

De digitalisering van de samenleving: Van techtopia naar sociologische nuance en vooruitgang

Bas Hofstra

Universitair docent, Vakgroep Sociologie/ICS, Radboud Universiteit, bas.hofstra@ru.nl

In ons werk, tijdens onze sociale interacties, en in onze vrije tijd zijn we in toenemende mate afhankelijk van digitalisering. Meer en meer geven we ons leven vorm met (online) digitale processen en informatie. De COVID-19-pandemie heeft laten zien hoe verregaand de Nederlandse samenleving is gedigitaliseerd. Het onderwijs werd van de ene op de andere dag digitaal en online gerealiseerd. Een groot deel van de beroepsbevolking werkte ineens vanaf een laptop op de keukentafel, en op diezelfde keukentafels stonden dikwijls meerdere laptops of “tablets” voor het eerdergenoemde digitale onderwijs van kinderen. De eerste screenshots met de vele webcamgezichten van collega’s en vrienden zo rond maart 2020 staan nu in ons collectieve geheugen gegrift.

Het tempo van deze digitalisering is hoog. Mocht de pandemie twee of zelfs één decennia eerder hebben toegeslagen, dan zouden deze plotselinge aanpassingen zich lastiger hebben voltrokken. In biologisch-epidemiologisch perspectief is één of twee decennia eerder of later bijna nietszeggend, maar met een digitaliseringsbril op is het een verschil van dag en nacht. Onze telefoons en laptops zijn nog niet bijzonder lang voorzien van hoge-resolutie camera’s, de eerste tablets kwamen slechts zo’n twaalf jaar geleden beschikbaar, en met de tragere internetsnelheden van minder dan tien jaar geleden zijn volle videoconferenties of zelfs simpel en toegankelijk videobellen lastig denkbaar. En stel, we gaan twee decennia terug: een populaire mobiele telefoon tijdens mijn middelbare schoolperiode had 64 miljoen keer minder opslagruimte dan mijn huidige telefoon. Dat ‘mobiele’ eigenlijk niet meer vóór telefoon komt, is veelzeggend.

Digitalisering gaat dus rap en digitalisering heeft implicaties voor hoe we leven en samenleven. In dit hoofdstuk reflecteer ik op sociologisch digitaliseringsonderzoek (in Nederland). Waarom is digitalisering in toenemende mate een cruciaal thema voor sociologen? Wat is digitalisering en hoe onderzochten én onderzoeken we het? En wat zijn de huidige bevindingen en toekomstige en ontwikkelingen binnen dit thema?

Onderzoek naar en met digitalisering

Nederland is van oudsher een digitale koploper. We horen bijvoorbeeld bij de Europese top als je kijkt naar het aantal huishoudens met internettoegang. In 2021 had 97 procent van de Nederlanders van 12 jaar of ouder thuis toegang tot snel internet (Centraal Bureau voor de Statistiek, 2021). Deze

digitalisering(sslag) heeft ook de interesse gewekt van (Nederlands) sociologisch onderzoek. Logischerwijs, aangezien veel sociale verschijnselen, soms terecht en soms onterecht, en soms direct en soms indirect, worden toegeschreven aan digitaliseringsdynamieken. Denk bijvoorbeeld aan de Arabische lente vanaf eind 2010, waarvan werd geclaimd dat snelle collectieve actie en massademonstraties mogelijk werden gemaakt door communicatie via sociale media – soms wordt de Arabische lente ook wel de “Facebook-revolutie” genoemd. Of denk, in Nederland, aan het plaatsje Haren in Groningen; een ongewild digitaal wijdverspreide verjaardagsuitnodiging veroorzaakte in 2012 tijdens ‘Project-X’ de zogenoemde ‘Facebook-rellen’ en miljoenen euro’s schade (Cohen et al. 2013).

Sinds ongeveer diezelfde tijd zijn ook begrippen als online nepnieuws, online echokamers, online polarisatie, en ‘big data’ onderdeel van ons collectieve vocabulaire. Recent bracht de COVID-19-pandemie sociologische vraagstukken gerelateerd aan digitalisering in beeld: welke gezinnen hadden bijvoorbeeld wel en welke niet de middelen om meerdere kinderen via meerdere laptops of tablets digitaal onderwijs te laten volgen, en vertaalt dat zich naar ongelijke onderwijsuitkomsten? En hoe zit dat met die 3 procent twaalfplussers die thuis geen toegang tot snel internet hebben?

Er zijn dus veel voorbeelden van sociale verschijnselen gelieerd aan digitalisering en digitalisering raakt sociologisch relevante thema’s zoals cultuur, sociale relaties en sociale ongelijkheid. Maar wat is digitalisering? Grofweg wordt digitalisering op twee manieren gedefinieerd. Enerzijds kun je het zien als het digitaliseren van analoge *processen*. Dit is nog steeds een vrij brede kapstok, waaraan je veel dynamieken kunt ophangen. Denk bijvoorbeeld aan het automatiseren van werkzaamheden die voorheen ‘met de hand’ gingen, zoals het werken met een digitale in plaats van een papieren planning, of aan het lezen van nieuwsberichten die Facebook aanbiedt op je ‘tijdlijn’ in plaats van het open slaan van een papieren krant.

Anderzijds kun je digitalisering ook iets enger definiëren als het digitaal maken van voorheen analoge *informatie*. Denk bijvoorbeeld aan het digitaliseren van oude papieren huwelijksakten (zie Knigge et al. 2014) of het inscannen van oude boeken (Google Books Library Project n.d.). Deze scheidslijn is niet altijd even strikt, en het één komt vaak voor het ander; vóór digitale *processen* (bijv. het aanklikken van een *Telegraaf*-bericht op Facebook) is digitale *informatie* cruciaal (bijv. het óók digitaal beschikbaar stellen van *Telegraaf*-artikelen).

Het is interessant dat het digitaliseringsonderzoek zelf ook min of meer gecategoriseerd kan worden met deze zelfde twee labels. Ik onderscheid bewust onderzoek *naar* digitalisering en onderzoek *met* digitalisering als twee richtingen in (Nederlands) sociologisch onderzoek. Dit zou je kunnen vatten onder de volgende twee onderzoeksrichtingen:

- a) Sociologisch onderzoek *naar* de oorzaken en gevolgen van digitaliseringsprocessen
- b) Sociologisch onderzoek *met* nieuwe gedigitaliseerde informatie en technieken

In beide richtingen is er een onderzoekstraditie, ook in Nederland, en zelfs met dit onderscheid blijven deze twee labels natuurlijk enorm vervlochten. Niet toevallig waren veel van de eerdergenoemde sociale verschijnselen in dit hoofdstuk gerelateerd aan sociale netwerken. Met wie gaan we om? Hoe geven we die sociale interacties vorm? Wat voor gevolgen heeft dit? Digitaliseringsprocessen en digitale informatie gaan tegenwoordig bovengemiddeld vaak hand in hand met nieuwe mogelijkheden tot (het onderzoeken van) sociale interacties. Om deze reden, en om enige focus aan te brengen in dit hoofdstuk, reflecteer ik in vogelvlucht met name op digitaliseringsonderzoek door de lens van één van mijn eigen expertises: het internet en sociale netwerken.¹

Onderzoek *naar* digitalisering: Nuance boven technopaniek en techtopia

Populaire ideeën over nieuwe technologieën en hun kwalijke of positieve invloeden op ons (sociale) leven kennen een rijke historie. In de Middeleeuwen dacht men bijvoorbeeld dat de ontwikkeling van boekdrukkunst de religieuze leer zou doen eroderen. Later bestond de vrees dat de telegraaf te snel te veel slecht nieuws zou verspreiden, en men zou weleens mentaal ongezondheid kunnen worden van fietsen. Meer recent zouden de elpee, computerspelletjes en de televisie ‘een slechte invloed hebben’, vaak op de jeugd.

Maar wat dan die slechte invloed exact is en wat de precieze beïnvloedingsmechanismen tussen een bepaalde technologie naar een ongewenste uitkomst zijn, blijft vaak onderbelicht. Een rol van sociologen was en is om deze beïnvloedingsmechanismen eerst te expliciteren en dan te confronteren met de empirie om vast te stellen of technologie echt ten koste gaat van sociale welvaart. Een treffend voorbeeld is bijvoorbeeld de vraag wat het causale, verklarende micromechanisme tussen bijvoorbeeld televisiekijken en een uitkomst, zeg, het hebben en onderhouden van sociale contacten is. In dit verband was er een algemene intuïtie van socioloog Robert Putnam (1995). Hij focust op de rol van televisiekijken bij een tweedeling tussen mensen geboren voor en na de Tweede Wereldoorlog. De oudere cohorten bezaten meer sociaal kapitaal dan de jongere cohorten, en hiervoor zoekt hij een verklaring in televisiekijken. De jongere cohorten groeiden op met de televisie in de huiskamer, de oudere niet. Tijd voor de televisie zou tijd wegnemen bij onderhouden van echte “face-to-face” sociale interacties (“time displacement”).

¹ Met sociale netwerken bedoel ik hier daadwerkelijke sociale relaties, die kunnen zowel plaatsvinden op sociale media als “offline”.

Zodoende zouden vanwege televisiekijken jongere cohorten minder sociaal kapitaal bezaten dan de oudere cohorten. Dit viel echter lastig los te zien van het mechanisme dat de televisie die ook nog een 'slecht wereldbeeld' veroorzaakt volgens Putnam: blootstelling aan slecht nieuws en geweld op de televisie maakt dat men minder sociale contacten opzoekt. Welke van de twee mechanismes – tijdsrestricties of het gemene wereldbeeld? – het juiste was, werd niet helemaal duidelijk. Dit is een voorbeeld van hoe technologie vaak wordt gezien als bedreiging voor sociale cohesie, maar toetsbare hypothesen over technologie of digitalisering afgeleid uit helder uitgelegde micromechanismen zijn en blijven daarbij vaak schaars.

Over de invloed van digitalisering en internet op sociale contacten werd aanvankelijk niet heel anders gedacht. Een sprekend voorbeeld van 'slechte' digitaliseringsuitkomsten met betrekking tot sociale netwerken is het werk van Kraut en collega's (1998) naar het vroege internet, zo rond 1995. Hun bevinding was dat internetgebruik negatief samenhangt met sociale contacten. Ze suggereerden dat tijd gependend op het internet andere sociale activiteiten verstoort (zie de overeenkomst met Putnam's claim) en dat internetgebruik sterke sociale relaties minder belangrijk maakt. Dit werd ook wel de 'internet paradox' genoemd: een nieuwe 'sociale' technologie zou ons juist minder sociaal maken.

Later werd deze bevinding opmerkelijk genoeg herzien door hetzelfde onderzoeksteam: een aantal jaar later vonden Kraut en collega's (2002) dat internetgebruik *positief* samenhangt met sociale contacten. De suggestie was dat in de tijd tussen beide onderzoeken veel meer mensen internetgebruiker werden, een verviervoudiging van 1995 tot 1999, en dat het daarom juist makkelijker werd om contacten online te onderhouden. Deze conclusie ondersteunt mijn eerdere claim over hoe dynamisch digitalisering is: een verspringende correlatie van negatief naar positief in slechts vier jaar tijd.

Bovenstaande tegengestelde bevindingen zijn goed te labelen onder wat ook wel 'dystopische' en 'utopische' blikken op het internet genoemd werden (zie Wellman 1997). De dystopisten zagen digitalisering en het (sociale) internet vooral als bedreiging voor 'echt' contact: via het internet gaat non-verbale communicatie verloren, is het contact maar oppervlakkig, en missen we kansen om te participeren in sociale activiteiten. Het gevolg is minder of in elk geval minder waardevol contact. De utopisten zagen digitalisering en het internet juist als een kans: in elke niche en voor elke interesse zijn er wel verwanten te vinden en je kunt omgaan met wie je wilt zonder enige vorm van kansenrestricties. Het gevolg is meer of in elk geval meer waardevol contact.

Als we de argumenten van beide kanten nader bekijken, zien we gangbare sociologische concepten over netwerkformatie. Dystopisten zien contactkansen als essentieel, terwijl utopisten persoonlijke voorkeuren benadrukken. In de context van het hedendaagse sociale internet zouden

beiden echter kritisch moeten zijn over hun eigen verwachtingen: louter contact op basis van voorkeuren leidt tot sterke homogeniteit en mogelijke polarisatie, terwijl er tegelijkertijd meer mogelijkheden dan ooit zijn voor contact. Wat we weten uit bestaand netwerkonderzoek is dat contactkansen en contactvoorkeuren niet opereren zonder elkaar, dus de hypothese-ontwikkeling moet genuanceerder en preciezer dan dergelijke ideeën.

Er is dus vaak niet één effect van digitalisering of 'het internet' op onze sociale contacten, en het theoretisch uitwerken van één zo'n mythe stuit al snel op theoretische tegenstellingen. Hier zien we dus weer een belangrijke rol in het specificeren van mechanismen en condities door sociologen. Het gaat niet over of je bepaalde facetten van het internet gebruikt zoals mythes vaak claimen ("Facebookgebruik veroorzaakt polarisatie!"), maar hoe je de vele verschillende facetten van het sociale internet in je dagelijkse leven integreert en hoe dat samenhangt met - en misschien andere processen versterkt in relatie tot - sociale contacten.

Van Deursen en Van Dijk (2013) stelden deze denkomslag ook vast: niet *of* maar *hoe* je het internet gebruikt hangt samen met sociologisch relevante uitkomsten. Hoe je het internet gebruikt kun je zien als 'digitale geletterdheid' en dat kan een bron zijn van sociale ongelijkheid. Nederlands sociologisch onderzoek kent hierin de laatste jaren terecht een groeiende en genuanceerde traditie. Een selectieve greep uit recente studies: Bayer en Hofstra legden uit dat het gebruik van de verschillende opties die sociale mediaplatforms bieden voor contact invloed heeft op het onderhouden van 'echte' sociale contacten (2020), Van Ingen en Matzat vonden dat de mobilisering van online hulp na vervelende gebeurtenissen gemakkelijker is voor theoretisch dan voor praktisch opgeleiden (2018), Spierings en collega's lieten zien dat politici op Twitter vaker onderling of met media dan met 'normale' mensen contact hebben (2019), Norbutas vond dat online drugstransacties met name binnen en niet tussen landen plaatsvindt (2018), Menshikova en Van Tubergen lieten zien dat anti-immigratie-opvattingen op Twitter talrijker zijn na berichten over immigratie in de media (2022), Kas en collega's vonden dat internationaal contact via AirBnB het vertrouwen tussen mensen uit verschillende landen niet bewijsbaar beïnvloedt (2022), en Hofstra en collega's lieten zien dat vrienden vaak online privacy instellingen van elkaar overnemen (2016).

Bovenstaande voorbeelden betreffen allemaal vrij recent Nederlands sociologisch onderzoek naar oorzaken en gevolgen van digitaliseringsprocessen, in het bijzonder over het sociale internet en sociale contacten. Onderzoek dat niet één digitaliseringseffect claimt, maar de waarde toont van het toetsen van populaire mythes aan verduidelijkte mechanismen en feitelijke gegevens. Daar was en is een rol voor de Nederlandse sociologen weggelegd, in het bijzonder in relatie tot vaak uitgesproken claims in het publieke debat over digitaliseringsprocessen.

Een uitdaging hier lijkt toegang voor wetenschappers tot die digitaliseringsprocessen. Het sociale internet en sociale mediaplatforms worden complexer, en de rol van algoritmes – in online aankopen, in online kijkgedrag, in vriendschapskeuzes, in toegang tot nieuws – krijgt steeds meer aandacht in maatschappelijke discussies. Maar om te begrijpen en te toetsen hoe dergelijke algoritmes mogelijk invloed hebben op ons gedrag, moeten we weten hoe zij exact in elkaar steken (om condities te verhelderen en verwachtingen te genereren over beïnvloedingsprocessen) en moeten er data over blootstelling aan die algoritmes en gedragssignalen ontsloten worden voor sociologen (om empirisch te toetsen wat mogelijke invloeden zijn). In beide richtingen kan vooruitgang geboekt worden; in de huidige situatie hebben technologiebedrijven en een select groepje wetenschappers (vaak verbonden aan prestigieuze Amerikaanse universiteiten) – mede door eigendom van platformen, data, en algoritmes – veel meer grootschalig *en* gedetailleerd inzicht in de beïnvloedingsmechanismen van ons (online) gedrag en sociale relaties dan de meeste sociale wetenschappers.

Dit roept vragen op. Is het wenselijk als het baanbrekendste netwerkonderzoek door technologiebedrijven uitgevoerd en gepubliceerd wordt in onze toptijdschriften? Wat als deze data niet ontsloten worden en dit onderzoek niet gerepliceerd kan worden? En dan zijn er nog tal van ethische vragen omtrent digitalisering, data-eigenaarschap, privacy, en de claim op het gebruik van dergelijke data door private en overheidsinstanties. Deze ethische vragen gaan verder dan de reikwijdte van dit essay, maar verdienen blijvende aandacht. Is het bijvoorbeeld wenselijk als gemeenten met nepaccounts op Facebook meekijken om verhalen van asielzoeker te verifiëren of falsifiëren (Volkskrant 2021)? En is de strijd om iemands online tijd en aandacht wel een gelijk speelveld wanneer miljarden dollar kostende supercomputers het sociale internet verregaand verpersoonlijken op basis van duizenden variabelen per individu?

Onderzoek *met* digitalisering: De (constante) digitale revolutie in de sociologie

Naast onderzoek *naar* digitaliseringsprocessen vindt er ook onderzoek *met* digitale informatiebronnen en technieken plaats. Waar het onderzoek *naar* vaak digitalisering als verschijnsel op zichzelf bestudeert, focust het onderzoek *met* voornamelijk op het stellen en beantwoorden van sociologische vragen die lastig zijn te beantwoorden zonder digitale informatie en geavanceerde data-analysetechnieken. Onderzoek *met* digitale informatiebronnen en -technieken kun je op twee manieren benaderen.

De eerste benadering behelst een trend sinds de jaren zestig. “*De grote hoeveelheid gegevens maakte mechanisatie van de bewerking noodzakelijk*” (1963: 345). Dit is wat Ivan Gadourek stelde toen hij de Nederlandse sociologie liet kennismaken met de digitale analyse van

vragenlijstgegevens uit een grote steekproef. De zorgvuldigheid van de analyses, de precisie waarmee hij deze beschreef, en de aansluiting met de iets volwassenere Amerikaanse kwantitatieve sociologie vond vruchtbare bodem in Nederland. Via en mede door zijn promovendus, en mede-oprichter van het *ICS* in 1986, Reinhard Wippler is de analyse van digitale gegevens sindsdien niet meer weg te denken uit Nederlandse sociologieprogramma's. Met recht een (digitale) revolutie in de Nederlandse sociologische traditie. Niettemin omdat onder Wippler's promovendi zich maar liefst negentien hoogleraren bevinden (Raub 2021). En de sociologen uit deze school die (voornamelijk) digitale of digitaal-gemaakte vragenlijstgegevens digitaal analyseren zijn talrijk: mijn proefschrift was de 251^{ste} uit de *ICS*-stal, en de steller staat momenteel al boven de 300.

De tweede benadering van onderzoek *met* digitalisering vloeit hieruit voort, ondanks dat het in de literatuur vaak als een nieuwe ontwikkeling wordt besproken. De eerst-natuurkundige-toen-socioloog Duncan Watts noemt deze tweede zienswijze zelfs expliciet een 'digitale revolutie' in de sociale wetenschappen (2011). Hij en anderen (Lazer et al. 2009) claimen dat door een toenemende beschikbaarheid aan digitale data en rekenkracht we steeds beter zowel grootschalig als gedetailleerd inzicht in menselijk gedrag krijgen. Waar Gadourek's analyses in 1963 nog met ponskaarten en een enkele centrale computer ergens op de campus verliepen kun je nu statistische analyses aanvuren met een telefoon terwijl je in de trein zit. Met name digitale voetafdrukken – denk aan berichtjes op Twitter, of aan 'check-ins' op bepaalde locaties, of aan forumberichten – zijn hierin volgens Watts en anderen – disclaimer, hier reken ik mezelf ook toe – leidend.

Deze tak van sociologisch onderzoek is gelieerd aan (misschien zelfs synoniem voor) het zogenaamde 'computational social science' of 'computational sociology'. Deze alsmaar groeiende richting definieer ik dikwijls pragmatisch als sociologische onderzoeks-aanpak: probleem-gedreven, theoretisch-empirische sociologie *waarin het empirische gedeelte nieuwe datatypes en/of methoden gebruikt anders dan 'traditionele' vragenlijsten of vragenlijstmethodiek*. De laatste schuingedrukte zinsnede is hierin de leidende afwijkende factor ten opzichte van de 'reguliere' kwantitatieve sociologie. Dit betekent vaak, maar niet altijd, onderzoek met behulp van data die groter, anders, en gevarieerder zijn dan vragenlijstinformatie – oftewel 'big data', zie Corten (2019) voor een gedetailleerde bespreking van definities omtrent deze term –, en behelst vaak, maar niet altijd, dat de digitale voetafdruk op het internet gezien wordt als een reflectie van iemands offline gedrag. Met andere woorden: door iemands digitale voetafdrukken onder de loep te nemen (bijvoorbeeld iemands Twitervrienden en onderlinge berichten over COVID-19-vaccinaties) leren we iets substantieels over de persoon in kwestie (bijvoorbeeld de verandering van denkbeelden over vaccinatie door sociale contacten).

Door dit op grote schaal te doen – soms zelfs op populatieniveau – kunnen we sociale verschijnselen ontwarren die voorheen lastig te bestuderen waren met ‘regulier’ vragenlijstonderzoek. Argumenten hiervoor zijn dat sommige meningen of gedragingen door sociale wenselijkheid in vragenlijsten lastiger te meten zijn dan gedragssignalen gevonden in digitale sporen, en dat met name sociale netwerk- en verspreidingsdynamieken vaak lastig te bestuderen zijn met vragenlijstonderzoek.

Deze nieuwe benadering van onderzoek *met* digitalisering in combinatie met de sneltreinvaart waarmee het sociale internet zich ontwikkelde is in het laatste decennium één van de vruchtbaarste richtingen binnen het (Nederlandse) sociologie en zelfs sociaalwetenschappelijke landschap gebleken. De integratie van probleem-gestuurde sociologische vragen met nieuwe en gevarieerde digitale informatie en technieken heeft nieuwe inzichten gegeven in Nederlandse sociale structuren. Zo hebben we bijvoorbeeld geleerd met behulp van Hyves-netwerken of en hoe sociale netwerken op gemeenteniveau relateren aan economische activiteit binnen die gemeenten (Norbutas en Corten 2018). Ander onderzoek met Hyves liet zien hoe een groot gedeelte van de Nederlandse bevolking binnen een ‘kleine wereld’ leeft, dat wil zeggen: men is via slechts enkele handdrukken verwijderd van vele anderen (de bekende sociologische frase “six degrees of separation”), terwijl er tegelijkertijd veel groepsvorming is binnen het netwerk van Nederlanders (Corten 2012). Cruciale informatie, gegeven dat bijvoorbeeld ziektes zich makkelijker kunnen verspreiden in netwerken met deze eigenschappen. (Hier blijkt de beperkte houdbaarheid van digitaliseringsonderzoek en de flexibiliteit die sociologen in dit onderzoeksveld zich eigen moeten maken; anno 2022 zou dergelijk onderzoek met het inmiddels ter ziele zijnde Hyves onmogelijk zijn.) Ook is er onderzoek van ondergetekende onder Nederlandse jongeren met behulp van Facebookinformatie dat laat zien waarom kennissen vaak dezelfde etniciteit hebben (Hofstra et al. 2017) en hoe dat vervolgens samenhangt met etnische vooroordelen (Hofstra 2022). En onderzoek onder honderdduizenden wetenschappers laat zien hoe vrouwen en minderheden beter gedijen in organisaties als zij mentoren hebben die op hen lijken (Hofstra 2022).

Ook digitalisering in de vorm nieuwe onderzoekstechnieken om deze nieuwe datatypes te bestuderen heeft deuren geopend (wederom, afwijkend van vragenlijstmethodiek). Denk aan het ontwikkelen van automatische tekstanalyse die knipbijvoorbeeld signalen over onderwijsachtergrond kan oppikken (Corten et al. 2021) of de toepassing van tekstanalyse voor het meten van sentiment in geschreven tekst (Menshikova en Van Tubergen 2022).

Onderzoek *met* dit type nieuwe digitale sporen gaat overigens lang niet altijd over het internet of sociale netwerken. Onderzoek dat gebruikmaakt van het grootschalig digitaliseren van eeuwenoude huwelijksakten laat bijvoorbeeld zien hoe afkomst steeds minder belangrijk werd voor

statusverwerving in het 19^{de}-eeuwse Nederland (Knigge et al. 2014). Weer ander onderzoek met digitale informatie over Nederlandse wetenschappers laat zien hoe het Nederlandse wetenschapssysteem onderhevig is aan zogenaamde Mattheus-effecten (Bol et al. 2018) waar soms (redelijk toevallig) succes wordt beloond met nog meer succes. En er is Nederlands onderzoek met Amerikaanse gegevens over gender- en rasverschillen bij het opsporen van vermiste kinderen met digitale politierapporten (Van de Rijt et al. 2018).

Dit is slechts een selectieve greep uit dit onderzoek, maar stuk voor stuk beantwoorden deze studies belangrijke sociologische vragen met behulp van informatie die beschikbaar is gekomen door digitalisering(stechnieken), maar *zonder* vragenlijsten. Ondanks deze succesvol gebleken onderzoeksaanpak zijn er ook hier uitdagingen. In onderzoek met nieuwe digitale informatiebronnen (van het internet) zijn sociologische probleemstellingen, theorieën, verwachtingen en sociologenkennis over steekproefbias en statistische inferentie vaak een essentiële toevoeging binnen interdisciplinaire onderzoeksteams. Echter, om een rol van betekenis te blijven spelen in dit nieuwe en groeiende veld zal er vooruitgang geboekt moeten worden op een aantal vlakken.

Ten eerste is een heroverweging nodig van het sociologisch methodencurriculum. Onderzoek met nieuwe, grote digitale informatiebronnen vereist vaak het bewerken en analyseren van grote hoeveelheden (denk aan miljoenen rijen) ongestructureerde data (denk aan teksten). Dit vergt andere programmeervaardigheden – zoals parallelrekenen, tekstanalyse of ‘machine learning’ – dan het bewerken en analyseren van variabelen en observaties in vragenlijstgegevens. Maar dit is iets waar de gemiddelde sociologiestudent in het huidige methodenonderwijs nu nog amper mee in aanraking komt. Deze vaardigheden zijn ook niet louter van belang voor aspirant wetenschappers. Om afgestudeerde sociologen op te leiden tot waardevolle toevoeging binnen in toenemende mate datagedreven organisaties, zullen zij kennis moeten hebben – ofwel als uitvoerder ofwel als verbinder tussen organisatielagen – van dergelijke datatypen en -mogelijkheden. Mijn indruk is dat we internationaal en in vergelijking met andere disciplines in Nederland achterlopen in (invloed op) sommige initiatieven, zoals recente brede studies *computational social science*. Een consequente sociologische vinger in de pap is daar cruciaal.

Ten tweede zijn blijvende investeringen in onderzoekinfrastructuur en -ondersteuning nodig. Nieuwe digitale informatie is vaak groot en complex en het bewerken daarvan op, zeg, een doorsnee werklaptop blijkt soms lastig. Samenwerken in grotere interdisciplinaire teams en toegang tot en ondersteuning bij supercomputers – en er zijn momenteel al de nodige platforms – zijn daarbij essentieel. Ook het linken van waardevolle en gedetailleerde vragenlijst- of registerdata aan nieuwe datatypen lijken vruchtbare routes in bovenstaande onderzoeksvoorbeelden. Het toegankelijk kunnen koppelen van die digitale sporen aan bijvoorbeeld vragenlijstgegevens vergt het

aanpassen van werkprocessen naar een veilige en tegelijkertijd toegankelijke onderzoeksinfrastructuur.

Vooruitgang

Een zogenoemde 'digitale revolutie' (naar Watts 2011) in sociaalwetenschappelijk onderzoek had de Nederlandse sociologie al bereikt. De nieuwe digitale revolutie lijkt een voortvloeisel uit de digitalisering en kwantificering van de Nederlandse sociologie die is ingezet sinds de jaren zestig, maar is succesvol en blijft veelbelovend. Ook onderzoek *naar* digitalisering kent een genuanceerde traditie in de Nederlandse sociologie. Bovenstaande tekst geeft alle aanleiding om in deze richtingen te blijven ontwikkelen. Er zijn echter uitdagingen. En nadenken over het methodenonderwijs ten behoeve van een nieuwe generatie sociologen, en investeren in toegankelijke onderzoekinfrastructuur voor innovatief sociologisch digitaliseringsonderzoek dat belangrijke kennis oplevert over sociale relaties, ongelijkheid, en cultuur in Nederlandse blijft essentieel.

Referenties

- Bayer, J. B., & Hofstra, B. (2020). Toward curation and personality-driven social networks. *Nature Human Behaviour*, 4(2), 123-125.
- Bol, T., de Vaan, M., & van de Rijt, A. (2018). The Matthew effect in science funding. *Proceedings of the National Academy of Sciences*, 115(19), 4887-4890.
- Centraal Bureau voor de Statistiek. (2021). Internettoegang en internetactiviteiten; persoonskenmerken. Geraadpleegd op 25 mei 2022 via <https://www.cbs.nl/nl-nl/cijfers/detail/84888NED>.
- Cohen, M. J., Brink, G. J. M., Adang, O. M. J., van Dijk, J. A., & Boeschoten, T. (2013). Twee werelden: You only live once. Geraadpleegd op 25 mei via <https://research.utwente.nl/en/publications/twee-werelden-you-only-live-once>.
- Corten, R. (2012). Composition and structure of a large online social network in the Netherlands. *PLoS One*, 7(4), e34760.
- Corten, R. (2019). Big data in sociologisch onderzoek. *Sociologie*, 15(1), 45-68.
- Corten, R., Nadi, S., & Frank, L. (2021). Predicting educational background using text mining. Geraadpleegd op 1 juni 2022 via <https://osf.io/preprints/socarxiv/rds8z/>.
- Van Deursen, A. J., & Van Dijk, J. A. (2014). The digital divide shifts to differences in usage. *New Media & Society*, 16(3), 507-526.
- Gadourek, I. (1963). *Riskante gewoonten en zorg voor eigen welzijn*. J.B. Wolters: Groningen.

- Google Books Library Project. (n.d.). Google Books Library Project. Geraadpleegd op 1 juni 2022 via <https://books.google.com/intl/en-GB/googlebooks/library.html>.
- Hofstra, B., Corten, R., & Van Tubergen, F. (2016). Understanding the privacy behavior of adolescents on Facebook: The role of peers, popularity and trust. *Computers in Human Behavior*, 60, 611-621.
- Hofstra, B., Corten, R., Van Tubergen, F., & Ellison, N. B. (2017). Sources of segregation in social networks: A novel approach using Facebook. *American Sociological Review*, 82(3), 625-656.
- Hofstra, B. (2022). Interethnic weak ties online and out-group attitudes among Dutch ethnic majority adolescents. *European Societies*, 1-30.
- Hofstra, B., McFarland, D. A., Smith, S., & Jurgens, D. (2022). Diversifying the professoriate. *Socius*, 8, 23780231221085118.
- Van Ingen, E., & Matzat, U. (2018). Inequality in mobilizing online help after a negative life event: the role of education, digital skills, and capital-enhancing Internet use. *Information, Communication & Society*, 21(4), 481-498.
- Kas, J., Delnoij, J., Corten, R., & Parigi, P. (2022). Trust spillovers in the sharing economy: Does international Airbnb experience foster cross-national trust? *Journal of Consumer Behaviour*, 21(3), 509-522.
- Knigge, A., Maas, I., Van Leeuwen, M. H., & Mandemakers, K. (2014). Status attainment of siblings during modernization. *American Sociological Review*, 79(3), 549-574.
- Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukophadhyay, T., & Scherlis, W. (1998). Internet paradox: A social technology that reduces social involvement and psychological well-being? *American Psychologist*, 53(9), 1017.
- Kraut, R., Kiesler, S., Boneva, B., Cummings, J., Helgeson, V., & Crawford, A. (2002). Internet paradox revisited. *Journal of Social Issues*, 58(1), 49-74.
- Lazer, D., et al. (2009). Computational social science. *Science*, 323(5915), 721-723.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 415-444.
- Menshikova, A., & van Tubergen, F. (2022). What drives anti-immigrant sentiments online? A novel approach using Twitter. *European Sociological Review*, forthcoming.
- Norbutas, L. (2018). Offline constraints in online drug marketplaces: An exploratory analysis of a cryptomarket trade network. *International Journal of Drug Policy*, 56, 92-100.
- Norbutas, L., & Corten, R. (2018). Network structure and economic prosperity in municipalities: A large-scale test of social capital theory using social media data. *Social Networks*, 52, 120-134.

- Putnam, R.D. (1995). Tuning in, tuning out: The strange disappearance of social capital in america. *Ps Political Science and Politics*, 28(4), 664-683.
- Raub, W. (2021). In memoriam: Reinhard Wippler (1931–2020). Geraadpleegd op 1 augustus 2022 via <https://www.uu.nl/en/organisation/sociology/reinhard-wippler>.
- Van de Rijt, A., Song, H. G., Shor, E., & Burroway, R. (2018). Racial and gender differences in missing children's recovery chances. *PLoS one*, 13(12), e0207742.
- Spierings, N., Jacobs, K., & Linders, N. (2019). Keeping an eye on the people: Who has access to MPs on Twitter? *Social Science Computer Review*, 37(2), 160-177.
- Watts, D. J. (2011). *Everything is obvious: * Once you know the answer*. Currency: Sydney.
- Wellman, B. (1997). The road to utopia and dystopia on the information highway. *Contemporary Sociology*, 26(4), 445-449.