuur 3). De overlap zit met name in begrijpelijkheid, juistheid en precisie. Begrijpelijkheid is de basis en gaat over de mate waarin we iets begrijpen. Het is de eerste fase van groepsprocessen. Als we iets



.....
 Figuur 2: Kwaliteitsdimensies in ISO/IEC 25012/25024



.....
 Figuur 3: Relatie tussen denkkwaliteit en datakwaliteit

begrijpen, dan kunnen we bepalen of iets waar is of niet – de juistheid van ons denken. Juistheid is de mate waarin iets overeenkomt met de werkelijkheid. Het zal duidelijk zijn dat dit een kernvraag is als het gaat over kwaliteit van denken. Als iets niet waar is, dan is het ruis (of fake news) en moeten we het gewoon negeren. Als je hier breder naar kijkt, dan zijn ook de datakwaliteitsdimensies plausibiliteit en traceerbaarheid hierbij van belang. Plausibiliteit is de mate waarin de data geloofwaardig zijn. Traceerbaarheid is de mate waarin de totstandkoming van data expliciet is gemaakt. De denkkwaliteit precisie gaat over de mate waarin alle relevante informatie om te kunnen oordelen aanwezig is. Dat is wat de datakwaliteitsdimensies compleetheid en precisie dekken.

Het belang van context

Een dieper inzicht in de relatie tussen kwaliteit van denken en kwaliteit van data is dat context allesbepalend is. Een deel van de context is het gebruik van de data. Het gebruik van de data bepaalt welke kwaliteit voldoende is. Bij het uitvoeren van bijvoorbeeld een gedetailleerde berekening heb je ook gedetailleerde data nodig (met een hoge precisie). Context is daarnaast ook alles wat je nodig hebt om iets te begrijpen. Daarbij geldt dat voor een geautomatiseerde bewerking vaak slechts een deel van deze context relevant is, namelijk dat deel dat relevant is voor de bewerking. Als mensen hebben we vaak de behoefte om beter te begrijpen wat iets betekent. We zijn geïnteresseerd in het onderzoeken van onze behoeftes en zoeken veel meer nuances dan computers. Dat is ook ons onderscheidend vermogen en de reden dat het nog lang zal duren voordat we kunnen worden vervangen door computers.

Het is niet eenvoudig om te benoemen wat allemaal relevant is voor mensen om iets goed te begrijpen. Het gaat waarschijnlijk om veel van wat de maker van de data wist, maar wat nooit expliciet is gemaakt of zal worden gemaakt [Woudenberg, 2002]. Een fundamentele vraag is dus of we zeker weten of we de data hebben begrepen zoals de maker het bedoeld heeft. Een deel van dit soort context kan expliciet worden gemaakt door metadata. Metadata kunnen ons in ieder geval een tekstuele uitleg geven van wat de maker bedoeld heeft met bepaalde woorden. Metadata

kunnen ook een ander belangrijk deel van de context inzichtelijk maken: de totstandkoming van de data (traceerbaarheid). Als we weten wie de data heeft gecreëerd, op welk moment in de tijd en welke bewerkingen het heeft ondergaan, dan hebben we al veel meer zicht op de context. Het kan ons vertrouwen in de nauwkeurigheid van de data verhogen.

Een ethisch perspectief

Een toenemend relevante vraag is in hoeverre data wel eerlijk zijn naar anderen. Of data iedereen gelijk behandelt. We leven nu eenmaal in een wereld waarbij mensen belangen hebben. Data zijn niet waarde vrij. Ze dragen de waarden en normen van de mensen die ze gecreëerd hebben. Het is daarom belangrijk om te begrijpen wat deze normen en waarden waren en of ze acceptabel zijn. Een ethisch perspectief is een expliciet onderdeel van kritisch denken zoals bijvoorbeeld beschreven door Elder en Paul [2013]. Daarnaast is er ook in de context van datamanagement toenemende aandacht voor ethiek. Zo is er in de Data Management Body of Knowledge een hoofdstuk geheel gewijd aan data-ethiek [DAMA, 2017]. Ethiek is inherent subjectief en zal contextueel beoordeeld moeten worden. Er zijn wel allerlei praktische handreikingen voor het omgaan met ethiek zoals de Ethische Data Assistent van de Utrecht Data School en het Ethiek-kompas van Kennisnet [Pijpers e.a., 2020]. De kern van dit laatste instrument is dat je de waarden benoemt die je zelf het belangrijkste vindt, argumenten voor en tegen verzamelt en deze weegt op basis van de waarden die je hebt benoemt.

Conclusies

De voorgaande paragrafen hebben vooral duidelijk proberen te maken wat de uitdagingen zijn bij het omgaan met een toenemende stroom aan data en informatie. De kernboodschap is dat zowel een hogere kwaliteit van denken als een hogere kwaliteit van data noodzakelijk is. Deze kwaliteiten hebben ook een directe relatie; een hogere datakwaliteit leidt ook tot een hogere denkkwaliteit. Voor beiden geldt dat context allesbepalend is.

Het dieperliggende doel van deze bijdrage als geheel is vooral om de lezer aan het denken te zetten. Het beoogt organisaties te onthaasten en ruimte te creëren voor de cruciale menselijke interactie die nodig is om elkaar te begrijpen, om te oordelen en om te besluiten. Uiteindelijk kunnen data en data-uitwisseling intersubjectieve betekenisgeving nooit vervangen. Communicatie is niet objectief en onze hersenen zijn niet zomaar te vervangen door geautomatiseerde beslisregels.

- Stimuleer een kritische denkhouding bij jezelf en bij anderen door het stellen van vragen.
- Besteed expliciete aandacht aan het definiëren en meten van de kwaliteit van data.
- Voorzie data van relevante context zodat anderen ze kunnen begrijpen en beoordelen.

Bronnen

- DAMA International (2017). Data Management Body of Knowledge, Technics Publications LLC.
- Elder, L. & Paul, R. (2013). *Critical Thinking – Tools for Taking Charge of Your Professional and Personal Life, 2nd edition*. New York: Pearson Education.
- Harari, Y.N. (2017). *Homo Deus: A Brief History of Tomorrow*. New York: Random House.
- ISO, (2008). ISO/IEC 25012 – Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model, first edition.
- Pijpers, R., Bomas, E., Dondorp L. & Ligthart, J. (2020). *Waarden wegen – Een ethisch perspectief op digitalisering in het onderwijs*. Rapport: Kennisnet.
- Woudenberg, R. van (2002). *Filosofie van taal en tekst*. Eindhoven: Damon vof.