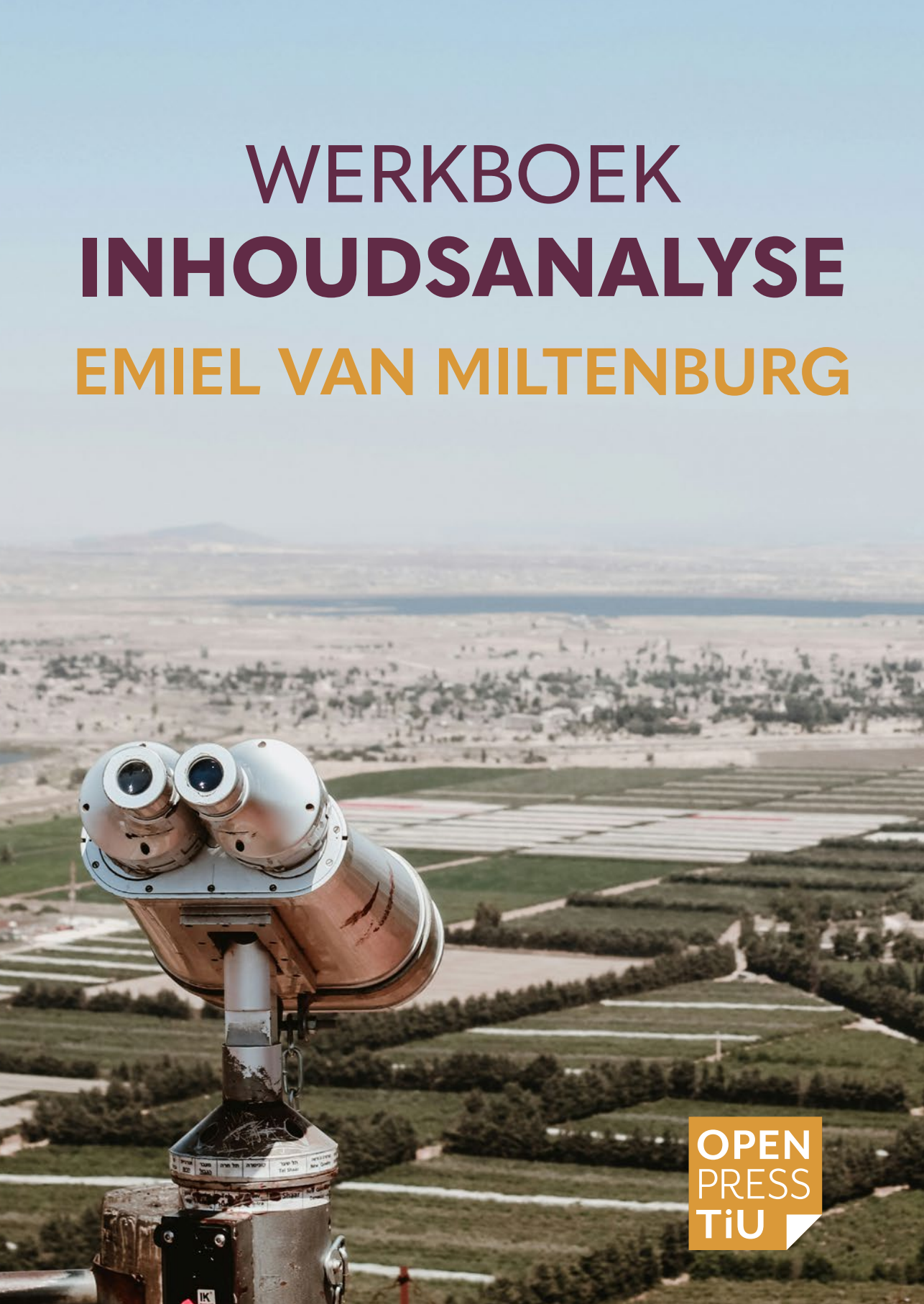
WERKBOEK
INHOUDSANALYSE

EMIEL VAN MILTENBURG



OPEN PRESS TiU

WERKBOEK INHOUDSANALYSE

WERKBOEK INHOUDSANALYSE

EMIEL VAN MILTENBURG

OPEN PRESS TiU

Open Press Tilburg University
Tilburg, The Netherlands

Design, typesetting, copyediting: LINE UP boek en media bv
Omslagfoto: Briana Tozour, verkregen via Unsplash.

ISBN: 9789403769509
DOI: https://doi.org/10.56675/wbiha

VERANTWOORDING

Dit boek is geschreven voor de cursus *Onderzoekspracticum 3: Inhoudsanalyse* die gegeven wordt in de opleiding *Communicatie en Informatiewetenschappen* van Tilburg University. Dit vak is in 2020 ontwikkeld door Christine Liebrecht en Emiel van Miltenburg. In 2023 is de cursus gegeven door Marie Barking en Emiel van Miltenburg. Dit boek is gebaseerd op het collegemateriaal dat we voor deze cursus hebben gemaakt. De uiteindelijke tekst is geschreven door Emiel van Miltenburg. Het heet een *Werkboek* omdat het de bedoeling is dat je ook echt aan de slag gaat met een eigen inhoudsanalyse tijdens het lezen van dit boek. De opdracht die we hiervoor in Tilburg gebruiken wordt omschreven in Bijlage C, met aanvullende informatie in Bijlage D en E. De rest van het boek is zo algemeen mogelijk opgeschreven, zodat de tekst ook bij andere opleidingen kan worden gebruikt. Veel van de genoemde studies komen wel uit Tilburg, om onze studenten bekend te maken met het onderzoek dat hier uitgevoerd wordt.

Dit boek had niet tot stand kunnen komen zonder mijn collega's. Ik wil graag Christine Liebrecht en Marie Barking bedanken voor de fijne samenwerking bij het geven van de cursus. Zij hebben me allebei geholpen met hun feedback op verschillende onderdelen van dit boek. Daarnaast wil ik Maria Mos bedanken voor haar uitgebreide commentaar op de volledige eerste versie van dit boek. Ik ben ook dankbaar voor het commentaar van Byron G. Adams, Martijn Goudbeek, Jeanine Guidry, Naomi Kamoen, Emiel Krahmer en Damian Trilling op verschillende onderdelen van het boek.

Natuurlijk blijf ik zelf verantwoordelijk voor alle fouten of onduidelijkheden die er nog in staan. Alle feedback blijft welkom voor de volgende editie(s) van dit boek. Je kunt me hiervoor een bericht sturen via het volgende adres:
C.W.J.vanMiltenburg@tilburguniversity.edu

INHOUDSOPGAVE

Bijlagen

HOOFDSTUK 1

Wat is een inhoudsanalyse eigenlijk?

1.1 INLEIDING

Als communicatiewetenschapper bestudeer je hoe mensen hun boodschappen vormgeven om hun ideeën en standpunten over te brengen op anderen, hoe anderen die boodschappen verwerken, en hoe verschillende media dat proces beïnvloeden. In deze cursus richten we ons op de boodschappen zelf: hoe zien ze eruit, en hoe kun je verschillende soorten boodschappen (foto's, teksten, video's) systematisch analyseren? Om deze vragen te beantwoorden, gebruiken we een methode genaamd Inhoudsanalyse (Engels: *content analysis*). Bij deze methode maak je eerst een verzameling boodschappen (een corpus), die je vervolgens analyseert en van labels voorziet. Dat labelen noemen we *coderen*.

1.1.1 KWANTITATIEVE INHOUDSANALYSE

Wij richten ons op *kwantitatieve* inhoudsanalyses, waarbij je eerst instructies opstelt om de data systematisch te kunnen coderen.[1] Vervolgens kunnen we de verschillende labels tellen, om zo verschillen tussen verschillende soorten boodschappen in kaart te kunnen brengen. Kwantitatieve inhoudsanalyse is meestal *deductief*: op basis van de literatuur kunnen we een onderzoeksvraag en bijbehorende hypothesen opstellen, waarna we de literatuur ook weer kunnen gebruiken om te bepalen wat er eigenlijk gecodeerd moet worden.

1 In sommige vakgebieden wordt er niet "coderen" maar "annoteren" gezegd, mogelijk omdat "coderen" te veel doet denken aan het Engelse woord voor "programmeren." Het woord "annoteren" betekent ongeveer: informatie toevoegen aan bestaande data.

1.1.2 KWALITATIEVE INHOUDSANALYSE

Naast een *kwantitatieve* inhoudsanalyse kun je ook een *kwalitatieve* inhoudsanalyse uitvoeren. Kwalitatieve inhoudsanalyses zijn inductief: vanuit de data kun je dan op zoek gaan naar regelmatigheden, waarna je als onderzoeker die patronen gaat beschrijven, die je kunt organiseren aan de hand van verschillende thema's. Deze werkwijze komt je misschien bekend voor als je eerder een interviewstudie of een focusgroep hebt uitgevoerd.

1.1.3 OVER DIT BOEK

Het idee achter dit boek is dat dit eerste hoofdstuk je een globaal overzicht geeft van de methode, waarna de volgende hoofdstukken je stap voor stap begeleiden door een groepsopdracht (zie Bijlage C) waarbij je samen met anderen een inhoudsanalyse uitvoert. De groepsopdracht is ontworpen voor een semester bij Tilburg University: in zeven weken werken studenten naar een onderzoeksvoorstel toe (zie Bijlage D), dat wordt nagekeken tijdens de eerste tentamenweek (*midterm*). Vervolgens passen de studenten hun plannen aan, waarna ze het onderzoek uitvoeren. Uiteindelijk krijgen studenten een cijfer voor de groepsopdracht (zie Bijlage E voor de richtlijnen) en een tentamen, waarbij het tentamen de individuele kennis en vaardigheden toetst die ze tijdens de groepsopdracht hebben opgedaan.

1.1.4 NASLAGWERKEN

Er zijn verschillende auteurs die de methodologie voor het uitvoeren van kwantitatieve inhoudsanalyses hebben beschreven. Populaire naslagwerken zijn bijvoorbeeld geschreven door Klaus Krippendorff (2018) en Kimberly Neuendorf (2017). Een goede optie is ook het werk van Riffe et al. (2023), of het Duits/Engelstalige handboek van Oehmer-Pedrazzi et al. (2023). Het laatste Nederlandstalige boek over inhoudsanalyse is geschreven door Fred Wester (2006). Dit werkboek is deels gebaseerd op deze naslagwerken. Twee grote verschillen met eerdere boeken over inhoudsanalyse zijn de vorm (een werkboek in plaats van een naslagwerk) en de sterke nadruk op onderzoeksethiek, waaraan ook een eigen hoofdstuk is gewijd.

1.2 EEN STAPPENPLAN

In deze cursus hanteren we het stappenplan voor het uitvoeren van een inhoudsanalyse, zoals beschreven door Neuendorf (2017, hoofdstuk 2).

1.2.1 STAP 1: ONDERWERP EN MOTIVATIE

Bedenk *wat* je gaat bestuderen en *waarom* je dat wil gaan doen. Neuendorf (2017) (in haar flowchart en in Hoofdstuk 4) benadrukt hierbij vooral het belang van theoretische overwegingen om een inhoudsanalyse uit te voeren. Daarnaast kunnen ook maatschappelijke (en zelfs ethische) overwegingen een rol spelen. Belangrijk is in ieder geval dat je een duidelijke onderzoeksvraag en bijbehorende hypotheses opstelt. Hierover, en over stap 2, lees je later meer in het hoofdstuk over *Onderwerpen en onderzoeksvragen*.

1.2.2 STAP 2: CONCEPTUALISATIE

Specificeer welke variabelen je allemaal gaat bestuderen, en zorg bij iedere variabele voor een heldere definitie. Op dit moment ga je vaak heen en weer tussen de theorie en een eerste selectie met voorbeelden die je alvast hebt verzameld. Sluiten de definities aan op de voorbeelden die je hebt gevonden?

1.2.3 STAP 3: OPERATIONALISATIE

Hoe ga je te werk? Wat ga je precies verzamelen en coderen, en hoe ga je de data vervolgens analyseren? Neuendorf (2017) benadrukt hier vooral dat de categorieën die je gebruikt (i) zo algemeen mogelijk en (ii) *mutually exclusive* moeten zijn (met andere woorden: ze moeten elkaar uitsluiten). Dit zullen we bespreken in het hoofdstuk *Een codeboek ontwikkelen.*

1.2.4 STAP 4: EEN CODEERSCHEMA OPSTELLEN

Neuendorf (2017) onderscheidt hier twee methodes voor het coderen van data:

1) Handmatig coderen. Hierbij moet je een codeboek (een soort handleiding) opstellen, en een codeerformulier maken. Het is belangrijk dat het codeboek zo duidelijk is dat verschillende mensen dezelfde data op dezelfde manier zouden coderen. Dit zullen we bespreken in het hoofdstuk *Een codeboek ontwikkelen.*
2) Automatisch coderen. Hierbij laat je de computer het werk doen. Neuendorf (2017) gaat ervan uit dat je dit doet aan de hand van bestaande of zelfgemaakte woordenlijsten. Voor iedere categorie is er dan een aparte lijst met woorden, bijvoorbeeld *positieve* en *negatieve* woorden. De computer kan die woorden dan automatisch gaan tellen in de data. Tegenwoordig is er veel meer mogelijk dan alleen het gebruik van woordenlijsten, maar dat bespreken we later in het hoofdstuk *Mensen en computers.*

1.2.5 STAP 5: SAMPLING

In het ideale geval gebruik je volgens Neuendorf (2017) gewoon alle data die er bestaat: een census. Om praktische redenen is dit vaak niet mogelijk, dus zul je een manier moeten bedenken om een representatieve selectie te maken van de data, om zo toch conclusies te kunnen trekken over de gehele populatie. In het hoofdstuk over *Sampling* leer je hier meer over.

1.2.6 STAP 6: CODEURS TRAINEN EN BETROUWBAARHEID BEREKENEN

Als je data hebt verzameld en een codeboek hebt opgesteld, kun je de codeurs trainen (leren omgaan met het codeboek en het codeerformulier), en vervolgens kijken in hoeverre ze op betrouwbare wijze kunnen coderen. Met andere woorden: doen de verschillende codeurs allemaal hetzelfde, of verschillen de codeurs in hoe ze de data gecodeerd hebben? Als er veel verschillen zijn tussen de codeurs, dan moet het codeboek herzien worden. Deze onderwerpen bespreken we in de hoofdstukken *Een codeboek ontwikkelen* en *Betrouwbaarheid en statistiek.*

1.2.7 STAP 7: CODEREN

Als je in een pilotstudie hebt bevestigd dat de codeurs de data betrouwbaar kunnen categoriseren, kun je overgaan tot de hoofdcodeerfase.

1) Als je de data door mensen laat coderen, raadt Neuendorf (2017) aan om tenminste twee onafhankelijke codeurs te gebruiken, die minstens 10% van de data dubbel coderen. Daaraan zouden we kunnen toevoegen dat je in ieder geval genoeg moet laten dubbel coderen om aan het eind van het onderzoek te kunnen bevestigen dat de data inderdaad betrouwbaar gecodeerd is. (Met andere woorden: 10% is niet altijd genoeg.) In het hoofdstuk over *Betrouwbaarheid en statistiek* lees je hier meer over.
2) Als je de data door computers laat coderen, raadt Neuendorf (2017) aan om steekproefsgewijs te controleren of je het eens bent met de computer. Tegenwoordig is dat niet meer genoeg, maar is het aan te raden om de computer systematisch te controleren. In het hoofdstuk over *Mensen en computers* lees je hier meer over.

1.2.8 STAP 8: BETROUWBAARHEID BEREKENEN

Er bestaan verschillende statistische maten om te controleren in hoeverre verschillende codeurs het met elkaar eens zijn. Die maten geven een indicatie van de betrouwbaarheid van je studie; als de codeurs het grotendeels oneens waren, hoe kun je dan nog conclusies trekken uit de resultaten van je studie? In het hoofdstuk over *Betrouwbaarheid en statistiek* lees je hier meer over.

1.2.9 STAP 9: RAPPORTAGE

Uiteindelijk moet je de resultaten ook overzichtelijk opschrijven. In het hoofdstuk *Rapportage* lees je hier meer over.

1.3 INHOUDSANALYSE VERSUS EXPERIMENTELE STUDIES

Hierboven hebben we al gezien dat je met een inhoudsanalyse goed kunt beschrijven *hoe* mensen eigenlijk communiceren. Het is dus een goede methode om te beschrijven welke patronen we kunnen waarnemen in het communicatiegedrag van verschillende groepen mensen, of in verschillende communicatiesituaties. Daarnaast brengt iedere onderzoeksmethode zijn eigen voor- en nadelen met zich mee. Laten we beginnen met een bekend voorbeeld: een experiment.

1.3.1 VOOR- EN NADELEN VAN EEN EXPERIMENTELE STUDIE

Het voordeel van een experiment is bijvoorbeeld dat je maximale controle hebt over jouw onderzoek: participanten worden toegewezen aan specifieke condities, waarbinnen ze blootgesteld worden aan bepaalde stimuli. Voor en na de blootstelling kun je metingen uitvoeren, waarna je kunt analyseren of de verschillende condities een verschil teweeg hebben gebracht

tussen de participanten. Als dat zo is, dan kun je stellen dat er een causaal verband is tussen de experimentele manipulatie en het verschil in de metingen. Bijvoorbeeld: informeel taalgebruik zorgt ervoor dat participanten een chatbot menselijker vinden dan een chatbot die je heel formeel behandelt.

Het nadeel van experimenten is dat ze vaak kunstmatig zijn; participanten weten dat ze deelnemen aan een experiment, en dat het dus gaat om een simulatie. Het is niet altijd duidelijk of ze zich "in het echt" ook zo zouden gedragen.

1.3.2 VOOR- EN NADELEN VAN EEN INHOUDSANALYSE

Het voordeel van inhoudsanalyses is dat je (meestal) werkt met bestaande data, waardoor je geen participanten nodig hebt. Het is een onopvallende methode, waarbij je de zender niet beïnvloedt. Omdat de datapunten uit het leven gegrepen zijn, hebben inhoudsanalyses vaak een hoge ecologische validiteit. Het nadeel van inhoudsanalyses is dat je minder controle hebt over de data; je moet werken met de gegevens die beschikbaar zijn, of die binnen afzienbare tijd beschikbaar gemaakt kunnen worden. Daarnaast is het vaak niet mogelijk om causale verbanden vast te stellen. Je weet na een inhoudsanalyse misschien wel of er een verschil is tussen bepaalde groepen, maar je kunt niet altijd vaststellen wat de oorzaak is van dat verschil.

1.4 DE ROL VAN EEN INHOUDSANALYSE

Wat is de rol van inhoudsanalyses in de wetenschap? Kort samengevat gaat het om de toegevoegde waarde van het beschrijven van de stand van zaken in de wereld. De redacteuren van het Journal of Quantitative Description: Digital Media (JQD:DM) beschrijven descriptieve kennis als noodzakelijk voor verschillende onderdelen van het sociaal-wetenschappelijke proces, specifiek: het genereren van hypothesen, het inschatten van het belang van het onderwerp, en om meer te kunnen zeggen over de generaliseerbaarheid van onderzoek in het algemeen.[2] Hieronder bespreken we deze onderdelen in meer detail.

1.4.1 HYPOTHESEN OPSTELLEN

Het is vrijwel onmogelijk om een zinvolle hypothese op te stellen zonder te weten hoe de wereld er überhaupt uitziet. Eerst moeten we weten wat er gaande is, voordat we iets kunnen zeggen over de oorzaken en gevolgen daarvan. Dus als je bijvoorbeeld de invloed van influencers op social media wil onderzoeken, dan is het goed om eerst vast te stellen wat die influencers eigenlijk doen, en hoeveel dat eigenlijk gebeurt. Die eerste stap kun je zetten door middel van een inhoudsanalyse. Daarna, in vervolgonderzoek, kun je het onderzoek gaan

2 De volledige tekst ten tijde van het schrijven van deze inleiding is hier te vinden: https://web.archive.org/web/20240906235123/https://journalqd.org/about

afbakenen: we beperken ons tot een specifieke groep mensen met een specifiek soort gedrag, om vervolgens van die groep te onderzoeken waarom ze dat eigenlijk doen (bijvoorbeeld via interviews), en welke invloed dat heeft op andere sociale mediagebruikers (bijvoorbeeld via focusgroepen of een experiment). De redacteuren van JQD:DM beklagen zich erover dat er vaak experimenten worden ontworpen zonder dat er eerst gekeken wordt naar de prevalentie [dat wil zeggen: hoe vaak iets voorkomt] van oorzaken en effecten die er bestudeerd worden. Dat brengt ons op het tweede punt.

1.4.2 HET INSCHATTEN VAN HET BELANG VAN HET ONDERWERP

De redacteuren van JQD:DM zijn minder concreet over dit punt, maar stellen vooral dat kwantitatieve beschrijving helpt om het belang van een onderwerp in vervolgonderzoek beter te kunnen beargumenteren. Waar zou je dan aan moeten denken?

Vaker is belangrijker.
Wanneer je vaststelt dat een bepaald fenomeen vaak voorkomt, kun je op basis van die observatie beargumenteren dat het daarom ook belangrijk is dat we begrijpen welke rol dat fenomeen speelt in ons dagelijks leven.

Uitzonderlijke gevallen zijn belangrijk.
Je zou de argumentatie ook kunnen omkeren. Als iets *meestal* op een bepaalde manier gebeurt, dan kan het ook interessant zijn om te kijken naar de uitzonderingen. Waarom wijken mensen soms af van de standaard? Zijn onze theorieën bestand tegen die uitzonderingen? Een inhoudsanalyse kan in ieder geval een vertrekpunt geven om een verklaring te vinden, zeker als je ontdekt dat mensen zich *systematisch* anders gedragen in bepaalde situaties.

1.4.3 GENERALISEERBAARHEID

Als wetenschappers stellen we ons ten doel om generaliseerbare kennis te vergaren. Maar om te kunnen voorspellen of bevindingen uit de ene situatie ook van toepassing zijn in een andere situatie, moeten we wel begrijpen hoe beide situaties eruitzien. Theorieën nemen vaak de vorm aan van een logische implicatie: als A, B, en C, dan gebeurt er X. Met experimenten kun je vaststellen of er een causaal verband is tussen ABC en X, en met beschrijvend onderzoek kun je vaststellen of er in een bepaalde situatie inderdaad sprake is van A, B, en C.

1.5 AAN DE SLAG

Je weet nu ongeveer wat een inhoudsanalyse is. Maar de beste manier om echt te begrijpen wat een inhoudsanalyse is en wat je ermee kan, is om eerdere onderzoeken te lezen en om zelf een inhoudsanalyse uit te voeren. Het volgende hoofdstuk beschrijft de groepsopdracht, waar we dit semester mee aan de slag gaan. Daarin zul je een eigen inhoudsanalyse gaan

opzetten en uitvoeren. Voor nu vragen we je om de onderstaande vragen door te nemen en te beantwoorden voor twee artikelen.

1.5.1 KRITISCHE VRAGEN OVER INHOUDSANALYSES

Deze vragen kun je stellen over ieder artikel dat een inhoudsanalyse bevat. Door de vragen te beantwoorden, begrijp je beter hoe de studie in elkaar zit. Hoewel de meeste concepten al bekend zouden moeten zijn uit de methodologiecursus, kan het zijn dat je nog niet bekend bent met alle terminologie. In dat geval kun je de begrippenlijst raadplegen aan het eind van dit werkboek (Bijlage A).

1) Algemeen: wat is de onderzoeksvraag?
2) Sampling: samenstelling van het corpus.
   a) Wat voor samplingstrategie hebben de auteurs gebruikt?
   b) Hoeveel items hebben de auteurs verzameld?
   c) Stel je zou zelf een vergelijkbaar corpus willen samenstellen; bevat het artikel dan alle relevante details om de datacollectiemethode te repliceren, of denk je dat er iets ontbreekt?
3) Coderen: labels toewijzen aan de data.
   a) Op welke categorieën zijn de data geanalyseerd?
   b) Waar komen die categorieën vandaan? Waar zijn ze op gebaseerd?
   c) Wat zijn de *data collection units*? Met andere woorden: welke discrete units worden er geannoteerd?
   d) Hoe zorgen de auteurs ervoor dat de annotaties betrouwbaar zijn?
   e) Welke categorie was het meest betrouwbaar te coderen? En welke het minste?
4) Statistiek en rapportage
   a) Hoe is de resultatensectie gestructureerd? (Bijvoorbeeld: in welke volgorde worden verschillende soorten statistiek gepresenteerd?)
   b) Welke statistieken worden er gerapporteerd over de geannoteerde data?
   c) Hoe worden de resultaten gevisualiseerd? (Bijvoorbeeld: tabellen, grafieken.)
   d) Als er geen of beperkte visualisaties zijn: hoe zou je de resultaten anders kunnen weergeven?
5) Ethiek
   a) Zijn er specifieke ethische aandachtspunten voor deze studie?
   b) Is het duidelijk hoe de auteurs hiermee om zijn gegaan?
   c) Zie je hier mogelijkheden tot verbetering?
6) Algemeen oordeel
   a) Wat zijn de sterke punten van deze studie?
   b) Wat zijn de minder sterke punten van deze studie. (Denk aan het design, of aan de (on)volledigheid van de rapportage.
   c) Als er minder sterke punten zijn, hoe zou je die dan kunnen oplossen?

Deze kritische vragen staan ook nog in de appendix op één pagina (zonder andere tekst), zodat je ze makkelijk kunt printen om erbij te houden als je een artikel leest.

1.5.2 ARTIKELEN

Het recente handboek van Oehmer-Pedrazzi et al. (2023) bevat meerdere overzichten van inhoudsanalyses die in verschillende domeinen van de communicatiewetenschap gepubliceerd zijn. Bijvoorbeeld:

- Bedrijfscommunicatie (Lischka, 2023).
- Wetenschapscommunicatie (Wicke, 2023)
- Gezondheidscommunicatie (Samson-Himmelstjerna, 2023)
- Politieke communicatie (Blassnig, 2023; Van der Velden & Loecherbach, 2023)
- Online comments en wangedrag (Esau, 2023; Naab & Küchler, 2023)
- …en nog veel meer (zie de inhoudsopgave van het boek).

Gebruik de overzichtsartikelen om twee publicaties te vinden waarin een inhoudsanalyse uitgevoerd wordt. Beantwoord vervolgens de leesvragen voor beide publicaties. Deze oefening bereidt je ook voor op de groepsopdracht (zie Bijlage C) die hoort bij deze cursus.

Studietip

Als je iedere week één of twee artikelen leest (uit deze lijst of uit een betrouwbaar tijdschrift) en daarbij de leesvragen beantwoordt, begrijp je steeds beter hoe inhoudsanalyses opgebouwd zijn, en hoe je de resultaten zou kunnen rapporteren.

HOOFDSTUK 2

Onderwerpen en onderzoeksvragen

2.1 INLEIDING

De eerste vraag die je als onderzoeker zult moeten beantwoorden is: *wat ga ik eigenlijk bestuderen?* Een antwoord geven op die vraag is nog niet zo eenvoudig, omdat er drie vragen aan ten grondslag liggen:

1) Wat voor vragen kun je beantwoorden met een inhoudsanalyse?
2) Wat is er theoretisch relevant om te bestuderen?
3) Wat is er maatschappelijk relevant om te bestuderen?

In dit hoofdstuk gaan we dieper in op deze verschillende vragen.

2.2 WAT IS ER MOGELIJK?

Zoals ook al besproken in hoofdstuk 1, is het vaak niet mogelijk om causale verbanden aan te tonen met een inhoudsanalyse. Dat komt omdat inhoudsanalyses meestal uitgevoerd worden met bestaande data, waarbij er geen experimentele manipulatie heeft plaatsgevonden. Daardoor kunnen we niet vaststellen of een verandering in een bepaalde onafhankelijke variabele leidt tot een verandering in een afhankelijke variabele. We kunnen wel een beschrijving geven van de wereld zoals zij is, waarbij we ook kunnen kijken of bepaalde variabelen met elkaar correleren. Het is dan alleen niet mogelijk om eventuele *confounders* uit te sluiten.

2.2.1 UITZONDERING: INHOUDSANALYSES BINNEN EXPERIMENTEN

Inhoudsanalyses kunnen ook uitgevoerd worden binnen een andersoortige methode, zoals een experiment. Een goed voorbeeld is de studie van Braun et al. (2019), die participanten in een experiment hebben gevraagd om een tafelvoetbalwedstrijd te spelen, waarna de participanten een wedstrijdverslag moesten schrijven. Die wedstrijdverslagen werden vervolgens geanalyseerd om te kijken hoe de teksten van winnaars en verliezers van elkaar verschillen. Hier is dus wel sprake van een causaal verband. De onafhankelijke variabele is hierbij het perspectief (winnaar of verliezer), en de afhankelijke variabelen zijn verschillende tekst-eigenschappen. Andere voorbeelden van inhoudsanalyses binnen een experimenteel design zijn de studies van Antheunis et al. (2011) en Mui et al. (2017).

2.3 DOELEN EN VRAGEN

We hebben hierboven vastgesteld dat inhoudsanalyses meestal beschrijvend van aard zijn. Maar dat klinkt nog best abstract; wat bedoelen we eigenlijk met een beschrijving? Hieronder vind je drie mogelijke invalshoeken die je kunt innemen bij het formuleren van je onderzoeksvraag. Bij iedere invalshoek staan ook voorbeelden van mogelijke onderzoeksvragen geformuleerd.

1) In kaart brengen van de situatie waar je in geïnteresseerd bent:
   a) Wat voor soort voedsel wordt er eigenlijk gepromoot door influencers?
   b) Welke gespreksstrategieën hanteren bedrijven eigenlijk in hun webcare-gesprekken?
2) Beschrijven van veranderingen over tijd:
   a) Vergeleken met 20 jaar geleden, is de man-vrouw verhouding in talkshows veranderd?
   b) Is het klimaatprobleem aanweziger dan vroeger in de verkiezingsprogramma's?
3) Toetsen van verwachtingen over de verdeling van data:
   a) Onderbreken mannen vrouwen inderdaad vaker in een gesprek dan andersom?
   b) Verontschuldigen luchtvaartmaatschappijen zich inderdaad bijna nooit op Twitter? (Zie ook het voorbeeld in de volgende sectie.)

Wetenschappelijke vragen staan nooit op zichzelf, maar sluiten aan op huidige discussies in de literatuur. Hierover lees je meer in Sectie 2.4.

2.3.1 TEST JEZELF

Kun je nog meer onderzoeksvragen bedenken die passen bij de drie bovenstaande invalshoeken?

2.4 THEORETISCHE RELEVANTIE

In het eerste hoofdstuk hebben we al gesproken over de rol van inhoudsanalyses in sociaal-wetenschappelijk onderzoek, waarbij we hebben gezien dat inhoudsanalyses vaak een belangrijke eerste stap kunnen zijn in het beantwoorden van grotere vragen. Samen met andere onderzoeksmethodes kun je dan laten zien hoe de wereld in elkaar zit. In deze paragraaf kijken we naar de theoretische relevantie van individuele studies.

Een inhoudsanalyse is theoretisch relevant als het onderzoek ons helpt om openstaande vragen in de literatuur te beantwoorden, of als het onderzoek ons helpt om weer nieuwe vragen te kunnen stellen. Daarnaast kun je verwachtingen of tegenstrijdigheden in de wetenschappelijke literatuur toetsen aan de werkelijkheid.

2.4.1 EEN VOORBEELD

Van Hooijdonk & Liebrecht (2021) gebruikten een inhoudsanalyse om communicatiestrategieën in webcaregesprekken in kaart te brengen. Webcaregesprekken zijn interacties tussen consumenten en organisaties op sociale media, waarbij de aanleiding van het gesprek vaak is dat een (potentiële) klant een vraag of klacht heeft over een product of service van een organisatie. De organisatie kan dan verschillende communicatiestrategieën hanteren, zoals: informatie geven, doorverwijzen naar een ander kanaal, actie ondernemen, enzovoorts. Van Hooijdonk & Liebrecht (2021) richtten zich daarbij in het eerste deel van hun studie op verontschuldigingen die al dan niet worden aangeboden door luchtvaartmaatschappijen wanneer iemand klaagt op Twitter.

De studie van Van Hooijdonk & Liebrecht (2021) heeft een duidelijke maatschappelijke relevantie: dankzij zulke studies weten we beter hoe klantenservicemedewerkers zouden moeten reageren om klanten tevreden te stellen. Maar wat is nu de theoretische of wetenschappelijke relevantie van deze studie?

Gebrek aan kennis in de literatuur: Er was wel veel geschreven over de aanwezigheid van verontschuldigingen in webcaregesprekken, maar nog weinig over de vorm. Van Hooijdonk & Liebrecht (2021) beargumenteren in hun artikel dat het ook uitmaakt (i) hoe je een verontschuldiging aanbiedt, en (ii) welke andere responsstrategieën je daarnaast nog gebruikt in combinatie met de verontschuldiging.

Onenigheid in de literatuur: Van Hooijdonk & Liebrecht (2021) stellen in hun artikel dat de literatuur niet eenduidig is over de rol van verontschuldigingen in webcaregesprekken. Enerzijds erken je met een verontschuldiging dat je een fout hebt gemaakt, wat kan leiden tot reputatieschade, maar anderzijds laat je met een verontschuldiging ook zien dat je de klant vooropstelt, wat een positief effect kan hebben op je reputatie. Die onenigheid maakt het interessant om te kijken wat er eigenlijk gebeurt in de praktijk: welke keuzes maken organisaties nou eigenlijk op dit vlak?

Ecologische validiteit: In het tweede, experimentele deel van hun studie, keken Van Hooijdonk & Liebrecht (2021) naar de effecten die verontschuldigingen kunnen hebben op

het beeld dat klanten hebben van een organisatie die deze gesprekstrategie hanteert. Doordat ze eerder een inhoudsanalyse hadden uitgevoerd, konden ze ervoor zorgen dat hun experimentele stimuli en condities aansluiten op de dagelijkse praktijk. Dit verhoogt de ecologische validiteit van hun experiment.

Het laatste punt laat de meerwaarde zien van de inhoudsanalyse in de studie van Van Hooijdonk & Liebrecht (2021), maar op zichzelf geeft dit argument het artikel geen grotere theoretische relevantie; ecologische validiteit is hier een *studie-interne reden* om een inhoudsanalyse uit te voeren. Het zou natuurlijk wel zo kunnen zijn dat je ziet dat de studies die momenteel uitgevoerd worden binnen een bepaald vakgebied eigenlijk allemaal wat kunstmatig zijn. Op zo'n moment kun je wel beargumenteren dat een inhoudsanalyse kan bijdragen aan de ecologische validiteit van studies binnen het vakgebied. De inhoudsanalyse laat dan zien hoe stimuli of experimentele condities eruit zouden moeten zien.

2.4.2 HOE VIND JE EEN THEORETISCH RELEVANTE ONDERZOEKSVRAAG?

Er zijn verschillende manieren om een relevante onderzoeksvraag te formuleren. Hieronder beschrijven we er drie.

Algemene vragen binnen communicatiewetenschap

We begonnen dit boek met een algemene definitie van communicatiewetenschap:

> *Als communicatiewetenschapper bestudeer je hoe mensen hun boodschappen vormgeven om hun ideeën en standpunten over te brengen op anderen, hoe anderen die boodschappen verwerken, en hoe verschillende media dat proces beïnvloeden.*

Die definitie kunnen we verder verkennen door te kijken naar verschillende communicatiesituaties. In een klassieke situatie met een zender en een ontvanger kun je je bijvoorbeeld afvragen:

- ...hoe mensen of organisaties hun boodschap afstemmen op hun (beoogde) publiek.
- ...hoe mensen of organisaties hun boodschap aanpassen voor verschillende kanalen (kranten, televisie, sociale media, een eigen website, ...)
- ...hoe mensen reageren op die boodschappen.
- ...welke mensen reageren op die boodschappen.

Voorbeeld 1: Het artikel van Huibers & Verhoeven (2014) bevat drie onderzoeksvragen waarin je een aantal van de bovenstaande vragen kunt herkennen:

1) "Welke webcarestrategieën zetten Nederlandse organisaties in om te reageren op klachten van stakeholders?"
2) "Welk effect hebben deze webcarestrategieën op de corporate reputatie?"

3) "Welke invloed heeft een CHV [*Conversational Human Voice; een warme, informele, menselijke communicatiestijl*] op de effectiviteit van webcarestrategieën?"

Het domein waar de auteurs naar kijken, *webcare*, richt zich specifiek op de manier waarop organisaties reageren op vragen die op sociale media gesteld worden door *stakeholders* (in dit geval vaak: klanten of consumenten in het algemeen). Daarbij is het van belang dat de reputatie van de organisatie beschermd wordt, en de vraag is hoe organisaties hun communicatiestijl daarop aanpassen. De organisatie is hier de zender, en het publiek is de ontvanger. In gelijkwaardige gesprekken heb je niet echt een duidelijke zender of ontvanger, maar zie je dat mensen hun boodschappen op elkaar aanpassen. Afhankelijk van de specifieke situatie kun je je bijvoorbeeld afvragen:

- ...hoe mensen zich door de tijd aanpassen aan elkaar.
- ...welke onderwerpen er wel/niet besproken worden.
- ...welke communicatieve normen er bestaan binnen verschillende groepen of gemeenschappen.
- ...wat de verschillen zijn in vorm en inhoud tussen gesprekken die gevoerd worden via verschillende kanalen (bijvoorbeeld *face-to-face* versus online gesprekken).

Voorbeeld 2: Swerts et al. (2021) hebben gekeken naar interacties tussen Nederlanders en Vlamingen. Eerder onderzoek heeft laten zien dat mensen zich op verschillende niveaus kunnen aanpassen aan elkaar. Zo kunnen ze bijvoorbeeld hun woordkeuze (bijvoorbeeld: *microgolf* versus *magnetron*), uitspraak (Nederlands of Vlaams) of hun grammatica (typisch Nederlandse of Vlaamse zinsconstructies) aanpassen op de gesprekspartner. In hun studie hebben Swerts et al. (2021) twee experimenten uitgevoerd om te kijken of en hoe Nederlanders en Vlamingen zich aan elkaar aanpassen. Zoiets zou je natuurlijk ook met een inhoudsanalyse kunnen onderzoeken aan de hand van bestaande gesprekken (bijvoorbeeld op televisie, YouTube, of in podcasts). In plaats van nationaliteit zou je ook kunnen kijken naar andere (sociaal-economische) factoren.

Openstaande vragen

Onderzoek is nooit af. Vrijwel alle artikelen eindigen met een discussie en conclusie waarin aangegeven wordt wat de bevindingen zijn, en welke vragen er nog open staan voor vervolgonderzoek. Er zijn daarbij twee verschillende soorten studies:

1) Rapportage van onderzoek dat door de auteurs zelf is uitgevoerd.
2) Overzichtsartikelen die samenvatten wat er binnen een bepaald deelgebied al is gedaan.

Overzichtsartikelen kunnen erg nuttig zijn omdat ze een langetermijnperspectief schetsen: wat moet er allemaal nog gedaan worden? Welke Grote Vragen blijven nog onbeantwoord? Dat perspectief ontbreekt soms als je losse artikelen leest, omdat auteurs vaak schrijven voor

vakgenoten die een gedeelde achtergrond hebben. Hieronder bespreken we verschillende manieren om onderzoeksideeën te ontwikkelen op basis van bestaande artikelen.

Voorbeeld 3: Guidry et al. (2015) hebben gekeken naar online debatten over vaccinaties op sociale media. In deze studie werden 800 posts op het *Pinterest*-platform geanalyseerd en gecategoriseerd om te kijken wat de balans is tussen voor- en tegenstanders van vaccineren (de meerderheid was tegen), welke argumenten er werden gebruikt, en hoe die argumenten werden ondersteund met behulp van afbeeldingen. De auteurs gebruiken voor hun analyse een specifiek theoretisch model (het Health Belief Model; HBM), maar stellen in hun conclusie voor dat het relevant zou zijn om een vergelijkbare studie uit te voeren maar dan aan de hand van andere communicatietheorieën (*Theory of Reasoned Action* en het *Extended Parallel Processing Model*). Daarnaast geven de auteurs aan dat het relevant zou zijn om andere sociale media (*Facebook* en *Instagram*; *TikTok* bestond toen nog niet) te onderzoeken om te kijken of daar dezelfde normen gelden als op *Pinterest*.

Meer diepgang, of verschillende domeinen

Het artikel van Guidry et al. (2015) is ook nuttig om een andere manier te illustreren om aan een onderzoeksvraag te komen. Als je het artikel leest, zie je dat de auteurs een specifiek codeboek gebruiken om de data te analyseren. Je kunt dat codeboek dan ook bekijken en nadenken over manieren om de analyse uit te breiden. Bijvoorbeeld: welke retorische stijlfiguren worden er eigenlijk gebruikt om argumenten op sociale media kracht bij te zetten?

Wat je ook ziet is dat de auteurs zelf al aangeven dat het nuttig zou kunnen zijn om andere sociale media te onderzoeken. Dit wordt niet altijd genoemd als vervolgonderzoek, maar vaak kan het nuttig zijn om te kijken of observaties van het ene platform ook gelden voor het andere platform. (We moeten natuurlijk niet overhaast generaliseren van één platform naar alle sociale media in het algemeen.) Daarbij is het goed om ook na te denken over de verschillende *affordances* of design-eigenschappen van de verschillende platforms. Hoe denk je dat die eigenschappen de aard van de posts op het platform zouden kunnen beïnvloeden?

2.4.3 TEST JEZELF

Als je wil oefenen met het concept "theoretische relevantie" dan zijn er drie opdrachten die je voor jezelf kunt uitvoeren. Iedere opdracht is net wat uitdagender dan de vorige.

1) Markeer in het artikel van Van Hooijdonk & Liebrecht (2021) alle zinnen of alinea's waar de theoretische relevantie van de studie besproken wordt.
   a) Waar in het artikel wordt de theoretische relevantie besproken?
   b) Hoe doen de auteurs dat eigenlijk?
2) Zoek zelf een wetenschappelijk artikel en maak een lijst van argumenten waarom dat artikel theoretisch relevant is volgens de auteurs.

3) Formuleer een nieuwe onderzoeksvraag naar aanleiding van een wetenschappelijk artikel over een onderwerp naar keuze. Zorg voor een goede onderbouwing van de relevantie van de onderzoeksvraag.

2.5 MAATSCHAPPELIJKE RELEVANTIE

De tijd van wetenschap vanuit de ivoren toren is allang voorbij. Wetenschappers zijn onderdeel van de maatschappij, en hebben een duidelijke verantwoordelijkheid tegenover de maatschappij: universitair onderzoek wordt betaald door de Nederlandse belastingbetaler, en daar staat tegenover dat wij ervoor zorgen dat we relevant onderzoek doen. Maar ook los van deze financiële kwestie geldt dat wij ons aan de universiteit in een bevoorrechte positie bevinden. Simpel gezegd: we zijn het aan onze stand verplicht om, met alle kennis en middelen die we aangereikt hebben gekregen, de maatschappij te helpen verbeteren. *Noblesse oblige*[1]. Of in modernere bewoordingen: "with great power comes great responsibility" (Stan Lee).

2.5.1 HOE KAN ONDERZOEK DE MAATSCHAPPIJ HELPEN?

Onderzoek kan mensen in de praktijk helpen om beter geïnformeerde keuzes te maken, en het kan overheden helpen om regels en richtlijnen op te stellen. Niet alle onderzoeken dragen direct bij aan deze doelen. Vaak is het zo dat er meerdere onderzoeken nodig zijn voordat er duidelijke conclusies getrokken kunnen worden, maar iedere volgende studie zorgt er weer voor dat we weer een stapje verder kunnen komen. Zo kan een inhoudsanalyse bijvoorbeeld:

- In kaart brengen welke gespreksstrategieën er worden gebruikt door verschillende bedrijven in webcare-gesprekken op sociale media (Huibers & Verhoeven, 2014), waarna vervolgonderzoek kan aantonen welke soorten boodschappen het meest effectief zijn om klanten tevreden te houden en het gevoel te geven dat ze gehoord worden als ze online hun beklag doen.
- In kaart brengen in hoeverre influencers op sociale media reclame maken die gericht is op minderjarigen (Qutteina et al., 2019), waarna vervolgonderzoek kan aantonen wat de effecten zijn van de verschillende soorten reclameboodschappen waarmee jongeren worden geconfronteerd. Beleidsmakers kunnen deze informatie dan gebruiken om regels op te stellen waar influencers zich aan moeten houden.
- In kaart brengen wat voor boodschappen over vaccinatie er gedeeld worden op sociale media (Guidry et al., 2015), waarmee platforms op hun verantwoordelijkheden gewezen worden om schadelijke informatie te modereren.

1 https://nl.wikipedia.org/wiki/Noblesse_oblige

2.5.2 ETHISCHE VRAGEN

Achter het idee van maatschappelijke relevantie zit een ethische vraag: hoe kunnen wij onze beperkte tijd het best gebruiken om een bijdrage te leveren aan de maatschappij? Daarnaast kun je je ook afvragen wie er eigenlijk baat heeft bij het onderzoek dat je doet. Zijn dat kwetsbare mensen in de maatschappij, of juist de mensen die het eigenlijk al heel goed hebben? Deze vragen hebben geen eenduidig antwoord. Als je een bijdrage levert aan het bedrijfsleven, dan kan dat bijvoorbeeld ook een positief effect hebben op de gehele Nederlandse economie. Verschillende onderzoekers maken hier verschillende keuzes, maar het is belangrijk dat je in ieder geval beseft dat er een keuze is, en dat die keuze gevolgen kan hebben voor andere mensen.

2.5.3 KADERS VOOR MAATSCHAPPELIJKE RELEVANTIE

In een inleidende methodologiecursus verwachten we niet dat je in één keer de wereld gaat verbeteren, maar het is wel goed om te weten dat er verschillende referentiekaders zijn om maatschappelijk relevant onderzoek uit te voeren.

Nationale wetenschapsagenda

In 2014 werd de Nationale Wetenschapsagenda (NWA) geïntroduceerd door de toenmalige minister van Onderwijs, Cultuur, en Wetenschap, Jet Bussemaker. Het idee was simpel: alle burgers van Nederland konden een vraag insturen voor de wetenschap. Op basis van die vragen zouden we dan kunnen vaststellen wat men in Nederland eigenlijk zou willen weten, en welke onderzoeksgebieden meer aandacht zouden moeten verdienen. Uiteindelijk zijn er bijna 12.000 vragen ingediend, die zijn samengevat in 140 zogenaamde clustervragen. Je kunt deze vragen lezen via de website: https://vragen.wetenschapsagenda.nl Een deel van de Nederlandse subsidies voor wetenschappelijk onderzoek is gekoppeld aan de NWA, waardoor onderzoekers moeten laten zien dat hun onderzoeksvoorstellen relevant zijn voor de verschillende clustervragen.

Er is ook kritiek geweest op de nationale wetenschapsagenda. Kort samengevat komt die kritiek hierop neer: waarom zouden we onze koers laten bepalen door amateurs? Wetenschappers zijn toch opgeleid om zelf te bepalen wat relevant onderzoek is? Zij weten als geen ander waar de grenzen van onze kennis liggen, en waar mogelijk ruimte is voor een wetenschappelijke doorbraak. Doet de nationale wetenschapsagenda dan geen afbreuk aan de vrijheid en onafhankelijkheid van wetenschappers? Uiteindelijk voorziet de NWA maar een deel van het geld dat middels subsidies verdeeld wordt aan de Nederlandse wetenschap. Dat betekent (gelukkig) dat ook onderwerpen buiten het blikveld van de NWA de aandacht kunnen krijgen die ze verdienen.

UNESCO Sustainable Development Goals

In 2015 heeft de algemene vergadering van de Verenigde Naties een agenda opgesteld voor duurzame ontwikkeling richting het jaar 2030. In deze agenda staan zeventien duurzame ontwikkelingsdoelen (*Sustainable Development Goals; SDGs*) centraal (geciteerd van de VN-website unric.org):

1) Beëindig armoede overal en in al haar vormen
2) Beëindig honger, bereik voedselzekerheid en verbeterde voeding en promoot duurzame landbouw
3) Verzeker een goede gezondheid en promoot welzijn voor alle leeftijden
4) Verzeker gelijke toegang tot kwaliteitsvol onderwijs en bevorder levenslang leren voor iedereen
5) Bereik gendergelijkheid en empowerment voor alle vrouwen en meisjes
6) Verzeker toegang tot duurzaam beheer van water en sanitatie voor iedereen
7) Verzeker toegang tot betaalbare, betrouwbare, duurzame en moderne energie voor iedereen
8) Bevorder aanhoudende, inclusieve en duurzame economische groei, volledige en productieve tewerkstelling en waardig werk voor iedereen
9) Bouw veerkrachtige infrastructuur, bevorder inclusieve en duurzame industrialisering en stimuleer innovatie
10) Dring ongelijkheid in en tussen landen terug
11) Maak steden en menselijke nederzettingen inclusief, veilig, veerkrachtig en duurzaam
12) Verzeker duurzame consumptie- en productiepatronen
13) Neem dringend actie om klimaatverandering en haar impact te bestrijden
14) Behoud en maak duurzaam gebruik van oceanen, zeeën en maritieme hulpbronnen
15) Bescherm, herstel en bevorder het duurzaam gebruik van ecosystemen op het vasteland, beheer bossen en wouden duurzaam, bestrijd woestijnvorming, stop landdegradatie en draai het terug en roep het verlies aan biodiversiteit een halt toe
16) Bevorder vreedzame en inclusieve samenlevingen met het oog op duurzame ontwikkeling, verzeker toegang tot justitie voor iedereen en bouw op alle niveaus doeltreffende, verantwoordelijke en toegankelijke instellingen uit
17) Versterk de implementatiemiddelen en revitaliseer het wereldwijd partnerschap voor duurzame ontwikkeling

We hebben hier geen ruimte om alle doelstellingen in detail te bespreken, maar de bovenstaande punten spreken in ieder geval tot de verbeelding: hoe zouden we dit allemaal voor elkaar moeten krijgen? En in het kader van deze cursus: hoe zou een inhoudsanalyse hieraan kunnen bijdragen?

Voorbeeld
Natuurlijk kunnen we genderongelijkheid (doel 5) niet direct bereiken met een kleine studie, maar we kunnen het probleem wel verder in kaart brengen door bijvoorbeeld te laten zien hoe mannen en vrouwen in vergelijkbare situaties toch heel anders behandeld worden in de media. Door hierover te rapporteren, maken we het probleem concreet. Dat maakt de ongelijkheid bespreekbaar, omdat we het niet meer hebben over discriminatie in abstracte zin, maar over voorbeelden uit het echte leven die ervaren worden door echte mensen van vlees en bloed.

2.5.4 ERVARINGSDESKUNDIGEN

Een laatste manier om je onderzoek praktische relevantie te geven is om ervaringsdeskundigen te betrekken in het proces. Wanneer je data uit of over een specifieke gemeenschap wil bestuderen, kan het helpen om bij een lid van die gemeenschap te controleren of er eigenlijk behoefte is aan jouw onderzoek, en of het perspectief dat je hanteert goed aansluit bij de manier waarop zij hun situatie ervaren. In het Engels wordt in deze context vaak de slogan *Nothing about us without us* gehanteerd om het belang aan te geven van ervaringsdeskundigen.

Voorbeeld: dikke mensen
Het is goed om je te realiseren dat je meestal niet de eerste bent die geïnteresseerd is om een bepaalde gemeenschap te bestuderen, en niet iedereen heeft even positieve ervaringen met buitenstaanders. Daarom is het belangrijk om voorzichtig en respectvol om te gaan met je doelgroep. Een recente studie van Payne et al. (2023) geeft een voorbeeld van hoe je dit kunt doen. De auteurs (die deels naar eigen zeggen ook deel uitmaken van de doelgroep) beschrijven "hoe je op ethische wijze dikke mensen kunt betrekken in je onderzoek." Dat begint al bij (1) **het taalgebruik**: de term 'dik' heeft de voorkeur voor veel van hun participanten boven meer normatieve beschrijvingen zoals "mensen met overgewicht" of medische aanduidingen ("obese mensen"). Het is goed om mensen uit je doelgroep te vragen hoe ze zichzelf zouden omschrijven, en die voorkeurstermen te hanteren. Daarnaast is (2) de **positionaliteit van de onderzoeker** ook belangrijk: hoe verhoud jij je tot de doelgroep, en op basis van welke motivatie ben je van plan de doelgroep te bestuderen? In hoofdstuk 8: *Ethiek* gaan we dieper in op positionaliteit en reflexiviteit. Een volgende punt dat de auteurs aandragen betreft (3) de **aannames over de doelgroep en hun problematiek**: in hoeverre komen jouw aannames overeen met de ervaringen en ideeën van mensen uit je doelgroep? De auteurs beschrijven hoe sommige mensen aannemen dat alle dikke mensen willen of zouden moeten afslanken, terwijl dat niet altijd het geval is. We leven in een diverse wereld waarbij niet iedereen dezelfde lichaamsbouw heeft. Een alternatief perspectief komt vanuit de *Health at any size*-beweging, die gezondheid en gezonde voeding centraal stelt, zonder dat mensen noodzakelijkerwijs moeten afvallen. Ten slotte moedigen de auteurs iedereen aan om (4) de **wensen en behoeften van de doelgroep** centraal te stellen door ervaringsdeskundigen bewust te betrekken in het onderzoeksproces.

2.6 ETHISCHE BEZWAREN

Hierboven hebben we gekeken naar de vragen die je je kunt beantwoorden met een inhoudsanalyse, en de manieren waarop je die vragen kunt onderbouwen. Met andere woorden: redenen om een onderzoek *wel* uit te voeren. Maar het is ook belangrijk om de omgekeerde vraag serieus te nemen: zijn er redenen om het onderzoek *niet* uit te voeren, of om op zijn minst terughoudend te zijn in het uitvoeren van de studie en het publiceren van de resultaten? Die redenen liggen vaak in het domein van de ethiek, een onderwerp waar ook een hoofdstuk aan gewijd is. Hier noemen we alvast één belangrijk concept: *dual use*.

2.6.1 DUAL USE

De term *dual use* verwijst naar het idee dat een product voor meerdere doeleinden gebruikt kan worden. Specifiek gaat het vaak om producten die zowel maatschappelijke als militaire toepassingen hebben, en meer algemeen wordt de term gebruikt voor producten die zowel voor positieve als negatieve doeleinden gebruikt kunnen worden. Een paar algemene voorbeelden:

- GPS-technologie wordt zowel gebruikt voor navigatiesystemen als voor de aansturing van raketten. Daarnaast worden GPS-trackers (nuttig tegen diefstal en om verloren spullen te vinden) soms gebruikt om mensen te stalken.
- AI-systemen voor automatische beeld- en spraakherkenning kunnen gebruikt worden om makkelijker afbeeldingen te kunnen zoeken of voor automatische ondertiteling, maar ook mensen te surveilleren en te bespioneren.

Vergelijkbare problemen spelen bij inhoudsanalyses. Zeker bij communicatiewetenschap, waar we vaak berichten tussen verschillende groepen mensen bestuderen, en juist bij maatschappelijk belangrijke onderwerpen.

Voorbeeld: calls to action

Rogers et al. (2019) keken naar zogenaamde *calls to action*; oproepen aan medeburgers op sociale media om in actie te komen. Door handmatig zulke calls te identificeren, konden ze bestuderen of er een verband is tussen het aantal oproepen op sociale media, en de daadwerkelijke opkomst bij protesten tegen de overheid. De auteurs schrijven in hun artikel dat ze bewust een aantal details hebben weggelaten uit hun studie, omdat ze niet willen dat overheden (specifiek: de Russische overheid) hun werk gebruiken om activisten op te sporen en te censureren.

HOOFDSTUK 3
Dataverzameling en sampling

3.1 INLEIDING

Zonder data kun je ook geen inhoudsanalyse uitvoeren. Maar waar komt die data vandaan? Hoeveel data heb je eigenlijk nodig voor een betrouwbare inhoudsanalyse? Hoe verzamel je die data eigenlijk? En mag je zomaar alles verzamelen? In dit hoofdstuk gaan we dieper in op deze en meer vragen over dataverzameling. Na het lezen van dit hoofdstuk weet je hoe je een dataverzamelingsplan moet opstellen, zodat je dat kunt toevoegen aan je onderzoeksvoorstel.

3.2 WAT KUN JE VERZAMELEN?

Inhoudsanalyse is een algemene methode om alle mogelijke soorten data te coderen. Zo kun je bijvoorbeeld denken aan:

- **Afbeeldingen**: foto's in kranten of nieuwswebsites, stripverhalen, kindertekeningen, reclamefolders, grafieken, posts op sociale media.
- **Tekst**: gesprekken met chatbots, krantenartikelen, boeken, ondertiteling, tijdschriften, gespreksverslagen, posts op sociale media.
- **Video**: TV-uitzendingen, promotiefilms, online nieuwsberichten, filmpjes van sociale media
- **Geluidsopnames**: radio, podcasts, gesprekken.
- **Multimedia**: combinaties van bovenstaande.

Voor de dataverzameling maakt het eigenlijk niet uit wat je verzamelt; dezelfde principes gelden voor alle soorten data.[1] Wel is het zo dat je goed moet nadenken over de opslag van de data. Video's nemen bijvoorbeeld meer ruimte in beslag dan tekst, waardoor het niet altijd haalbaar is om alle data lokaal op te slaan.[2] Verder kun je de meeste data op een vergelijkbare manier coderen, maar voor sommige studies heb je gespecialiseerde software nodig. Bijvoorbeeld als het belangrijk voor je is om te coderen waar in een afbeelding verschillende visuele elementen zich bevinden ten opzichte van elkaar. Dat is lastig om in woorden te omschrijven, maar makkelijk om te markeren in een afbeelding.

3.3 WAAR HAAL JE DE DATA VANDAAN?

Als je weet wat voor soort data je wil gaan analyseren, dan kun je twee dingen doen: je kunt er natuurlijk voor kiezen om zelf data te verzamelen, of je kunt ervoor kiezen om gebruik te maken van een eerder samengesteld corpus.

3.3.1 WAT IS EEN CORPUS?

De term *corpus* (meervoud: corpora) wordt binnen de taalwetenschap gebruikt voor datasets die bestaan uit een grote verzameling teksten. Deze teksten zijn vaak (maar niet altijd) voorzien van verschillende soorten metadata (dat wil zeggen: informatie over de teksten uit het corpus). In het volgende hoofdstuk (Taal en inhoudsanalyse) leer je meer over verschillende analyseniveaus die er bestaan in de taalwetenschap. Corpora en andere datasets ontstaan niet zomaar, maar worden samengesteld vanuit een bepaalde gedachte, en vaak ook met een bepaald doel. Als je hier meer over wil lezen, dan is het werk van Sinclair (2004) een aanrader.

3.3.2 WAAR VIND JE CORPORA?

Wetenschappelijke corpora worden beschreven in twee soorten publicaties:

1) Allereerst zijn er studies die er volledig op gericht zijn om een algemeen corpus samen te stellen en beschikbaar te maken. Zulke studies worden vaak gepubliceerd in gespecialiseerde tijdschriften (zoals *Language Resources and Evaluation, Behavior Research Methods, Corpora*) of conferenties (bijvoorbeeld de *International Conference on Language Resources and Evaluation*; afgekort als LREC). Afhankelijk van het vakgebied gebruiken auteurs de term 'corpus' of het meer algemene 'dataset' om hun verzameling teksten, video's, afbeeldingen, of geluidsopnamen te beschrijven. Sommige auteurs kiezen ervoor om hun data zelf te

1 Voor dit hoofdstuk nemen we aan dat je goed hebt nagedacht over de aard van de data die je wil verzamelen. Hoewel dezelfde principes gelden voor de *verzameling* van de data, zijn er wel belangrijke verschillen tussen de analyse van bijvoorbeeld gesproken en geschreven taal.

2 Bovendien bevatten video's vaak ook meer persoonlijke informatie, waardoor het niet altijd wenselijk is om die data te verzamelen. Zie Hoofdstuk 8 over Ethiek en Hoofdstuk 9 over Datamanagement voor verdere discussie.

archiveren, terwijl anderen ze overdragen aan een nationale instantie. Een voorbeeld van zo'n instantie in Nederland is het *Instituut voor de Nederlandse Taal,* dat bijvoorbeeld het *Corpus Gesproken Nederlands* en het *Eindhovencorpus* beheert.

2) Vervolgens zijn er ook studies die een corpus samenstellen om zelf iets te kunnen zeggen over die data. Je vindt zulke studies in alle gangbare tijdschriften binnen de communicatiewetenschap. Waar je de gebruikte gegevens kunt downloaden verschilt per auteur en per tijdschrift. Sommige tijdschriften publiceren de gebruikte data als *Supplementary materials* op hun website, terwijl andere tijdschriften het delen van de data geheel aan de auteurs overlaten. Als je uiteindelijk niet kunt vinden waar de onderzoeksgegevens te downloaden zijn, kun je ook overwegen om een bericht te sturen naar de auteurs met het verzoek om het corpus te delen.

3.3.3 ZELF DATA ZOEKEN OM TE VERZAMELEN

Later in dit hoofdstuk gaan we dieper in op de praktische details, maar het is goed om alvast een intuïtie te hebben over het zelf verzamelen van data. Alles begint met een heldere definitie van de populatie: wie of wat hoort er wel of niet bij? Hierbij ga je aan het begin van je onderzoek vaak heen en weer tussen de data en de onderzoeksvraag.

Voorbeeld: groene influencers. Stel je voor dat je geïnteresseerd bent in 'groene influencers' en de manier waarop ze het klimaatprobleem presenteren op sociale media. (Bijvoorbeeld: als persoonlijke verantwoordelijkheid, of als gedeelde verantwoordelijkheid?) Dan is jouw eerste taak om die groep goed af te bakenen. Dat begint met het *definiëren* van jouw terminologie: wat bedoel je eigenlijk met de term 'groene influencer?' Vervolgens kun je die definitie verder *operationaliseren*: Op welke sociale media ga je kijken? Wie zitten er wel of niet in die groep? Is het eigenlijk belangrijk dat je een homogene groep bestudeert, of zoek je juist meer diversiteit? Voor het beantwoorden van die vragen kan het nuttig zijn om een aantal voorbeeldprofielen op te zoeken om te kijken of ze voldoen aan jouw definitie.

Het voorbeeld hierboven gaat over het gebruik van data van sociale media, maar daarnaast zijn er nog vele andere kanalen waar je gebruik van kunt maken. Denk aan kranten (relatief eenvoudig te verkrijgen via databases zoals *Nexis Uni, Delpher, Factiva*, en *Newsbank* die beschikbaar zijn via de universiteitsbibliotheek), websites, podcasts, of bijvoorbeeld de Handelingen van de Tweede Kamer (met vrijwel letterlijke gespreksverslagen van ieder debat uit de afgelopen decennia). Overal om je heen zie je wel communicatie-uitingen. Voor alle bronnen geldt hetzelfde principe: kijk om je heen, en bedenk wat je met dat soort data zou kunnen doen. Denk je dat je met zulke data een zinvol antwoord kunt geven op jouw onderzoeksvraag? Of misschien een aangepaste onderzoeksvraag? Als het antwoord "ja" is, kunnen we verder gaan kijken naar de praktische haalbaarheid.

3.4 HOEVEEL DATA MOET JE VERZAMELEN?

In het ideale geval verzamel je volgens Neuendorf gewoon alles (een census). Bijvoorbeeld alle afleveringen van het televisieprogramma *De wereld draait door* dat jarenlang op de Nederlandse televisie te zien was, of alle wekelijkse persconferenties van de premier van Nederland. Als dat niet kan, moet je in ieder geval genoeg data verzamelen om gegronde conclusies te trekken. Daarbij moet je zorgen voor een representatieve steekproef. Hierover hebben we het later in dit hoofdstuk bij de paragraaf over Sampling.

3.4.1 WAAROM EIGENLIJK EEN CENSUS? EN REDENEN OM DAT NIET TE DOEN

Het lijkt heel logisch om gewoon alle data te verzamelen, omdat je dan niet meer hoeft na te denken over sampling. Als je alles verzamelt, kun je de hele populatie bestuderen, waarbij de resultaten een volledige beschrijving geven van de data. Maar er zijn goede redenen om te twijfelen aan het advies van Neuendorf (2017) om te werken met alle data die je tot je beschikking hebt.

Als een steekproef al genoeg is om gegronde conclusies te trekken, waarom zou je dan nog meer data gaan verzamelen en coderen? Zeker het coderen kan erg tijdrovend zijn. Daarbij gaat het niet alleen om jouw tijd als onderzoeker (en als mens!), maar ook om geld dat (in ieder geval aan de universiteit) door de maatschappij in ons is geïnvesteerd. Dat maakt de vraag hoeveel data je moet verzamelen extra belangrijk. Tegelijkertijd is die vraag niet eenvoudig te beantwoorden, omdat het antwoord afhangt van vele factoren (bijvoorbeeld: de hoeveelheid variatie in de data). Een beperkende factor is in ieder geval de praktische haalbaarheid, die we hieronder bespreken.

3.4.2 PRAKTISCHE HAALBAARHEID

Praktische haalbaarheid kunnen we opsplitsen in twee onderdelen: de dataverzameling en het codeerproces. Hoeveel tijd je kwijt bent aan deze onderdelen is voor ieder project weer anders, maar het is vaak wel mogelijk om een goede inschatting te maken door eerst te meten hoeveel tijd het kost om een kleine hoeveelheid data te verzamelen en te coderen, en die gegevens vervolgens te extrapoleren naar grotere hoeveelheden data.

Voorbeeld: grafieken coderen op basis van visuele toegankelijkheid.

In een recente masterscriptie heeft Probst (2023) gekeken naar de visuele toegankelijkheid van grafieken in gepubliceerde onderzoeksartikelen. Dat betekent bijvoorbeeld dat er gekeken werd naar het contrast tussen lijngrafieken en de achtergrond, het gebruik van redundantie (bijvoorbeeld zowel kleur als patronen om verschillende categorieën van elkaar te onderscheiden). De figuren waren eenvoudig automatisch te verzamelen, waardoor Probst een Excelbestand had met alle figuren uit alle artikelen die gepresenteerd zijn op een wetenschappelijke conferentie. De uitdaging lag dus bij het coderen: hoeveel tijd zou dat kosten?

Iedere inschatting begint met één of meerdere aannames, bijvoorbeeld: iedere grafiek kost me 2 minuten om te coderen. Met deze aanname kun je verder rekenen, waardoor we het volgende kunnen concluderen:

- 60 minuten per uur, gedeeld door 2 minuten per grafiek, is 30 grafieken per uur.
- Als je 8 uur per dag rekent, dan zijn dat 240 grafieken per dag.
- Met vijf werkdagen codeer je 5 * 240 = 1.200 grafieken per week, bij full-time coderen.

Op papier klinkt dat misschien haalbaar, maar de werkelijkheid is dat een planning als deze niet haalbaar is. Het is absoluut niet gezond om alleen maar te coderen; het zou betekenen dat je bijna de hele week in dezelfde houding achter de computer zit, waarbij je dezelfde repetitieve taken uitvoert. We hebben dus *realistische* aannames nodig. Bijvoorbeeld:

- Ieder uur neem ik 10 minuten pauze. (Om RSI te voorkomen wordt vaak aangeraden om ieder half uur even 5 minuten pauze te nemen.)
- Per dag kan ik twee uur coderen voordat ik weer wat anders moet doen.

Wanneer we daarmee verder rekenen, komen we op heel andere getallen uit:

- 50 productieve minuten per uur, betekent 25 grafieken per uur.
- Met twee uur per dag kan ik 50 grafieken per dag coderen.
- Met vijf dagen per week kan ik 250 grafieken per week coderen.

Natuurlijk verschilt het ook van situatie tot situatie hoeveel je kan coderen per dag. Bijvoorbeeld: Hoe repetitief/afwisselend is het coderen? Daarnaast maakt je planning ook uit: twee uur achter elkaar is best veel, maar misschien lukt het wel om twee keer 1,5 uur te coderen op een dag. Hoe dan ook, het is essentieel dat je een goede planning maakt en je daar ook aan houdt. Inhoudsanalyses kun je niet uitstellen tot een week voor de deadline.

Als je een planning hebt gemaakt ben je nog niet klaar. Je kunt wel *denken* dat het maar twee minuten kost om een item te coderen, maar daarmee weet je het nog niet zeker. Dit is een van de vele redenen om een pilotstudie uit te voeren voordat je begint met het echte werk.

3.4.3 VARIATIE, EFFECTGROOTTE, EN STEEKPROEFGROOTTE

Naast praktische overwegingen speelt ook het karakter van de data een rol bij de sampling.[3] In het algemeen geldt de stelregel: hoe groter de variatie, des te groter dient je steekproef te zijn (Williams & Williams, 2020). Als je bijvoorbeeld wil bestuderen of bepaalde eigenschappen van berichten op sociale media meer *engagement* (bijvoorbeeld het aantal likes en comments) opleveren, dan heb je relatief veel data nodig, omdat het aantal likes en comments behoorlijk fluctueert. Daarnaast wordt de vereiste samplegrootte ook beïnvloed door de verwachte *effectgrootte*: hoe groot verwacht je dat de verschillen zullen zijn tussen de groepen die je wil vergelijken? Als het effect maar klein is, dan heb je een grotere steekproef nodig om dat effect aan te tonen dan wanneer je verwacht dat er een flink effect zal zijn. Het probleem is dat het vaak lastig is om de precieze variatie en waarschijnlijke effectgrootte in te schatten.[4] Voor de groepsopdracht bij dit boek kun je de samplegrootte baseren op eerder onderzoek en een berekening van de praktische haalbaarheid, maar Lakens (2022) geeft terecht aan dat deze motivaties geen garantie geven voor een samplegrootte waaruit je harde conclusies kunt trekken. Het bespreken van alle verschillende uitdagingen (en oplossingen) bij het bepalen van de samplegrootte gaat te ver voor dit boek, maar Lakens (2022) geeft een goed overzicht van de verschillende motivaties en benaderingen om een samplegrootte te kiezen.

3.5 SAMPLING

Sampling is een ander woord voor het nemen van een steekproef. De samenstelling van de steekproef hangt af van de manier waarop je de selectie maakt. Allereerst kunnen we een onderscheid maken tussen *probability sampling* (op basis van kans) en *non-probability sampling* (alternatieve benaderingen, die vaak minder generaliseerbaar zijn maar nog steeds relevante inzichten kunnen opleveren). Waarschijnlijk heb je de sampling-strategieën hieronder eerder in een statistiek- of methodologiecursus langs zien komen, dus we zullen ze slechts kort bespreken.

---

3 Bij *kwalitatief* onderzoek wordt vaak gesproken over de *saturatie* (ook wel: verzadiging) die optreedt bij het verzamelen van de data: op een gegeven moment voegen nieuwe documenten, interviews, of focusgroepen steeds minder toe ten opzichte van de eerder verzamelde data. Je ziet dan vooral meer van hetzelfde, en dat levert geen nieuwe inzichten meer op. Op dat moment maken onderzoekers de *inschatting* (je weet het natuurlijk nooit helemaal zeker) dat het geen zin heeft om meer data te verzamelen (Saunders et al., 2018). Het moment van verzadiging hangt af van de hoeveelheid variatie in de data: hoe groter de onderlinge variatie, des te meer data moet je verzamelen voordat saturatie optreedt. Bij kwantitatief onderzoek is saturatie niet relevant, omdat men dan niet geïnteresseerd is in het beschrijven van verschillende thema's, maar in het kwantificeren van reeds bekende categorieën.

4 Zie bijvoorbeeld Albers & Lakens (2018) voor een bespreking van de problemen die optreden wanneer je de samplegrootte wil bepalen aan de hand van een pilotstudie.

3.5.1 PROBABILITY SAMPLING

Zoals gezegd hebben alle vormen van probability sampling met kans te maken, en voor ieder item in de populatie kun je berekenen wat de kans is dat deze wordt geselecteerd in je steekproef. De eenvoudigste methode is **Simple Random Sampling**, waarbij je items volledig willekeurig selecteert. Dat kan door lootjes uit een hoed te trekken, maar natuurlijk ook met een automatische *randomizer* (bijvoorbeeld in Python[5], of in je favoriete spreadsheetprogramma[6]). Simple random sampling kent ook een aantal nadelen. Zo vereist random sampling bijvoorbeeld dat je toegang hebt tot de volledige populatie, en kan het voorkomen dat bepaalde sub-populaties niet voldoende vertegenwoordigd zijn in je uiteindelijke sample.

Stratified random sampling zorgt voor een oplossing van het probleem van ontbrekende subpopulaties. Het idee is dat je niet gelijk een willekeurige steekproef neemt onder de gehele populatie, maar dat je eerst verschillende strata (subpopulaties of deelgroepen) onderscheidt, en binnen iedere deelgroep een steekproef neemt zodat het volledige sample representatief is voor de gehele populatie, inclusief de subpopulaties. Het grootste nadeel aan stratified random sampling is dat het een goed overzicht van de populatie vereist, met een lijst van alle leden van de verschillende subpopulaties.

Een alternatief voor random sampling is **Systematic Sampling**. Hierbij kies je een willekeurig startpunt en een intervalwaarde N, en vervolgens selecteer je ieder Nde item dat je tegenkomt. De intervalgrootte wordt afgestemd op de grootte van de populatie, dus stel je voor dat er 500 uitzendingen zijn van een televisieprogramma, maar je wil maar 100 afleveringen coderen, dan kies je een interval van 5 afleveringen zodat je uiteindelijk een mooie spreiding hebt van geselecteerde afleveringen. Het nadeel van systematic sampling is dat er een bias kan ontstaan in je selectie, bijvoorbeeld omdat er een vaste roulatie is van presentatoren in het televisieprogramma, en je door de keuze van de intervalgrootte altijd afleveringen met dezelfde presentator treft.

Bij **cluster sampling** groepeer je alle berichten waarin je geïnteresseerd bent in clusters, waarna je vervolgens willekeurig een cluster selecteert om verder te analyseren. Het grote voordeel van cluster sampling is dat de dataverzameling geconcentreerd is in een beperkte periode (of een beperkt aantal periodes/locaties), waardoor je als onderzoeker veel tijd en geld kunt besparen. Neuendorf (2017) geeft de studie van C. A. Lin (1997) als voorbeeld. Voor deze studie heeft de auteur een week lang video-opnames gemaakt van alle reclames die tijdens *prime time* (8 tot 11 uur 's avonds) uitgezonden werden op de drie belangrijkste televisienet-

5 Gebruik hiervoor bijvoorbeeld de sample()-functie uit de random-module. Je kunt deze procedure reproduceerbaar maken door vooraf een zogenaamde *random seed* in te stellen. Bijvoorbeeld random.seed(1). Wanneer je dat doet, krijg je altijd dezelfde (pseudo-)willekeurige selectie.

6 In Excel kun je met de =RAND()-functie een willekeurig kommagetal tussen 0 en 1 laten genereren. Wanneer je de rijen sorteert op de kolom met willekeurig gegenereerde getallen, kun je de eerste N rijen selecteren. Deze procedure is niet reproduceerbaar, maar als je dat wil zijn er alternatieven voor de RAND-functie.

werken in de Verenigde Staten. Omdat de week willekeurig is gekozen, zou deze representatief moeten zijn voor de reclames die normaal op de Amerikaanse televisie vertoond moeten worden. Daarbij hoort natuurlijk wel de kanttekening dat sommige weken kunnen afwijken van de rest, bijvoorbeeld doordat er een groot evenement plaatsvindt waarbij iedereen wil aanhaken. (Denk bijvoorbeeld aan het WK voetbal, waarbij de reclames opeens allemaal om voetbal lijken te draaien.) Hier heb je dus ook een rol als onderzoeker om ervoor te waken dat je niet toevallig de verkeerde periode hebt uitgekozen.

Ten slotte is het ook mogelijk om bovenstaande sampling-strategieën te combineren, via **multistage sampling**. De naam zegt het eigenlijk al: in meerdere stappen, bijvoorbeeld door eerst een aantal clusters te selecteren, en vervolgens binnen die clusters een gestratificeerde, willekeurige steekproef te nemen. Of je kunt cluster sampling toepassen binnen eerder geïdentificeerde clusters (**multistage cluster sampling**). Omdat er eindeloos veel mogelijkheden zijn om data te verzamelen, is er niet altijd een passende term voorhanden die je kunt gebruiken om jouw sampling-strategie te omschrijven. Het belangrijkste is dan dat je helder communiceert hoe de steekproef tot stand is gekomen, zodat het onderzoek herhaalbaar is. De categorieën die we hier beschrijven zijn vooral praktische hulpmiddelen om je te helpen nadenken over de juiste methode voor jouw onderzoek.

3.5.2 NON-PROBABILITY SAMPLING

Voor non-probability sampling is de relatie tussen de steekproef en de populatie minder eenvoudig te bepalen, en daardoor is het ook lastiger om te zeggen of de conclusies uit je steekproef ook gelden voor de gehele populatie.

Convenience sampling is wellicht de bekendste methode: je kiest de gemakkelijk verkrijgbare gevallen. Bijvoorbeeld met studenten als participanten in experimentele studies. Maar 'alle psychologieboeken in de bibliotheek van de universiteit' is ook een voorbeeld van een *convenience sample* als je iets wil zeggen over psychologieboeken in het algemeen.[7] Voor een representatieve steekproef zijn er meer criteria dan alleen beschikbaarheid, maar het is wel een eenvoudige, goedkope en snelle manier om aan data te komen.

Bij **purposive of judgment sampling** beoordeel je als onderzoeker zelf of een item in het sample opgenomen of niet, op basis van een of meer criteria. Bijvoorbeeld: wanneer je geïnteresseerd bent in de posts van influencers binnen een specifiek genre, is er misschien geen lijst voorhanden die je kunt gebruiken om auteurs te selecteren. Een optie is om dan zelf te bepalen wie in aanmerking komt.

Ten slotte noemen we nog **quota sampling**: dit is eigenlijk de niet-kansgedreven tegenhanger van stratified sampling. Hierbij selecteer je zelf items voor je dataset, waarbij je ervoor

7 Als je iets wil zeggen over de collectie van de universiteit, dan is het geen convenience sample maar een census van alle psychologieboeken die we in Tilburg hebben.

zorgt dat de numerieke verhoudingen tussen de items in je dataset overeenkomen met de verhoudingen in de populatie.

3.5.3 RAPPORTAGE

Verschillende auteurs hanteren verschillende definities voor dezelfde termen. De tekst hierboven volgt de terminologie van Treadwell (2023), maar het kan zijn dat andere auteurs net iets anders bedoelen wanneer ze het bijvoorbeeld hebben over *stratified sampling*. Daarom volstaat het in je verslag niet om alleen te zeggen welke vorm van sampling je hebt gebruikt, maar moet je ook (kort) uitleggen wat je precies hebt gedaan.

3.5.4 NOG EVEN OVER POPULATIES

Bij het praten over sampling hebben veel mensen (en ikzelf heb hier soms ook last van!) de neiging om te vervallen in algemeenheden over "de populatie." Maar het is goed om je af te vragen wie of wat dat eigenlijk zijn. Zijn dat bijvoorbeeld influencers, gebruikers van sociale media, jongeren, Nederlanders, Europeanen, Westerlingen, of mensen in het algemeen? De generaliseerbaarheid van je studie naar "de populatie" is afhankelijk van hoe je die populatie definieert. (Ook goed om te bedenken: waar de populatie in experimentele studies beperkt is tot levende wezens, kunnen populaties voor een inhoudsanalyse ook bestaan uit levenloze dingen, zoals boeken of reclames.)

Daarnaast kun je je afvragen wat het doel is van je studie. Wil je eigenlijk wel generaliseerbare inzichten opdoen, of gaat het je meer om het beter begrijpen van een specifieke casus? Soms is het voldoende om te laten zien dat iets bestaat, zonder verdere claims te willen doen over andere gevallen. Bijvoorbeeld als een theorie zegt dat je als bedrijf beter geen openbare excuses kunt aanbieden voor je eigen fouten omdat dat mogelijk imagoschade kan opleveren, is het erg nuttig om te weten dat dit in de praktijk toch regelmatig voorkomt (Van Hooijdonk & Liebrecht, 2021).

3.6 SAMPLING EN SOCIALE MEDIA

Representativiteit van je sample hangt af van de manier waarop je de data verzamelt. In het ideale geval zou je beginnen met een uitputtende lijst die de gehele populatie beschrijft. (Dat wordt ook wel een sampling frame genoemd.) Bijvoorbeeld: alle debatten in de Tweede Kamer, waarvan de notulen allemaal netjes gearchiveerd zijn. Voor sociale media is dat een stuk lastiger, omdat zo'n lijst vaak niet bestaat, of in ieder geval niet zomaar beschikbaar is. McGrady et al. (2023) bespreken de uitdagingen die komen kijken bij het beschrijven van de vele video's die op YouTube staan; alleen al het tellen van het aantal video's is een probleem,

laat staan het maken van een volledige lijst.[8] Dat bemoeilijkt de discussie over wat nou eigenlijk een representatieve steekproef is.

3.6.1 KRITISCH KIJKEN NAAR ZOEKRESULTATEN

In plaats van een losse lijst van alle posts of video's op een platform, kun je ook een lijst met zoektermen opstellen en de zoekfunctie gebruiken om een voorselectie te maken van relevante items. Het probleem met sociale media (en eigenlijk vrijwel alle websites) is dat ze gesloten zijn; we weten simpelweg niet hoe de website in elkaar zit, en hoe de zoekfunctie werkt. Dat heeft ook implicaties voor de representativiteit van de zoekresultaten: als je niet weet hoe de zoekresultaten geselecteerd en geordend zijn, kun je niet zomaar generaliseren van de eerste pagina naar alle relevante items op het platform. Daarom is het belangrijk om kritisch te kijken naar de zoekresultaten. Bijvoorbeeld:

1) Zie je een patroon in de manier waarop de resultaten geordend zijn?
2) Gebruik dezelfde zoektermen in verschillende situaties:
   a) Krijg je dezelfde resultaten op verschillende computers?
   b) Krijg je dezelfde resultaten op verschillende tijdstippen?
   c) Krijg je dezelfde resultaten als je in/uitgelogd bent?
   d) Krijg je dezelfde resultaten als iemand anders die hetzelfde probeert?

Deze vragen doen denken aan het idee van de *filter bubbel*.[9] Pariser (2011) beschrijft hoe gepersonaliseerde zoekresultaten ervoor kunnen zorgen dat wij een gekleurd beeld van de werkelijkheid krijgen doordat we voornamelijk berichten krijgen die aansluiten op ons eigen perspectief, maar die niet het volledige beeld geven van wat er in de wereld aan de hand is. Er is in de jaren daarna veel discussie geweest over het bestaan van filter bubbels, met tegenstrijdige resultaten (Michiels et al., 2022; Moeller et al., 2018).[10] Belangrijk voor ons is in ieder geval dat je je bewust bent van het feit dat zoekresultaten gekleurd *kunnen* zijn, en dat je goed nadenkt over de implicaties hiervan op de generaliseerbaarheid van je onderzoek.

8 De auteurs hebben het aantal video's benaderd door uit te proberen hoeveel mogelijke video-*identifiers* (unieke combinaties van cijfers en letters die horen bij een specifieke video) al gekoppeld zijn aan een video, en te kijken hoeveel pogingen succesvol waren. Door die getallen te extrapoleren naar alle mogelijke identifiers, konden ze een ruwe schatting geven. Volgens die schatting stonden er eind 2022 zo'n 10 miljard video's op YouTube, en dat aantal is groeiende.

9 Een verwante term is *echo chambers*. Deze term verwijst naar groepen of netwerken binnen sociale media waar door een gebrek aan diversiteit maar één mening te horen is die constant herhaald wordt. Een bespreking van dit idee valt buiten het bereik van dit werkboek.

10 Veel van de onderzoeksresultaten zijn tegenstrijdig of lastig met elkaar te vergelijken omdat er verschillende definities en operationalisaties van filter bubbels gehanteerd worden.

3.6.2 KRITISCH KIJKEN NAAR HET PLATFORM

Gillespie (2013) heeft een overzicht gemaakt van zes maatschappelijk relevante dimensies om algoritmes te analyseren en bespreken. Deze dimensies zijn ook nuttig om na te denken over representativiteit. Hieronder staan de drie meest relevante dimensies, met originele Engelse termen tussen haakjes. Definities zijn vertaald uit het oorspronkelijke artikel, gevolgd door een korte toelichting, maar het is de moeite waard om het oorspronkelijke artikel helemaal te lezen.

1) **Inclusie en uitsluiting** (*patterns of inclusion*): welke informatie wordt er geïndexeerd (dat wil zeggen: opgenomen in het platform om doorzocht te kunnen worden), wat wordt er uitgesloten, en hoe wordt de data bewerkt om opgenomen te kunnen worden? Bijvoorbeeld: wat voor *content* mag je eigenlijk plaatsen op sociale media? Alles wat tegen de richtlijnen ingaat wordt verwijderd (dit noemt Gillespie *exclusion*).[11] Soms wordt het plaatsen van bepaalde informatie wel toegestaan, maar zorgt het systeem ervoor dat die informatie door veel minder mensen gezien wordt (dit noemt Gillespie *demotion*). Verder geldt nog dat alle informatie in een specifieke vorm gegoten moet worden. Alles wat daar niet in past, komt niet op het platform terecht.
2) **De evaluatie van relevantie** (*the evaluation of relevance*): de criteria waarmee algoritmen bepalen wat relevant is, hoe die criteria voor ons verborgen blijven en hoe het systeem het gevolg is van politieke keuzes over wat legitieme of gepaste kennis is. Er wordt weleens gezegd dat computers objectief zijn. Dat klopt: ze doen precies wat je zegt dat ze moeten doen, maar achter ieder systeem zitten ontwerpkeuzes die zelf niet objectief zijn. Want hoe bepaal je nu eigenlijk wat bovenaan de zoekresultaten moet staan? Iedere keuze die een bedrijf hier maakt, is het gevolg van wat dat bedrijf zelf belangrijk vindt.[12]
3) **Verstrengeling met het gebruik** (*entanglement with practice*): hoe gebruikers zich in hun gedrag voegen naar de algoritmen waarvan ze afhankelijk zijn, en hoe gebruikers algoritmen kunnen veranderen in terreinen voor politieke strijd, soms zelfs om de politiek van het algoritme zelf te bevragen. In het kort komt deze dimensie erop neer dat we technologie niet los kunnen zien van de gebruikers van die technologie, en hoe zij gebruik maken van (werkelijke of ingebeelde) eigenschappen van die technologie om hun doelen

11 En waar komen die richtlijnen eigenlijk vandaan? De meeste grote technologiebedrijven komen uit de Verenigde Staten, waar ze andere normen en waarden hebben dan in Nederland. (Denk bijvoorbeeld aan discussies over seksualiteit.) Daarnaast worstelen ze vaak met het probleem dat ze gebruikers hebben over de hele wereld, waarbij ieder land zijn eigen regels heeft over wat er wel of niet is toegestaan.

12 Deze keuzes zijn bovendien ook niet statisch: online platforms blijven hun algoritmen doorlopend aanpassen, en dat kan invloed hebben op je onderzoek. Een voorbeeld hiervan is de studie van Guess et al. (2023) in het wetenschappelijke tijdschrift *Science*. De auteurs onderzochten posts op Facebook tijdens de Amerikaanse verkiezingen van 2020. Bagchi et al. (2024) merkten vervolgens op dat er tijdens die periode maar liefst 63 wijzigingen zijn doorgevoerd in de *news feed* van het socialemediaplatform. (Thorp & Vinson (2024) geven een samenvatting van de discussie.)

te bereiken. Gillespie geeft hier het voorbeeld van de website Flickr, waar gebruikers foto's kunnen delen met elkaar. Veel gebruikers kiezen ervoor om de stijl en inhoud van hun foto's zo aan te passen dat ze een zo groot mogelijke kans hebben om gezien te worden op het platform. Op YouTube zie je soms ook dat YouTubers expres een typefout maken in hun video's, zodat meer mensen commentaar plaatsen onder hun video's ("je hebt een woord verkeerd gespeld!!11!"). Het idee is dat het aantal comments signaleert aan het algoritme hoe interessant de video is, ongeacht de inhoud van die comments. Wat je ziet op sociale media is dus niet alleen het gevolg van het algoritme achter het platform, of van de gebruikers op het platform, maar van de interactie tussen die twee.

Kort gezegd zijn zoekresultaten op verschillende platforms dus niet onbevooroordeeld. Eerder heb je al gelezen dat die zoekresultaten gekleurd kunnen zijn, en nu hebben we een beter beeld van de redenen daarachter: de inhoud die je te zien krijgt is het gevolg van de ontwerpkeuzes die gemaakt zijn bij de ontwikkeling van die platforms, en die nog steeds gemaakt worden bij de doorontwikkeling ervan. Het is lastig om in deze context te praten over *representativiteit*, omdat je daarbij altijd een referentiekader nodig hebt. We kunnen in ieder geval zeggen dat foto's op Instagram en video's op YouTube niet representatief zijn voor foto's en video's *in het algemeen*. Het is ingewikkelder om iets te zeggen over de representativiteit van zoekresultaten voor de boodschappen die een bepaalde doelgroep ontvangt via sociale media.

3.6.3 TEST JEZELF

Kijk naar de video *Is Instagram changing art?*[13] op het YouTube-kanaal van *The Art Assignment*. Hierin bespreekt Sarah Urist Green de invloed van Instagram op de kunstwereld, en op onze beleving van kunst. In de video beschrijft ze meerdere ontwikkelingen die hebben te maken met sampling.

1) Welke ontwikkelingen in de video hebben te maken met sampling?
2) Wat betekenen deze ontwikkelingen voor de kunstwerken die je te zien krijgt
   a) ...als gebruiker van sociale media?
   b) ...als bezoeker van musea?
3) Welke van de door Gillespie geïdentificeerde dimensies zijn hier relevant?
4) Wat is het verband tussen deze dimensies en de ontwikkelingen die Urist Green bespreekt?

13 https://www.youtube.com/watch?v=eb0i070d3fA

3.7 HOE VERZAMEL JE DATA?

3.7.1 ZELF DATA VERZAMELEN

De nadruk in deze cursus ligt op het zelf verzamelen van onderzoeksdata. Er zijn twee manieren dat te doen: handmatig en automatisch.

Handmatige dataverzameling
betekent dat je zelf naar een website, app, archief, of gewoon naar buiten gaat om de data te verzamelen (bijvoorbeeld foto's maken in een winkelstraat). In eerste instantie kun je het beste handmatig beginnen: verzamel eerst een paar datapunten en kijk eens hoe de data eruit ziet. Als het goed is weet je dan ook hoeveel tijd het kost om de data te verzamelen, en of het de moeite waard is om het proces te automatiseren. Het is in ieder geval belangrijk om alle relevante informatie te noteren, zodat de dataverzameling herhaalbaar is. Met andere woorden:

- Iemand anders zou moeten kunnen reconstrueren hoe je de data verzameld hebt.
- Iemand anders zou zelf een vergelijkbare dataset moeten kunnen samenstellen op basis van de instructies en gegevens die je opgeschreven hebt.

Automatische dataverzameling
betekent dat je de computer het werk laat doen. Of het de moeite waard is om de dataverzameling te automatiseren is afhankelijk van de tijd die het kost om de dataverzameling voor te bereiden. Je hebt meestal de keuze uit:

- Bestaande hulpmiddelen: programma's die ontwikkeld zijn om data te verzamelen van een bepaalde website of een bepaald online platform. Een voorbeeld van zo'n programma is 4CAT (Peeters & Hagen, 2022), dat goed samenwerkt met Zeeschuimer, een *plugin* voor de Firefox-browser waarmee je posts van sociale media kunt verzamelen. Het online platform Communalytic kan ook gebruikt worden om data te verzamelen van sociale media (Gruzd et al., 2022). Er bestaan ook meer algemene browser-plugins zoals *Webscraper*, waarmee je systematisch data kunt verzamelen van verschillende websites.
- Een zelfgeschreven script: je kunt ook zelf een oplossing programmeren om automatisch data te verzamelen. Dit is nooit de eerste optie, omdat het meer tijd kost dan een bestaande oplossing. Maar uiteindelijk is het toch vaak nodig om een script te schrijven om ervoor te zorgen dat de data in het juiste format staat. Een uitgebreide handleiding vind je in het boek van Van Atteveldt et al. (2022).[14]

14 Het boek van Van Atteveldt et al. (2022) is gratis te lezen via: https://cssbook.net/

3.7.2 ANDERE MANIEREN VAN DATAVERZAMELING

Hieronder beschrijven we nog andere manieren van dataverzameling, die verschillende tekortkomingen van het zelf verzamelen kunnen compenseren.

Dataverzameling in een experiment.
Naast het verzamelen van bestaande data, kun je ook een inhoudsanalyse uitvoeren met data (tekst, video's, afbeeldingen) die je verzamelt via een experiment. Eerder werd het experiment van Braun et al. (2019) genoemd, waarin participanten zelf teksten moesten schrijven, die vervolgens bestudeerd werden via een inhoudsanalyse. Een ander voorbeeld is de studie van Antheunis et al. (2011), waarin participanten met elkaar moesten praten via de computer. Vervolgens werden de *chatlogs* onderworpen aan een inhoudsanalyse. Een laatste voorbeeld is de studie van Mui et al. (2017) waarin de auteurs video-opnames hebben gemaakt van hun participanten, en vervolgens de gezichtsuitdrukkingen van die participanten (specifiek: de aan- of afwezigheid van een glimlach) hebben gecodeerd.

Dagboekstudies
In plaats van een experimentele studie kun je er ook voor kiezen om participanten hun ervaringen bij te laten houden. Dit wordt ook wel een dagboekstudie (Engels: *diary study*) genoemd. Een mooi voorbeeld is de studie van Qutteina et al. (2019), die jonge adolescenten hebben gevraagd om berichten over voedsel op social media te verzamelen. Gedurende een periode van één week hebben de participanten tijdens hun normale gebruik van sociale media screenshots genomen van relevante berichten, en die opgestuurd naar de onderzoekers. Op deze manier verkregen de auteurs een representatieve sample van de verschillende marketingboodschappen die de doelgroep in het dagelijks leven ontvangt. Belangrijk hierbij is wel dat de participanten een training hebben ontvangen om relevante boodschappen te kunnen herkennen.

Deze voorbeelden laten zien dat het de moeite waard kan zijn om buiten de gebaande paden te treden, en er zo voor te zorgen dat je betekenisvolle data hebt om te analyseren.

3.8 WAT SLA JE ALLEMAAL OP?

3.8.1 ONLINE BRONNEN

Bij online bronnen is het belangrijk om, naast de data die je gaat coderen, ook alle relevante metadata op te slaan. (Metadata omvat alle gegevens over de data die je verzamelt.) Hieronder staat een overzicht van de gegevens die nuttig zijn om op te slaan:

- URL (zoek naar een permalink)
- Auteur (tenzij in strijd met AVG/GDPR)

- Datum
  - Van de zoekopdracht
  - Van de plaatsing
- Een identifier (unieke code die naar het bestand/bericht verwijst)
  - Van de oorspronkelijke bron
  - In jouw verzameling
- Andere relevante informatie

Het is belangrijk dat je niet zomaar alle beschikbare informatie opslaat. De huidige regelgeving staat toe dat je als onderzoeker informatie verzamelt die nodig is voor je onderzoek, maar we worden ook geacht om aan *data minimization* te doen. Met andere woorden: verzamel niet meer dan nodig. Dat betekent dat je vooraf goed moet nadenken over de manier waarop je de data later gaat gebruiken.

3.8.2 OFFLINE BRONNEN

Voor offline bronnen varieert de data die je moet opslaan. Denk vooraf in ieder geval aan:

- De vindbaarheid van de bron, in verband met de herhaalbaarheid van je studie.
- Relevante informatie voor de beschrijving van je sample.
- Relevante informatie voor je analyse.

3.8.3 BIJHOUDEN WAT JE NIET VERZAMELT

Indien mogelijk is het ook belangrijk om bij te houden wat je *niet* verzamelt. Vooral bij *purposive sampling* is het belangrijk om niet alleen bij te houden hoeveel je al gevonden hebt, maar ook hoe vaak een item niet voldeed aan je eisen, en eventueel: waarom die items niet voldeden aan je eisen. Met deze informatie ben je beter uitgerust om het belang en de generaliseerbaarheid van jouw studie te bespreken.

Voorbeeld

Stel je voor dat je alle berichten op sociale media wil verzamelen waar een bedrijf reageert met een doorverwijzing naar een online klachtenformulier, in plaats van een directe afhandeling van de klacht. Je bent benieuwd naar de manier waarop klagers reageren op zulke berichten: vinden ze het oké om zo doorgestuurd te worden, of worden ze boos? Om te bepalen of dit een veelvoorkomende of slechts een marginale situatie is, heb je informatie nodig over de verhouding tussen doorverwijzingen (waar je eigenlijk in geïnteresseerd bent) en andere reacties (waar je eigenlijk niet in geïnteresseerd bent).

3.9 HOE SLA JE DAT DAN OP?

Data kun je opslaan in een spreadsheet (te openen met bijvoorbeeld: Excel, Numbers, Google Sheets, LibreOffice Calc). Dat ziet er ongeveer zo uit als in Tabel 3.1:

Item	ID	Datum	Permalink	Bedrijf	Likes	Comments	Text
1	384859	2023-01-14	http://...	Pear	5483	434	We relea...
2	394868	2023-01-22	http://...	Pear	3593	327	One more...
3	209374	2023-01-28	http://...	Megasoft	23	82	Developer...

TABEL 3.1 *Voorbeeld van een spreadsheet met verzamelde data. Iedere rij staat voor een losse observatie, en de kolommen bevatten de eigenschappen van die observatie.*

3.9.1 STELREGELS VOOR HET OPSLAAN VAN DATA

Hieronder staan stelregels voor het opslaan van data.

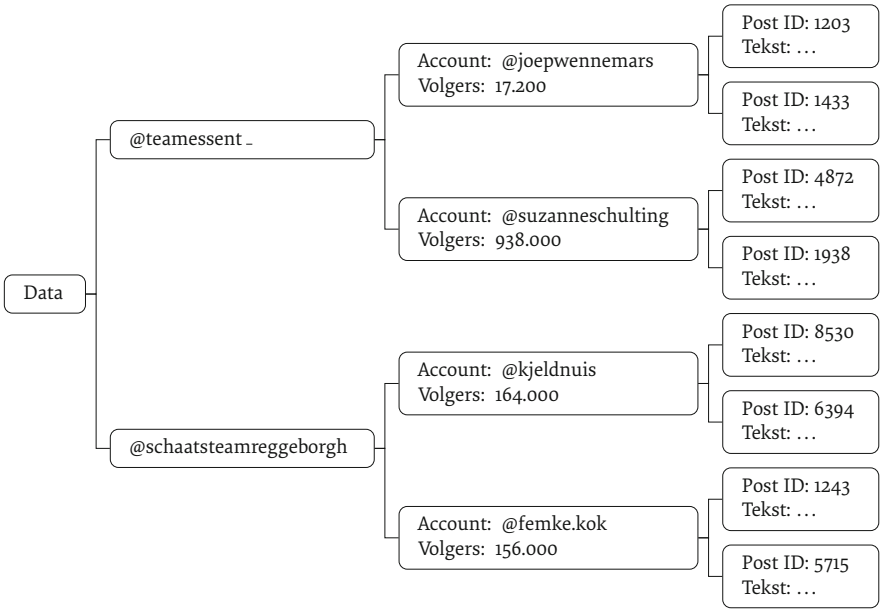
- Zorg voor logische eenheden:
  - Gebruik één regel per observatie, één kolom per eigenschap.[15]
  - Zet dus niet meerdere observaties op één regel.
  - Alle categorieën die je later in je inhoudsanalyse toevoegt zijn kolommen.
- Wees volledig in het opslaan van de metadata. Achteraf is het soms erg lastig om details terug te vinden.[16]
- Sla alle relevante data meteen op wanneer je deze verzamelt, zelfs wanneer je de data eenvoudig nogmaals kunt verzamelen (bijvoorbeeld wanneer je automatisch data verzamelt).
- Zorg er in ieder geval altijd voor dat duidelijk is waar de data vandaan komt.
- Sla de data twee keer op: één keer om verder te bestuderen, en één keer als backup.
- Zorg daarnaast nog voor een back-up van alle data, zodat er niets verloren kan gaan.

15 Als de volgorde van je items belangrijk is, voeg je meteen een kolom toe met oplopende cijfers. (Dit kan automatisch in Excel.) Zo kun je de data altijd weer sorteren in de juiste volgorde.

16 Daarnaast is het altijd mogelijk om details weg te laten in de analyse, maar het is niet mogelijk om diezelfde details achteraf nog toe te voegen.

3.9.2 HIERARCHISCHE (*MULTILEVEL*) DATA

Bij het analyseren van posts op sociale media wordt er vaak data verzameld op twee of zelfs meer niveaus. Naast de posts zelf wordt er dan bijvoorbeeld ook informatie verzameld over de accounts die de posts geplaatst hebben.[17] Dat zorgt voor een hierarchische structuur in de data, die geïllustreerd wordt in Figuur 3.1.



FIGUUR 3.1 *Fictieve dataset met posts op Instagram van schaatsers van verschillende teams. Voor ieder team zijn hier twee schaatsers uitgekozen, en van iedere schaatser zijn twee posts verzameld.*

Hierarchische data kun je opslaan als in Tabel 3.2. Het eerste wat opvalt aan deze tabel is dat er veel gegevens herhaald worden. Dat komt omdat de metadata over de accounts gekopieerd wordt voor ieder bericht in de dataset. Op deze manier kan het bestand eenvoudiger geanalyseerd worden door je favoriete statistiekprogramma.[18] Omdat het handmatig invoeren van de data vaak leidt tot (type)fouten, is het verstandig om het proces te automatiseren via een programmeertaal zoals Python of R. Als dat geen optie is, kun je lezen in Bijlage G hoe je dit in Excel zou kunnen doen.

17 In sommige gevallen worden er zowel eigenschappen van de accounts als eigenschappen van de posts gecodeerd. Voor beide codeertaken dien je dan ook te kijken naar de betrouwbaarheid van het codeerproces.

18 Om hierarchische data te analyseren wordt er vaak gebruik gemaakt van *mixed effects* modellen. Winter (2013) legt op een toegankelijke manier uit hoe deze modellen werken. (Een uitgebreidere uitleg is beschikbaar via zijn boek (Winter, 2019).)

Team	Account	Volgers	Post ID	Tekst	...
@teamessent_	@joepwennemars	17200	1203	...	...
@teamessent_	@joepwennemars	17200	1433	...	...
@teamessent_	@suzanneschulting	938000	4872	...	...
@teamessent_	@suzanneschulting	938000	1938	...	...
@schaatsteamreggeborgh	@kjeldnuis	164000	8530	...	...
@schaatsteamreggeborgh	@kjeldnuis	164000	6394	...	...
@schaatsteamreggeborgh	@femke.kok	156000	1243	...	...
@schaatsteamreggeborgh	@femke.kok	156000	5715	...	...

TABEL 3.2 *De data uit Figuur 3.1 als spreadsheet. Na de getoonde kolommen kunnen nog extra kolommen toegevoegd worden voor de verschillende categorieën die je gaat coderen (en natuurlijk voor andere relevante metadata).*

3.9.3 AFWIJKENDE SOORTEN DATA

Niet alle soorten data zijn eenvoudig in een spreadsheet te stoppen. Afbeeldingen, geluidsopnamen of video's kun je bijvoorbeeld niet zomaar invoegen. Het beste is dan om in je spreadsheet alle metadata op te slaan met een *verwijzing* naar je data, bijvoorbeeld de bestandsnaam, of een link naar het bestand. Het coderen kun je vervolgens nog steeds doen in Excel, of in een programma dat is uitgerust om andere mediabestanden te coderen. Of je het bestand zelf moet downloaden is afhankelijk van de situatie. Bijvoorbeeld: video's van YouTube hoef je niet te downloaden als je alles in Excel codeert, maar wel als je een speciaal programma hebt (zoals ELAN) om video's te annoteren.

3.10 DATAVERZAMELING BIJHOUDEN

Bij het verzamelen van data is het belangrijk om transparant te zijn over het proces. Dat betekent dat je precies moet bijhouden op welke datum je de data verzameld hebt, en hoe je dat precies gedaan hebt.

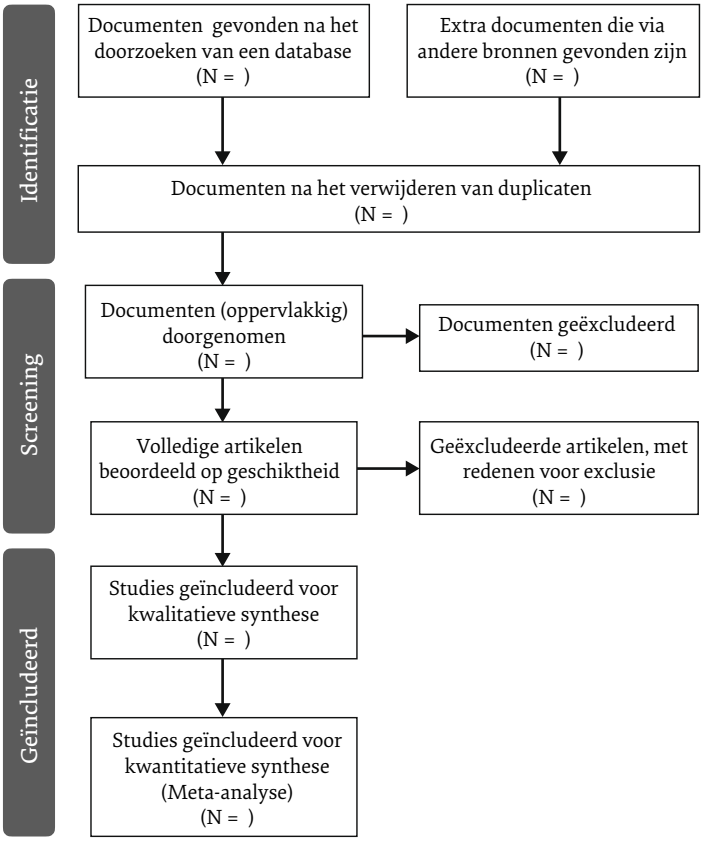
Als je de ***social media accounts* van meerdere groepen** vergelijkt, dien je ook aan te geven *hoe* je aan die groepering bent gekomen.[19] Bijvoorbeeld: Van Hooijdonk & Liebrecht (2021) hebben de webcarestrategieën van *budget* versus *full-service* vliegtuigmaatschappijen vergeleken op sociale media, waarbij ze gebruik maakten van een bestaande lijst van bekende maatschappijen. Er zit altijd een bepaalde mate van willekeur in de selectie van groepen of individuen die je gaat vergelijken, maar een referentielijst helpt om te verantwoorden waarom je precies die selectie hebt gemaakt.

19 *Waarom* je die groepen vergelijkt heb je al eerder uitgelegd in de inleiding, maar nu gaat het om het selectieproces.

Als je een **zoekmachine** hebt gebruikt, dien je aan te geven:

- Welke zoektermen je hebt gebruikt.
- Hoeveel zoekresultaten je gevonden hebt.
- Hoeveel items je daarvan gebruikt hebt.
- Hoe je die items geselecteerd hebt.

De PRISMA-benadering (Page et al., 2021) kan je helpen om dit proces beter te begrijpen, en om jouw dataverzameling te rapporteren. Hieronder zie je het PRISMA-stroomdiagram (Figuur 3.2), dat gebruikt wordt voor het beschrijven van systematische literatuurstudies.



FIGUUR 3.2 *PRISMA-stroomdiagram, gebaseerd op het diagram van de officiële website: http://prisma-statement.org/*

Een literatuurstudie klinkt misschien heel anders dan een inhoudsanalyse, maar eigenlijk zijn ze heel vergelijkbaar: in beide gevallen verzamel je gegevens, die je filtert op basis van verschillende karakteristieken, waarna je die gegevens op systematische wijze analyseert. In

een literatuurstudie zijn die gegevens onderzoeksartikelen, en een inhoudsanalyse kun je in principe toepassen op allerlei soorten data. Een voorbeeld van een inhoudsanalyse die de PRISMA-benadering hanteert is de studie van Vromans et al. (2019), waarin de auteurs medische beslishulpen (waarmee patiënten geïnformeerd worden over de voor- en nadelen van verschillende behandelingen) hebben geanalyseerd om te kijken of die beslishulpen verbeterd zouden kunnen worden. Daarbij zijn onderzoeksartikelen drie keer gecodeerd: eerst vond er een screening plaats op basis van het abstract en de titel (één categorie, twee niveaus: includeren of excluderen), daarna werden de papers volledig gelezen om te kijken of ze inderdaad relevant waren (één categorie, twee niveaus: relevant of niet relevant (met reden)), en tenslotte werden de beslishulpen gecodeerd aan de hand van bestaande richtlijnen (meerdere categorieën, meerdere niveaus).

PRISMA is dus nuttig als gids voor systematische reviews, maar helpt ook om na te denken over dataverzameling in het algemeen. De PRISMA-methode biedt geen richtlijnen over het daadwerkelijk coderen van de data, maar daarover lees je later meer.

3.11 ETHIEK EN DATAVERZAMELING

Dataverzameling roept ook veel ethische vragen op, van fundamentele vragen (mag het überhaupt?) naar meer praktische vragen (hoe kun je dit op een verantwoordelijke manier doen?). Als vakgebied staat de ethiek van online dataverzameling nog steeds in de kinderschoenen, wat ook betekent dat de antwoorden op veel vragen verschillen tussen verschillende onderzoekers, en dat de huidige *best practices* (bij wijze van spreken) morgen alweer anders kunnen zijn. Dat gezegd hebbende, is het wel belangrijk om de belangrijkste thema's te bespreken. Dat doen we in de volgende hoofdstukken:

Hoofdstuk 8: Ethiek
- Mag je zomaar data verzamelen van sociale media? (Niet alleen wettelijk, maar ook moreel.)
- Moet je eigenlijk ethische goedkeuring aanvragen voor het analyseren van publieke data?
- Wat betekent dit voor de privacy van de auteurs?
- Als je data verkrijgt via een samenwerking met een bedrijf, kun je dan nog wel onafhankelijk zijn als onderzoeker?

Hoofdstuk 9: Datamanagement
- Hoe kun je de data op een verantwoordelijke manier opslaan en archiveren?
- Mag je de data die je verzameld hebt ook weer delen met anderen?

Hoofdstuk 10: Rapportage

- Hoe verwijs je naar de data die je hebt verzameld? Moet je de auteur juist wel of niet citeren?

Meer weten?

Als je meer wil weten over ethiek en online dataverzameling, dan kun je bijvoorbeeld kijken op de website van de Association of Internet Research. Zij hebben een aparte pagina die gewijd is aan ethische vragen en richtlijnen: https://aoir.org/ethics/

HOOFDSTUK 4

Taal en inhoudsanalyse

4.1 INLEIDING

In dit hoofdstuk gaan we kijken naar de verschillende manieren waarop eerdere onderzoekers hebben nagedacht over taal, en manieren waarop we talige boodschappen kunnen analyseren. Dat doen we in vogelvlucht; taalwetenschap is een ontzettend breed en veelzijdig vakgebied, en wij hebben maar beperkt de tijd om ons hier verder in te verdiepen. Het doel van dit hoofdstuk is om je te leren welke verschillende analyseniveaus er zijn binnen de taalwetenschap, wat die niveaus inhouden, en wat die niveaus betekenen voor ons als we een inhoudsanalyse willen uitvoeren.

Geen enkel niveau is objectief beter of slechter dan een ander niveau. Het is onze taak als onderzoekers om te begrijpen welk niveau het meest geschikt is om de data te onderzoeken. Uiteindelijk is de onderzoeksvraag leidend om te bepalen vanuit welk perspectief we naar de data zouden moeten kijken.

4.2 VERSCHILLENDE ANALYSENIVEAUS, VERSCHILLENDE VAKGEBIEDEN

Talige uitingen kunnen we analyseren op verschillende niveaus, die allemaal weer geassocieerd worden met verschillende vakgebieden binnen de taalwetenschap. Voor sommige analyseniveaus wordt hieronder direct een verband gelegd tussen het niveau en het uitvoeren van een inhoudsanalyse, maar algemeen kunnen we ook stellen dat een beter begrip van taal de mogelijkheid vergroot om taaldata op verschillende manieren te analyseren.

4.2.1 GELUID

Voor gesproken taal beginnen we met geluid. Daarbij kunnen we binnen de taalwetenschap twee verschillende vakgebieden onderscheiden:

Fonetiek

Binnen de fonetiek bestuderen wetenschappers waarneembare spraak. Sommige mensen richten zich meer op spraakproductie waar anderen zich richten op de perceptie van verschillende klanken. Misschien heb je in een woordenboek ook weleens gezien dat er in een ander schrift wordt omschreven hoe je de verschillende woorden moet uitspreken. Dat gebeurt aan de hand van het International Phonetic Alphabet (IPA), dat door fonetici is ontwikkeld. Omdat spelling geen goede indicator is van de uitspraak van verschillende woorden, en omdat niet iedereen alle woorden hetzelfde uitspreekt, wordt IPA vaak gebruikt om geluidsopnames te annoteren en precies te omschrijven hoe mensen alles uitspreken. Fonetiek is relevant omdat de manier waarop we bepaalde woorden uitspreken vaak samenhangt met sociale variatie. Die variatie wordt bestudeerd binnen een ander vakgebied: **sociolinguïstiek**. Daarover lees je later meer.

Fonologie

Naast de fonetiek is er ook nog de fonologie. In dit vakgebied onderzoekt men hoe het klanksysteem werkt 'in ons hoofd' en hoe verschillende soorten klanken met elkaar interacteren. Net zoals we in het Nederlands een grammatica hebben voor de manier waarop je woorden kunt combineren tot zinnen, bestuderen fonologen hoe klanken wel of niet gecombineerd kunnen worden met elkaar, en welke regelmatigheden hierin bestaan tussen verschillende talen.

4.2.2 WOORDEN EN WOORDVORMEN

Naast geluiden kunnen we ook woorden en woordvormen bestuderen. De **morfologie** analyseert hoe woorden zijn opgebouwd uit kleinere betekenisvolle elementen, die 'morfemen' worden genoemd. Het woord 'huisje' is bijvoorbeeld een combinatie van het zelfstandige morfeem 'huis' en het verkleinende morfeem 'je.' Het meervoud 'huizen' is een samenstelling van 'huis' en het afhankelijke meervoudsmorfeem 'en.' Daarbij zien we ook een klankverandering optreden: van /s/ naar /z/, die gereflecteerd is in de spelling. (Deze klankverandering heeft te maken met de fonetiek/fonologie: de stemloze medeklinker /s/ assimileert met de (per definitie stemhebbende) klinkers naar de stemhebbende medeklinker /z/.)

Kennis van morfologie helpt ons om beter te begrijpen hoe dezelfde betekenisvolle elementen (morfemen zoals 'huis') verschillende verschijningsvormen kunnen hebben, en wat dit betekent voor het (automatisch) herkennen van deze elementen in een zin. Daarnaast kan de morfologie ons ook helpen wanneer je talen vergelijkt die een rijkere of armere morfologie hebben dan het Nederlands.

4.2.3 ZINNEN EN ZINSDELEN

Het volgende niveau is de zinsbouw (**syntaxis**) waarbij we kijken hoe zinnen zijn gestructureerd. Syntactici onderzoeken hoe je in natuurlijke talen zoals het Nederlands woorden wel of niet kunt combineren tot grammaticale zinnen. Kennis van syntax is bijvoorbeeld nuttig wanneer er sprake is van syntactische variatie, waarbij verschillende sprekers van het Nederlands er (meestal onbewust) voor kiezen om dezelfde woorden met dezelfde boodschap in een andere volgorde te combineren. Bijvoorbeeld:

- ...dat ik hem heb gezien.
- ...dat ik hem gezien heb.

Net als bij de verschillende manieren om bepaalde woorden uit te spreken, weerspiegelt syntactische variatie ook vaak sociale variabelen. Daarover lees je later meer in de paragraaf over sociolinguïstiek.

Een ander onderdeel van syntax is het idee dat verschillende woorden in een zin ook verschillende rollen kunnen hebben. Op de basisschool of de middelbare school heb je al geleerd over de verschillende *woordsoorten*, zoals werkwoorden, zelfstandig naamwoorden, bijvoeglijk naamwoorden, enzovoorts. Afhankelijk van je onderzoeksvraag, zou je deze rollen ook kunnen gebruiken door met een automatische *part-of-speech tagger* specifieke woordsoorten te herkennen in een tekst. Meer hierover kun je lezen in het hoofdstuk over mensen en computers.

4.2.4 LETTERLIJKE BETEKENIS

Tot nu toe hebben we alleen maar gekeken naar de *vorm* van taaluitingen, maar iedere uiting heeft natuurlijk ook een *betekenis.* Het vakgebied dat kijkt naar de letterlijke betekenis van uitdrukkingen in natuurlijke taal heet **semantiek**. Dit vakgebied wordt vaak opgesplitst in twee losse vakgebieden:

1) **Lexicale semantiek**: een deelgebied van de semantiek waarin gekeken wordt naar de betekenis van individuele woorden. Sommige onderzoekers kijken naar veranderingen in de betekenis van woorden over tijd, terwijl anderen onderzoek doen naar de mentale representatie van verschillende betekenissen.
2) **'Algemene' semantiek**: een deelgebied van de semantiek waarin gekeken wordt naar de betekenis van combinaties van woorden. In dit vakgebied worden onderwerpen besproken als *compositionaliteit*: is de betekenis van een zin het resultaat van de betekenis van de individuele woorden en de manier waarop ze gecombineerd worden? Of moeten we de betekenis van een zin op een andere manier afleiden?

Kennis van de semantiek is relevant om beter te begrijpen hoe vorm en betekenis zich tot elkaar verhouden, en om te zien hoe ambiguïteit kan optreden.

Vorm en betekenis

Taal heeft veel verschillende verschijningsvormen, waarvan de bekendste zijn: gesproken taal, geschreven taal, en gebarentaal. In Tilburg bestudeert Neil Cohn ook hoe stripverhalen en andere beelden gezien kunnen worden als instanties van een visuele taal, met haar eigen grammatica. In alle gevallen kun je zeggen dat mensen met een bepaalde taaluiting een bepaalde boodschap over willen brengen. Die boodschap kunnen we vaak nog uitsplitsen in een letterlijke boodschap (bestudeerd binnen de semantiek) en een verdere interpretatie van die boodschap als daadwerkelijke *bedoeling* van de zender (bestudeerd binnen de pragmatiek; waarover later meer).[1]

Als we de letterlijke boodschap van een bericht zien als de betekenis van dat bericht, dan kun je je afvragen: wat is de verhouding tussen de vorm van het bericht en haar betekenis? Daarover hebben semantici vele observaties gedaan die ook relevant zijn voor ons. Misschien de belangrijkste observatie:

De relatie tussen vorm en betekenis is niet één-op-één maar veel-op-veel.

Met andere woorden: er zijn verschillende (combinaties van) woorden met dezelfde betekenis, en er zijn verschillende betekenissen die uitgedrukt kunnen worden door dezelfde (combinaties van) woorden. Het bekendste probleem dat we hierdoor krijgen is ambiguïteit.

Ambiguïteit

Er zijn grofweg twee soorten ambiguïteit (die ook te koppelen zijn aan de twee deelgebieden van de semantiek):

1) **Lexicale ambiguiteit**, waarbij één woord verschillende betekenissen kan hebben (bijvoorbeeld: 'bank' kan verwijzen naar een meubel of een financiële instelling, en 'toren' kan verwijzen naar een gebouw of een schaakstuk).
2) **Structurele ambiguïteit**, waarbij een zin meerdere betekenissen kan hebben, afhankelijk van hoe je hem interpreteert. Het bekendste voorbeeld onder taalwetenschappers is: "ik zie de man met de telescoop", waarbij de ik-persoon in de ene betekenis een telescoop gebruikt om de man te zien, en het in de andere betekenis juist de man is die een telescoop heeft. De reden dat deze vorm van ambiguïteit "structureel" wordt genoemd, is omdat de woorden voor de verschillende betekenissen op een verschillende manier met elkaar gecombineerd worden; de onderliggende syntactische structuur verschilt tussen de twee betekenissen. (Alleen kun je dat niet zien wanneer de zin zo uitgeschreven op papier staat.)

Kennis over ambiguïteit is relevant voor inhoudsanalyse, omdat dubbelzinnigheden kunnen leiden tot verschillen in de manier waarop verschillende codeurs een tekst interpreteren. Bij

1 Dit is eigenlijk het klassieke communicatiemodel, waarover we nog kunnen opmerken dat de boodschap ook nog verstoord kan worden, wat extra uitdagingen oplevert voor de ontvanger (het *noisy channel*-model van Claude Shannon). Dit gaat echter te ver voor dit werkboek over inhoudsanalyse.

het opstellen van je codeboek dien je dan ook rekening te houden met ambiguïteit: kun je redelijkerwijs verwachten dat de codeurs een tekst eenduidig zullen coderen? En hoe ga je om met een codeertaak die inherent dubbelzinnig is?

4.2.5 TAALGEBRUIK

Naast de letterlijke betekenis kan een boodschap ook nog een extra lading krijgen door de context waarin een uitspraak gedaan is. Het vakgebied dat taalgebruik in context bestudeert heet **pragmatiek** (hoewel er ook wat overlap is met de sociolinguïstiek, waarover later meer). Dit vakgebied heeft (net als de semantiek) veel te danken aan de taalfilosofie, waarin filosofen zoals Austin, Grice, en Searle hebben geschreven over de manier waarop wij in een gesprek betekenis geven aan de uitspraken van onze gesprekspartners.

De maximen van Grice

Herbert Paul Grice is vooral bekend om zijn coöperatieve principe, het uitgangspunt dat mensen met hun uitspraken willen voortbouwen op wat eerder gezegd is, en toewerken naar het doel van het gesprek. Daarvoor heeft hij ook een aantal gedragsregels (of *maximen*) opgesteld:

- Wees eerlijk (kwaliteit)
- Wees informatief (kwantiteit)
- Wees relevant (relevantie)
- Wees duidelijk (helderheid)

Deze regels zijn geen harde eisen, maar eerder richtlijnen die we allemaal proberen te volgen wanneer we een gesprek hebben. Je kunt de maximen van Grice namelijk ook overtreden, en daarmee signaleren dat je meer of iets anders bedoelt dan wat je letterlijk gezegd hebt.

Een bekend voorbeeld van het overtreden van de maximen van Grice is de uitspraak van de Nederlandse presentatrice Chantal Janzen tijdens het Televizier Ring gala over een optreden van hockeyster Fatima Moreira de Melo die net een nummer had gezongen. Nadat het publiek had geapplaudisseerd zei Janzen: “Wat kan ze hockeyen, hè?” Daarmee overtrad Janzen het *maxime van relevantie*: hockeyprestaties waren op dat moment helemaal niet relevant. Met haar overtreding signaleerde Janzen haar eigenlijke bedoeling: het was een verschrikkelijk vals optreden geweest.

Met dit voorbeeld zien we ook een uitdaging voor inhoudsanalyse: om te begrijpen wat er speelt, moeten we onze conclusies soms baseren op niet-letterlijke informatie, die soms ook tegengesteld kan zijn aan de letterlijke boodschap. Janzen deed een *positieve* uitspraak over Moreira de Melo, maar haar boodschap was *negatief*. Dit soort voorbeelden is ook erg lastig voor computers om te begrijpen, omdat automatische taalverwerking nog niet zo goed is in het meenemen van context. (Meer hierover lees je in Hoofdstuk 7 over mensen en computers.)

Taalhandelingstheorie

Naast de cooperatieve principes van Grice is er nog een bekende pragmatische theorie: de Taalhandelingstheorie van Austin (1962) en Searle (1969). Zij beargumenteren op overtuigende wijze dat je met een uitspraak niet alleen iets *zegt*, maar dat je er ook iets mee *doet*. In de taalhandelingstheorie kunnen we verschillende soorten taalhandelingen onderscheiden:

- **Assertief**: proberen de ander zich erin te laten vinden.
  - Bijvoorbeeld: Taalhandelingen zijn erg nuttig!
- **Directief**: proberen de ander gedrag te laten vertonen.
  - Bijvoorbeeld: Kun je me de suiker aangeven?
- **Commissief**: proberen overeen te komen met propositionele inhoud.
  - Bijvoorbeeld: Ik kom vrijdag naar je verjaardag.
- **Expressief**: proberen oprechtheid uit te drukken
  - Bijvoorbeeld: Sorry voor mijn geschreeuw
- **Declaratief**: proberen een verandering teweeg te brengen in de wereld.
  - Je bent ontslagen!

Binnen de taalhandelingstheorie kunnen we iedere uitspraak analyseren in termen van drie niveaus:

1) Locutie: wat je zegt.
2) Perlocutie: wat je bedoelt te communiceren.
3) Illocutie: wat je daarmee hoopt te bereiken.

De taalhandelingstheorie geeft ons een conceptueel raamwerk en de terminologie om beter te begrijpen wat mensen eigenlijk doen wanneer ze een boodschap de wereld insturen. In de volgende paragraaf bespreken we ook een inhoudsanalyse van Van Hooijdonk & Liebrecht (2021) die impliciet gebruik maakt van de taalhandelingstheorie.

4.3 CODEREN EN ANALYSENIVEAUS

Nu we een overzicht hebben van de verschillende analyseniveaus binnen de taalwetenschap kunnen we kijken hoe die analyseniveaus terugkomen binnen inhoudsanalyses. Daarvoor gebruiken we twee voorbeelden: de studie van Van Hooijdonk & Liebrecht (2021) over webcare-strategieën, en de studie van Antheunis et al. (2011) over computer-gemedieerde communicatie.

4.3.1 INHOUDSANALYSE EN PRAGMATIEK: VAN HOOIJDONK & LIEBRECHT (2021)

Van Hooijdonk & Liebrecht (2021) bestuderen webcarestrategieën die gebruikt worden door vliegtuigmaatschappijen wanneer mensen klachten plaatsen op sociale media. Webcarestra-

tegieën corresponderen met verschillende soorten reacties die je als organisatie kunt geven, zoals *doorverwijzen, excuses aanbieden, sympathie tonen, informatie geven*, enzovoorts. Net als bij de wat meer abstracte taalhandelingen in de vorige paragraaf proberen luchtvaartmaatschappijen met hun boodschap dus ook iets te *doen*.

Voor Van Hooijdonk & Liebrecht (2021) is het belangrijk om op een betrouwbare manier de verschillende webcarestrategieën te kunnen herkennen. Natuurlijk hebben ze een codeboek opgesteld, waarin ze iedere categorie definiëren en voorbeelden geven die ze eerder hebben gevonden, maar ze geven ook een lijst van signalen die je kunt gebruiken om de verschillende strategieën te kunnen herkennen. Die signalen noemen ze IFIDs (*Illocutionary Force Indicating Devices*): "Linguistic elements that indicate the potential presence of a certain strategy." Deze afkorting is ook weer een verwijzing naar de theorie van Austin en Searle.

Als voorbeeld van IFIDs noemen Van Hooijdonk & Liebrecht (2021) uitdrukkingen als *DM, direct message, online form, helpdesk at the airport*. Deze uitdrukkingen worden vaak samen gebruikt met de conversationele strategie die 'doorverwijzen' heet. Natuurlijk is er geen garantie dat het voorkomen van één van bovenstaande uitdrukkingen *altijd* signaleert dat er sprake is van een doorverwijzing, maar het geeft de codeurs wel houvast om makkelijker te kunnen bepalen of de organisatie de klager doorverwijst. Wanneer je verschillende lijsten met IFIDs hebt opgesteld, zou je ook kunnen proberen om berichten automatisch te coderen. Daarover lees je later meer in Hoofdstuk 7 over Mensen en Computers.

4.3.2 INHOUDSANALYSE, EENHEDEN, EN SEMANTIEK/PRAGMATIEK: ANTHEUNIS ET AL. (2012)

Het is niet alleen belangrijk om het juiste analyseniveau te kiezen voor de verwerking van talige data; je moet ook nog eens nadenken over de juiste codeer-eenheden, bijvoorbeeld het hele bericht, een alinea, een zin, of een zinsdeel. Om dat te illustreren kijken we naar de studie van Antheunis et al. (2011) die hebben onderzocht in hoeverre de setting van een gesprek invloed hebben op de *self-disclosure* (met andere woorden: hoeveel iemand over zichzelf vertelt) die getoond wordt door de participanten die met elkaar communiceren. Antheunis et al. (2011) lieten participanten op drie verschillende manieren met elkaar communiceren: *face-to-face*, via de chat zonder webcam, en via de chat met een webcam. De data voor hun onderzoek bestond dus uit transcripten van de face-to-face gesprekken, en logs van de gesprekken via de computer. De uitingen van de participanten werden gecodeerd als (1) self-disclosure, (2) vraag, of (3) anders, maar dat ging niet zomaar in één keer.

Het grote **probleem** voor de auteurs was dat een uitspraak kan vallen onder meerdere categorieën. Ze konden dus niet de tekst opsplitsen in zinnen en die vervolgens coderen. De **oplossing** Antheunis et al. (2011) was om de uitspraken op te splitsen in losse semantische eenheden, die zij "utterances" noemen. (In de semantiek zouden dit misschien eerder *proposities* genoemd worden.) Een zin als "Mijn naam is Hugo en ik hou van ballet" werd bijvoorbeeld opgesplitst in twee losse proposities:

1) "Mijn naam is Hugo" en
2) "Ik hou van ballet."

Deze proposities werden vervolgens gecodeerd. Eigenlijk voeren Antheunis et al. (2011) hier dus *twee* codeertaken achter elkaar uit. In termen van de eerder besproken niveaus zou je kunnen zeggen dat het opdelen van zinnen in propositionele eenheden enige semantische kennis vereist, terwijl vragen en het geven van informatie over jezelf typische taalhandelingen zijn.

4.4 ABSTRACTIE EN CODEREN

Kijkend naar de verschillende analyseniveaus kunnen we stellen dat er een relatie lijkt te zijn tussen het abstractieniveau en de moeilijkheidsgraad van het coderen: hoe abstracter de analyse wordt, des te moeilijker is de annotatie ervan. Binnen de inhoudsanalyse noemen we concrete, aanwijsbare eigenschappen **manifest**, en meer abstracte, verborgen eigenschappen worden **latent** genoemd. Hieronder staan voorbeelden van vragen die gegrond zijn in de Morfologie (concreet, manifest):

- Wat zijn de losse woorden in deze zin?[2]
- Wat is de stam van ieder woord? (fiets+en)

Vragen die gerelateerd zijn aan de Pragmatiek (meest abstract) zijn bijvoorbeeld:

- Is deze uitspraak sarcastisch/ironisch bedoeld?
- Wat wil de spreker hiermee doen of zeggen? (Informatie geven, doorverwijzen, ...)

De moeilijkheidsgraad van het coderen hangt ook samen met de betrouwbaarheid van het codeerproces. Codeurs zullen eerder fouten maken bij taken die lastiger of subjectiever zijn. Dat wil niet zeggen dat we de pragmatiek maar moeten opgeven; vaak is dit juist de informatie waar we naar op zoek zijn! Maar het betekent wel dat we zorgvuldig te werk moeten gaan om overeenstemming te bereiken tussen de verschillende codeurs. Hierover lees je meer in het volgende hoofdstuk (3): een codeboek ontwikkelen.

Misschien heb je al gemerkt dat we de fonetiek voor het gemak even hebben overgeslagen. Hoewel geluid erg concreet is, kost het behoorlijk wat oefening om losse klanken en andere fonetische variabelen te herkennen. De oorzaak hiervan ligt in de aard van het signaal: waar geschreven taal bestaat uit discrete eenheden (bijvoorbeeld: letters), hebben we bij geluid te

2 Hier is wel enige nuancering op zijn plaats. Hoewel het voor het Nederlands relatief eenvoudig is om een zin op te splitsen in woorden, geldt dit niet voor alle talen. Bovendien verschillen taalkundigen ook van mening over de precieze definitie van 'woorden' (Haspelmath, 2023).

maken met een continuüm. Er is niets in het signaal dat aangeeft waar een bepaalde klank begint of eindigt; dat moeten we allemaal zelf doen.

4.5 SOCIOLINGUÏSTIEK

Sociolinguïstiek is een vakgebied waarbinnen onderzoekers bestuderen hoe taal en sociale groepen of situaties zich tot elkaar verhouden. Het is algemeen bekend dat hoe we praten laat zien wie we zijn of hoe je je voelt over je gesprekspartners. Daarvoor kijken sociolinguisten bijvoorbeeld naar de woordkeuze, uitspraak, of zinsvolgorde die we hanteren.

4.5.1 UITSPRAAK

Aan Brabanders hoef je niet uit te leggen dat je uitspraak kan laten zien waar je vandaan komt. Iedereen weet dat mensen in het zuiden des lands de <g> zachter uitspreken dan boven de rivieren, waar men een 'harde g' gebruikt. Dat bewustzijn zorgt er ook voor dat sommige Brabanders hun accent proberen te verbergen (om zich aan te passen) of juist extra benadrukken (om hun eigen identiteit te tonen) wanneer ze in de Randstad zijn. Daarmee is een fonetische variabele ook een indicator van sociale verhoudingen.

4.5.2 WOORDKEUZE EN MORFOLOGISCHE VERSCHILLEN

Misschien wel een van de bekendste observaties met betrekking tot woordkeuze is de patat-frietkaart, die hieronder weergegeven is in Figuur 4.1. Deze kaart is gebaseerd op een landelijke vragenlijst waarin mensen gevraagd werd hoe ze in het vet gebakken aardappelstaafjes noemen. Net als bij de 'zachte g' is woorkeuze een markeerder van sociale identiteit: mensen in het noordelijke deel van het Nederlandse taalgebied gebruiken vaker "patat" terwijl mensen onder de rivieren vaker "friet" zeggen.

Hetzelfde fenomeen zien we ook bij morfemen: in Brabantse dialecten krijgen lidwoorden (zoals *de*) en bijvoeglijk naamwoorden (zoals *blauw*) een extra achtervoegsel *-e(n)* voor mannelijke bijvoeglijk naamwoorden (zoals *auto*). Dus in het Standaardnederlands zegt men *de blauwe auto* en in het Brabants zegt men *de**n** blauwe**n** auto* (Doreleijers, 2024). Omdat deze regel typisch is voor het Brabants, wordt het achtervoegsel ook wel een *dialectmarkeerder* genoemd; hieraan kunnen we herkennen dat iemand Brabants praat. Dat geeft het achtervoegsel ook een tweede functie: hiermee kun je signaleren dat je Brabants *bent*.[3]

3 Doreleijers (2024) observeert dat verschillende sprekers van het Brabants ook overgeneraliseren door hetzelfde achtervoegsel ook te gebruiken voor onzijdige woorden: *en**e** koekske* in plaats van *een koekske* (Standaardnederlands: *een koekje*). Wanneer sprekers bepaalde dialectmarkeerders in de 'verkeerde' context gebruiken wordt dat *hyperdialectisme* genoemd. Er zijn verschillende theorieën over de oorzaken van hyperdialectismen, en wat dat zegt over zowel de sprekers van dialecten als de dialecten zelf, maar een bespreking hiervan gaat te ver voor dit boek. Je kunt er meer over lezen in het proefschrift van Doreleijers (2024).

FIGUUR 4.1 *De patat-frietkaart. Bron: Wikimedia, kaart gemaakt door gebruiker Cavit op basis van een eerdere kaart gemaakt door Jan Stroop bij het Meertens Instituut.*

4.5.3 WOORDVOLGORDE

Als laatste voorbeeld van een variabele die bestudeerd wordt door sociolinguïsten kunnen we kijken naar woordvolgorde. In Nederland zijn er twee volgordes waarin (hulp)werkwoorden geplaatst worden: de zogenaamde *groene* en de *rode* volgorde. Tabel 4.1 geeft voorbeelden van beide werkwoordsvolgordes.

Groene volgorde	Rode volgorde
Dat is de afstand, die door mij *gelopen is.*	Dat is de afstand, die door mij *is gelopen.*
Ik vind dat dit *uitgelegd moet worden.*	Ik vind dat dit *moet worden uitgelegd.*
Ik denk dat hij *drinken wil.*	Ik denk dat hij *wil drinken.*

TABEL 4.1 *Voorbeelden van zinnen met de 'groene' en 'rode' werkwoordsvolgorde.*

Voor deze variabele vond Pauwels (1953) een duidelijke geografische spreiding: de groene volgorde was dominant in Noord-Nederland (Groningen en Drenthe), en de rode volgorde

kwam het vaakst voor in Noord- en Zuid-Holland, Utrecht en Noord-Brabant.[4] Het moge duidelijk zijn na bovenstaande voorbeelden: ons taalgebruik hangt samen met onze afkomst en identiteit. Tegelijkertijd is ons taalgebruik niet statisch en onveranderlijk. Hieronder lees je meer over de manieren waarop we ons taalgebruik aanpassen aan de omgeving.

4.6 ACCOMMODATIE, CONVERGENTIE, EN DIVERGENTIE

Mensen praten niet altijd op dezelfde manier. Hoe ze praten hangt af van de situatie. Denk er maar eens over na: hoe praat je...

- ...als je praat met vrienden
- ...als je praat met iemand die je niet zo leuk vindt
- ...als je op je werk bent
- ...als je een sollicitatiegesprek hebt
- ...als je een presentatie moet geven
- ...als je op TV of op de radio komt
- ...als je een vlog of TikTok-video opneemt
- ...als je met je (groot)ouders praat

Het ligt in onze natuur om ons taalgebruik af te stellen op onze omgeving. Daarover zijn meerdere theorieën opgesteld, zoals *accommodation theory* (Giles et al., 1973) en *audience design* (Bell, 1984). Binnen accommodation theory kun je je op twee manieren aanpassen:

1) **Convergentie**: Meer praten zoals je gesprekspartners
2) **Divergentie**: Behouden of versterken van je eigen taalgebruik

4.6.1 ACCOMMODATIE EN INHOUDSANALYSE: EEN VOORBEELD

Van der Zanden et al. (2018) hebben onderzoek gedaan naar taalaccommodatie in online datingprofielen. Zo keken ze naar de mate waarin hoger opgeleide vrijgezellen hun teksten aanpasten aan het publiek op twee verschillende datingsites: een website voor een algemeen publiek, en een website voor hoger opgeleiden. De profielteksten op beide websites zijn automatisch verzameld en deels handmatig (hoeveel taalfouten worden er gemaakt?) en deels automatisch (hoe ingewikkeld is het taalgebruik?) gecodeerd. Zij vonden dat hoger opgeleiden deels hun taalgebruik aanpasten: de teksten op de website voor hoger opgeleiden bevatten inderdaad ingewikkelder taalgebruik.

4 Het boek van Pauwels is gratis beschikbaar via de *Digitale Bibliotheek voor de Nederlandse Letteren*: https://www.dbnl.org/tekst/pauw022plaa01_01/. De dialectkaarten staan helemaal achterin het boek.

4.7 NON-VERBALE COMMUNICATIE

Tot nu toe hebben we alleen gekeken naar *verbale* communicatie. Dat wil zeggen: de verschillende aspecten van gesproken en geschreven taal. Maar wanneer mensen met elkaar praten wordt er ook veel *non-verbale* informatie gecommuniceerd, via de prosodie (intonatie van wat er gezegd wordt), gelijktijdige gebaren (Engels: *co-speech gestures*), gezichtsuitdrukkingen, en de lichaamshouding die spreker en toehoorder aannemen. Deze signalen verrijken de interactie, en kunnen de betekenis van wat er gezegd wordt beïnvloeden of zelfs omkeren. Binnen een inhoudsanalyse kunnen non-verbale signalen gebruikt worden om meer te leren over de zenders van een boodschap of de deelnemers aan een gesprek, bijvoorbeeld:[5]

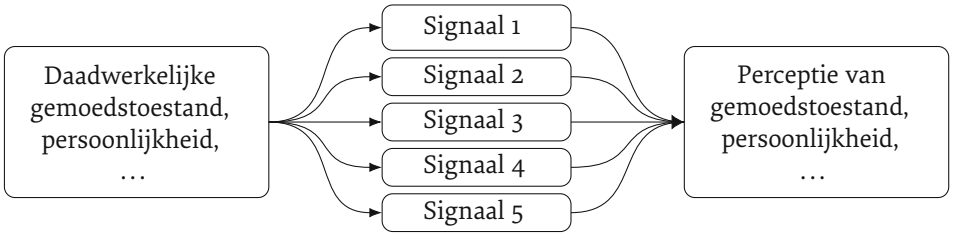
- Het daadwerkelijke of waargenomen karakter van een spreker. Bijvoorbeeld als iemand een dominante houding aanneemt, of zich juist erg klein maakt.
- De daadwerkelijke of waargenomen emotionele toestand van een spreker. Bijvoorbeeld als iemand glimlacht, of wild gesticuleert tijdens een gesprek.
- De daadwerkelijke of waargenomen verhoudingen tussen deelnemers aan een gesprek. Bijvoorbeeld als de gesprekspartners beiden een gesloten houding aannemen.

Zoals de formulering van bovenstaande punten al suggereert, moeten we voorzichtig zijn met de conclusies die we trekken uit non-verbale signalen. Hoe zeker kun je zijn van iemands gemoedstoestand, zonder het die persoon te vragen? We kunnen dit probleem illustreren aan de hand van Brunswik's Lens model (Brunswik, 1955b, 1955a; Osterholz et al., 2021), dat getoond wordt in Figuur 4.2. Aan de linkerzijde van het model zien we de daadwerkelijke gemoedstoestand of persoonlijkheid die we willen meten. Afhankelijk van je persoonlijkheid of gemoedstoestand vertoon je ook ander gedrag, waardoor we verschillende signalen kunnen waarnemen die in meer of mindere mate gecorreleerd zijn met je gemoedstoestand of persoonlijkheid. Op basis van jouw gedrag (in het model: al die non-verbale signalen) kunnen we ons een beeld vormen van jouw gemoedstoestand of persoonlijkheid. In dit proces is er veel ruis aanwezig, waardoor onze perceptie van iemands persoonlijkheid of gemoedstoestand niet altijd overeenkomt met de daadwerkelijke persoonlijkheid of gemoedstoestand.

Meer weten?

Je kunt meer lezen over non-verbale communicatie in het overzichtsartikel van Hall et al. (2019), of in het boek van Hall & Knapp (2013).

5 Voor sommige auteurs valt zelfs kledingkeuze onder de paraplu van non-verbale communicatie. Kleding is immers een manier om jouw persoonlijkheid (bijvoorbeeld bescheiden of uitbundig) uit te dragen. (Zie Hall et al. (2019) voor verdere discussie.)



FIGUUR 4.2 *Brunswik's Lens model.*

4.8 TOT SLOT

Dit hoofdstuk gaf een kort overzicht van de basis van de taalwetenschap, waarbij we hebben gekeken naar verschillende deelgebieden die relevant zijn als je een inhoudsanalyse wil doen. Maar dit is natuurlijk nog maar het topje van de ijsberg. Hopelijk heb je in ieder geval kunnen zien dat taalkundige kennis inderdaad relevant is voor het ontwerpen van een inhoudsanalyse.

Meer weten?

Als je meer wil weten over taalwetenschap, dan kun je bijvoorbeeld beginnen met één van de volgende boeken:

- Handboek taalkunde (2022, eerste editie onder redactie van Banga et al.)
- Taal en taalwetenschap (2023, derde editie onder redactie van Don et al.)

HOOFDSTUK 5
Een codeboek ontwikkelen

5.1 INLEIDING

Een codeboek is een soort handleiding die aangeeft hoe codeurs je data moeten coderen. Daarmee leg je de basis voor je inhoudsanalyse: hoe duidelijker de instructies zijn, des te betrouwbaarder is het codeerproces. In het boek van Wester (2006) wordt ook verwezen naar het codeboek als *waarnemingsinstrument*. Die omschrijving geeft al aan dat een codeboek niet neutraal is: het beperkt de blik van de codeurs, en geeft ook aan hoe ze naar de data moeten kijken. Daarom is het ook belangrijk dat je goed nadenkt over de validiteit van het codeboek.

In dit hoofdstuk leer je hoe codeboeken eruit zien, wat de vereisten zijn, en hoe je zelf een codeboek kunt ontwikkelen. Daarnaast bespreken we ook het codeer*formulier*, dat de codeurs moeten invullen aan de hand van het codeboek.

5.2 HET PROCES

Bij de ontwikkeling van een codeboek kunnen we verschillende stappen onderscheiden die hieronder worden besproken (§5.2.1–§5.2.4). Heel soms hoef je alle stappen maar één keer te doorlopen, maar meestal ontwikkel je het codeboek in een *iteratief* proces: je maakt een eerste versie, probeert het codeboek en -formulier uit, maakt wat aanpassingen, probeert het nog een keer uit, en zo ga je verder tot je tevreden bent over zowel het codeboek als het formulier.

5.2.1 VARIABELEN SPECIFICEREN

Het maken van een codeboek begint bij de variabelen waar je in geïnteresseerd bent. We hebben al verschillende voorbeelden gezien in artikelen die eerder besproken zijn in dit boek:

- Antheunis et al. (2011) waren geïnteresseerd in **self-disclosure** en het stellen van **vragen**.

- Braun et al. (2019) waren onder andere geïnteresseerd in **zelf-verwijzingen**, **verwijzingen naar anderen**, en **emotioneel taalgebruik**.
- Van Hooijdonk & Liebrecht (2021) waren geïnteresseerd in **responsstrategieën** in webcaregesprekken op Twitter.
- Van der Zanden et al. (2018) waren onder andere geïnteresseerd in de **taalfouten** die mensen maken in profielteksten op datingsites.

5.2.2 DEFINIËREN

Als je geïnteresseerd bent in een bepaalde variabele, dan is het een goed idee om die variabele precies te definiëren: wat bedoel je precies met een *taalfout*, *zelfverwijzing*, of *responsstrategie*? Wat hebben andere onderzoekers daar eigenlijk eerder over gezegd? Zijn er misschien ook verschillende sub-categorieën? (Bijvoorbeeld verschillende soorten taalfouten, responsstrategieën of emotioneel taalgebruik.)

5.2.3 OPERATIONALISEREN

De volgende stap is om te kijken hoe je de verschillende variabelen meetbaar kunt maken, oftewel operationaliseren. Daarvoor heb je twee onderdelen nodig: een **codeboek** en een **codeerformulier**. Deze onderdelen staan hieronder omschreven.

Een codeboek maken

Een codeboek is een handleiding waarin staat omschreven wat codeurs precies moeten doen en hoe ze te werk moeten gaan. In een codeboek definieer je de verschillende categorieën, en specificeer je waaraan je die categorieën kunt herkennen. Het is daarbij belangrijk om goede voorbeelden te geven, zodat de codeurs precies weten wat ze moeten doen wanneer ze aan de slag gaan. Op dit punt kun je je afvragen of het beter is om een eerder gebruikt codeboek of codeerschema te gebruiken, of om er zelf een te ontwikkelen. Beide mogelijkheden hebben voor- en nadelen.

Voordelen van bestaande codeboeken:

- Het kost minder tijd om een codeboek te ontwikkelen.
- Het codeboek sluit goed aan op eerdere literatuur.
- Het is eenvoudiger om resultaten te vergelijken met eerdere studies.

Nadelen van bestaande codeboeken:

- Bestaande codeboeken geven geen garantie dat ze ook zullen werken voor jouw data.
- Het gebruik van een bestaand codeboek is geen garantie voor betrouwbare resultaten. Je zult nog steeds moeten toetsen of de resultaten betrouwbaar zijn.

Algemeen advies omtrent het gebruik van bestaande codeboeken is om ze altijd kritisch te bestuderen, en om ze waar nodig aan te passen.

Een codeerformulier maken

Een codeerformulier is een bestand (meestal een spreadsheet) waarin codeurs voor ieder item de juiste waarden moeten invoeren. Het codeerformulier kan je onderzoek maken of breken. Als het codeerformulier goed ontworpen is, dan verloopt het codeerproces soepel en kun je de data eenvoudig importeren in jouw favoriete statistiekprogramma. Wanneer het formulier minder goed ontworpen is, dan kan dat invloed hebben op de kwaliteit van de data (dankzij allerlei invoerfouten), en kan het ook een stuk lastiger zijn om de data achteraf te analyseren. Het is dus belangrijk om goed na te denken over het ontwerp van het formulier. Er zijn vier manieren om data handmatig te coderen:

1) Via een spreadsheetprogramma (bijvoorbeeld Excel, Google Sheets, Calc). Je maakt dan één grote tabel waar je alle data in codeert.
2) Via een toegewijd programma (bijvoorbeeld ELAN, Wittenburg et al. (2006); MAST, Cardoso & Cohn (2022)). Je kunt dan de data laden, coderen, en de resultaten exporteren in een grote spreadsheet.
3) Via andere bestaande software (bijvoorbeeld Qualtrics). Je ontwerpt dan zelf een *workflow* om de data te coderen door bijvoorbeeld vragen te beantwoorden.[1]
4) Via een zelf ontworpen programma. Deze optie bespreken we verder niet in dit werkboek.

Wij bespreken in dit werkboek alleen de eerste optie, omdat de andere opties ofwel automatisch een goed geformatteerd bestand opleveren (opties 2 en 3), ofwel een bestand moet opleveren dat voldoet aan dezelfde vereisten als een codeerformulier in een spreadsheetprogramma (optie 4).

5.2.4 UITPROBEREN

Als je een goed beeld hebt van de categorieën die je wil gaan coderen, dan kun je samen met je groepsgenoten een aantal voorbeelden verzamelen om samen te coderen. Vaak geven concrete voorbeelden aanleiding tot een onderlinge discussie, waarna je het codeboek verder kunt aanpassen om toekomstige misverstanden te voorkomen. Het is vooral belangrijk om goed te communiceren met je onderzoeksteam, zodat je het eens bent over de gekozen benadering. Hieronder staat een stappenplan om de uitprobeerfase vorm te geven. Dit kun je ook zien als een vorm van training.

1 Je kunt bijvoorbeeld met Python automatisch een vragenlijst genereren en deze vragenlijst exporteren als een "Advanced Format TXT." Dat wordt dan een groot bestand met alle items onder elkaar, waarbij je bij ieder item dezelfde vraag stelt. Het gaat te ver voor dit werkboek om te vertellen hoe je dit precies moet doen, maar je kunt kijken naar de volgende voorbeelden:
https://github.com/hienhuynhtdn/GPT2andImplicitCausality/tree/main/experiment_2/human_ratings/setup
https://github.com/evanmiltenburg/ReproHum-definition-complexity/tree/main/Experiment/Setup

- Probeer het codeboek samen toe te passen op vijf berichten.[2] Pas het codeboek aan indien nodig.
- Bedenk fictieve berichten die uitdagend zijn om te coderen. Bespreek samen hoe je met die gevallen zou moeten omgaan (ook in het licht van de literatuur die jullie gebruikt hebben bij het opstellen van het codeboek).
- Codeer nu los van elkaar nog eens vijf dezelfde berichten. Bespreek jullie ervaringen, en kijk naar de overeenkomsten en verschillen tussen de resultaten. Pas het codeboek aan waar nodig.
- Bespreek de volgende stappen: gaan jullie nu aan de slag met een grotere hoeveelheid data (in een pilotstudie, zie §5.6), of wil je eerst nog een ronde oefenen?

5.3 HOE ZIET EEN GOED CODEBOEK ERUIT?

Een codeboek bevat ten minste drie onderdelen:

1) Definitie van de categorie. Deze definitie komt meestal uit de literatuur, of sluit in ieder geval goed aan op de literatuur die je gelezen hebt. Dat betekent ook dat je referenties dient te geven in het codeboek, zodat je transparant bent over de oorsprong van de categorie.
2) Specificatie van de eigenschappen waaraan je die categorie kunt herkennen.
3) Een voorbeeld om de definitie mee te illustreren, met uitleg in termen van bovengenoemde eigenschappen.

Daarnaast is het verstandig om ook voorbeelden te geven van twijfelgevallen die wel/niet horen bij de categorie. Door uit te leggen hoe die voorbeelden gecategoriseerd moeten worden, begrijpen codeurs beter wat ze moeten doen.

5.4 HOE ZIET EEN GOED CODEERFORMULIER ERUIT?

Codeerformulieren zijn meestal zo gestructureerd dat iedere rij staat voor een observatie, en iedere kolom voor een eigenschap van die observatie. Die eigenschappen vallen op te delen in:

- **Bestaande metadata**, zoals de bron, URL, auteur, datum, enzovoorts.
- **Organisatorische metadata**, zoals een uniek nummer dat je zelf aan iedere boodschap toegekend hebt, de groep waarbij de boodschap hoort binnen jouw studie, enzovoorts. Als je structuur wil aanbrengen binnen de data kun je daar ook extra kolommen voor aanmaken, bijvoorbeeld oplopende nummers voor ieder bericht binnen een gesprek.

2 Vijf is een willekeurig gekozen aantal. Kies in ieder geval een aantal dat haalbaar is om in korte tijd te analyseren, maar dat tegelijkertijd genoeg informatie geeft om een betrouwbaar codeboek te ontwikkelen.

- **Variabelen die je codeert**, dit zijn alle hoofdcategorieën in je codeboek.
  - Als bepaalde categorieën elkaar uitsluiten[3] (bijvoorbeeld 'algemeen sentiment' waarbij een bericht gecategoriseerd wordt als *positief*, *negatief*, of *neutraal*)', dan kun je ze in dezelfde kolom zetten.
  - Als meerdere labels tegelijkertijd op een bericht van toepassing kunnen zijn (bijvoorbeeld responsstrategieën zoals 'excuses aanbieden' en 'doorverwijzen'), dan heb je meerdere kolommen nodig waarin je die labels apart codeert (bijvoorbeeld met een '1' voor 'aanwezig' en een '0' voor 'afwezig').

Organisatorische metadata worden vaak vergeten bij het maken van codeerformulieren, maar deze metadata zijn ontzettend belangrijk bij de data-analyse, omdat je ze kunt gebruiken om de data te sorteren. Bijvoorbeeld als je gesprekken codeert en wil kijken naar het eerste bericht uit ieder gesprek, dan kun je de data in je spreadsheet gewoon sorteren op de kolom met berichtnummers. Daarbij is het wel belangrijk dat je genoeg organisatorische metadata toevoegt om de originele volgorde van je bestand te kunnen herstellen.

5.5 VOORBEELDEN VAN CODEBOEKEN

Laten we nu gaan kijken naar twee concrete voorbeelden van codeboeken, om beter te begrijpen hoe jouw codeboek er straks uit moet gaan zien.

5.5.1 VAN HOOIJDONK & LIEBRECHT (2018)

hebben berichten (Tweets) op het socialemediaplatform Twitter gecodeerd aan de hand van verschillende categorieën die zij associëren met het fenomeen Conversational Human Voice (CHV). Met andere woorden: elementen die je kunt toevoegen aan een bericht waarmee de boodschap warmer en menselijker overkomt. Deze gespreksstijl kun je contrasteren met een meer zakelijke (Corporate) communicatiestijl. Hieronder in Tabel 5.1 is een deel van hun codeboek weergegeven. Over dit codeboek geven Van Hooijdonk & Liebrecht (2018) aan dat het gebaseerd is op corpus- en literatuuronderzoek naar talige kenmerken van CHV. De bedoeling is dat codeurs per bericht coderen of het conversationeel linguïstisch element wel (1) of niet (0) voorkomt.

3 In het Engels wordt hiervoor de term *mutually exclusive* gebruikt.

Conversationeel linguistisch element	Bronnen
1) Personalisatie	Kelleher, 2009; Kelleher & Miller, 2006; Van Noort et al., 2014.
a. Ondertekening Met een ondertekening wordt duidelijk gemaakt van wie het bericht binnen de organisatie afkomstig is. IFIDs: Aan het einde van het bericht, naam (Wies) of initialen (WP), vaak gekenmerkt door een dakje (^). Voorbeelden: Mooie foto! :-) ^JB Ik ga het doorgeven. ^Sam	Crijns et al., 2017; Gretry et al., 2017; Huibers & Verhoeven, 2014; Kerkhof et al., 2011; Kwon & Sung, 2011; Lillqvist & Louhiala-Salminen, 2013; Park & Lee, 2013; Rybalko & Seltzer, 2010; Schamari & Schaefers, 2015; Van Hooijdonk & Liebrecht, 2015; Van Noort et al., 2014.
b. Persoonlijke begroeting De stakeholder wordt met zijn naam begroet door de organisatie. IFIDs: Aan het begin van het bericht, een begroetingswoord (Hallo, Ha, Beste, Hoi, Hi) en de naam van de stakeholder (anders dan zijn twitternaam). Voorbeelden: Hoi Stephanie, wat naar! Hallo Marije, excuses voor de vertraagde beantwoording	Crijns et al., 2017; Gretry et al., 2017; Lillqvist & Louhiala Salminen, 2013; Pollach, 2005; Van Hooijdonk & Liebrecht, 2015; Van Noort et al., 2014.
c. Persoonlijk aanspreken De stakeholder wordt persoonlijk aangesproken door de organisatie. IFIDs: De stakeholders' naam wordt genoemd (ander dan bij een persoonlijke begroeting), of informele persoonlijke of bezittelijke voornaamwoorden worden gebruikt (je, jij, jouw). Voorbeelden: Ik hoop dat er snel een oplossing wordt gevonden Colinda. We hebben je melding over hondenpoep doorgegeven.	Crijns et al., 2017; Gretry et al., 2017; Huibers & Verhoeven, 2014; Kelleher, 2009; Kelleher & Miller, 2006; Pollach, 2005; Kwon & Sung, 2011; Sparks et al., 1997; Van Hooijdonk & Liebrecht, 2015; Van Noort et al., 2014.

TABEL 5.1 *Deel van het codeboek van Van Hooijdonk & Liebrecht (2018).*

Test jezelf

Lees het bovenstaande fragment en beantwoord voor jezelf de onderstaande vragen.

- Hoe is het codeboek gestructureerd?
- Wat valt je op?
- Wat vind je goed aan het codeboek?
- Zie je verbeterpunten?
- Welke uitdagingen zie je voor de codeurs?
- Welke uitdagingen zie je voor de analyse van de resultaten?

Evaluatie van het codeboek

Als je de bovenstaande vragen voor jezelf hebt beantwoord, kunnen we samen kijken naar een aantal eigenschappen van het codeboek.

1) Het codeboek is georganiseerd aan de hand van hoofd- en subcategorieën. De hoofdcategorie hier is **Personalisatie**, en de subcategorieën zijn **ondertekening, persoonlijk begroeten,** en **persoonlijk aanspreken.**
2) Iedere subcategorie begint met een definitie, gevolgd door IFIDs (Illocutionary Force Indicating Devices, oftewel: uitdrukkingen die iets zeggen over de aard van de boodschap) en een aantal voorbeelden. Daardoor krijgen codeurs een goed beeld van iedere categorie.
3) Iedere (sub)categorie is voorzien van referenties. Daarmee is meteen duidelijk waar de categorieën precies vandaan komen. Bij het herzien van het codeboek zou je hierdoor ook eenvoudig kunnen controleren of de wijzigingen in lijn zijn met eerdere literatuur.
4) De voorbeelden zijn gebaseerd op echte berichten. Dat zorgt ook voor vertrouwen, dat je weet dat de categorieën ook echt zo voorkomen zoals ze worden omschreven.
5) Het coderen is zeer eenvoudig: per bericht hoeft er alleen maar een 1 of een 0 ingevuld te worden in het codeerformulier.

Codeformulier

Het formulier was gemaakt in Excel, en zag er ongeveer zo uit als in Tabel 5.2. In werkelijkheid was het formulier natuurlijk wat groter, met meer organisatorische metadata, en meer categorieën om te coderen.

Gespreks-nummer	Bericht van stakeholder	Bericht van organisatie	Personalisatie: persoonlijke begroeting	Personalisatie: aanspreek-vormen	Personalisatie: ondertekening
1	"..."				
1		"..."	1	0	1
1	"..."				
1		"..."	0	1	1
1	"..."				

TABEL 5.2 *Codeformulier behorend bij het codeboek van Van Hooijdonk & Liebrecht (2018)*

Het voordeel van een spreadsheetprogramma zoals Excel, Numbers, of Google Sheets is dat de data zonder veel moeite klaar is om te analyseren in je favoriete statistiekprogramma. Het nadeel van deze programma's is dat je eenvoudig typefouten kunt maken. Deels zijn deze te voorkomen (je kunt het programma zo instellen dat er alleen bepaalde waarden zijn toegestaan in een kolom) of te detecteren (door een formule te schrijven om alle waarden te tellen), maar het blijft belangrijk om goed op te letten en tussentijds back-ups te maken van je data.

5.5.2 VAN DER ZANDEN ET AL. (2018)

Zoals we eerder hebben gezien, keken Van der Van der Zanden et al. (2018) naar het taalgebruik in online datingprofielen. Daarbij hebben ze zich in hun codeboek vooral gericht op taalfouten.[4] (Andere eigenschappen van de teksten werden automatisch berekend, en dus niet handmatig gecodeerd.) Je kunt het codeboek hier downloaden: https://doi.org/10.34894/HLYUNB. Zoals je ziet, is het codeboek uitgebreider (en ingewikkelder) dan het bovenstaande voorbeeld.

Test jezelf

Neem het codeboek goed door voordat je verder leest.

- Hoe is het codeboek gestructureerd?
- Wat valt je op?
- Wat vind je goed aan het codeboek?
- Zie je verbeterpunten?
- Welke uitdagingen zie je voor de codeurs?
- Welke uitdagingen zie je voor de analyse van de resultaten?

Evaluatie van het codeboek

Als je de bovenstaande vragen voor jezelf hebt beantwoord, kunnen we samen kijken naar een aantal eigenschappen van het codeboek.

1) Algemeen kunnen we zien dat Van der Zanden en collega's in hun codeboek de volgende kolommen hebben gebruikt:
   - **Categorie**: naam van de categorie.
   - **Definitie**: definitie van de categorie.
   - **Uitleg en aandachtspunten**: een beschrijving die uitlegt hoe je de categorie kunt herkennen, en punten waar je op moet letten om de categorieën goed uit elkaar te houden en niet te verwarren.
   - **Subcategorie**: verschillende soorten taalfouten die binnen de eerdergenoemde categorie vallen.

4 Het codeboek bevat de volgende categorieën, waarbij iedere categorie een eigen definitie heeft: Grammaticale fout, Spelfout, Typfout, Spatiefout, Interpunctiefout, Hoofdletterfout, Overig.

- **Voorbeelden**: voorbeelden voor iedere subcategorie die genoemd wordt. Deze voorbeelden worden gebruikt om de subcategorieën concreter te maken.

2) De categorieën in het codeboek worden duidelijk van elkaar onderscheiden, en het codeboek anticipeert ook op mogelijk verwarrende situaties. Zo wordt er over dt-fouten (bijvoorbeeld *ik wordt* in plaats van *ik word*) expliciet aangegeven dat ze horen bij de incongruentiefouten (de vervoeging van het werkwoord past niet bij het onderwerp van de zin) en niet bij de spelfouten (per ongeluk een extra letter aan het woord toegevoegd).
3) Opvallend is dat de categorieën wel, maar de subcategorieën niet worden gedefinieerd in het codeboek. Of tenminste: er wordt wel een aanwijzende definitie gegeven door middel van voorbeelden, maar geen verdere uitleg zoals bij de hoofdcategorieën.
4) Een andere opvallende eigenschap aan het codeboek is dat de verschillende categorieën op verschillende niveaus worden gecodeerd. Sommige taalfouten (zoals grammaticale fouten) worden op zinsniveau gecodeerd, terwijl andere taalfouten (zoals spelfouten) op woordniveau worden gecodeerd. Hiervoor wordt door de auteurs speciale software gebruikt om relevante onderdelen van de verschillende teksten te kunnen selecteren en te categoriseren.
5) Ten slotte is het ook belangrijk om te beseffen dat het codeerproces eigenlijk twee stappen heeft: 1. de hoofdcategorie bepalen, 2. de subcategorie bepalen. Voor iedere stap zou je iets moeten zeggen over de betrouwbaarheid van die stap. (Dat brengt in dit geval nog wat complicaties met zich mee, maar dat kun je verder lezen in het artikel.)

5.6 PILOTSTUDIES

Wanneer je tevreden bent met je codeboek, kun je verder gaan met een *pilotstudie*. De term 'pilotstudie' verwijst naar de verkennende fase van het onderzoekstraject waarin je de hoofdstudie eerst 'in het klein' uitprobeert, om eventuele (vaak praktische) problemen voor te kunnen zijn. In een pilot kun je met je onderzoeksteam dezelfde subset van de data individueel coderen, om vervolgens de resultaten met elkaar te vergelijken. Ook kun je op dat moment alle relevante statistieken berekenen met de data die je gecodeerd hebt. Een pilotstudie zorgt dus voor een *proof-of-concept*. Daarbij kun je vragen stellen als:

1) Hoe betrouwbaar zijn de resultaten?
   a) Hoe duidelijk is het codeboek?
   b) Zijn alle codeurs het eens met elkaar?
2) Hoe realistisch is de planning?
   a) Hoeveel tijd kost het coderen nu echt?
   b) Hoeveel tijd gaat het straks kosten om dit uit te voeren?
3) Kun je de onderzoeksvraag straks wel beantwoorden?
   a) Staat de data straks in het juiste format om de resultaten te analyseren?
   b) Heb je genoeg informatie om de onderzoeksvraag te beantwoorden?

5.6.1 PROBLEMEN VOORKOMEN

Zoals de bovenstaande vragen al suggereren, kunnen we problemen bij inhoudsanalyses ruwweg categoriseren in drie groepen:

1) **Onbetrouwbare data**: wanneer er teveel variatie optreedt tussen de labels die worden toegekend door verschillende codeurs.
2) **Onrealistische planning**: wanneer het niet haalbaar lijkt te zijn om de geplande hoeveelheid data binnen de gestelde tijd te verzamelen.
3) **Onbruikbare data**: wanneer je erachter komt dat de onderzoeksvraag eigenlijk niet beantwoord kan worden met de data die je verzamelt.

Veel van bovenstaande problemen kun je (deels) al voorkomen voor de pilotfase. Je kunt bij het nadenken over de onderzoeksvraag en het opstellen van het codeboek alvast wat voorbeelden samen bekijken, en ze hardop analyseren en bespreken. Daarbij kun je vragen stellen als:

- Wat is er lastig of tijdrovend? Valt dat te vereenvoudigen?
- Voorzien jullie onenigheid? Hoe valt dat te voorkomen?
- Ontbreekt er iets in het codeerschema?
- Hoe gaan jullie deze data gebruiken om de onderzoeksvraag te beantwoorden?

5.6.2 STAPPENPLAN

Het stappenplan voor een pilotstudie is als volgt:

Stap 0: Bedenk hoe je studie eruit gaat zien. Je kunt geen pilotstudie uitvoeren zonder de benodigde materialen. Zorg voor een duidelijk codeboek en een helder codeerformulier. Bespreek vooraf met je groep wat het doel is van de pilotstudie en wat er van iedereen verwacht wordt.

Stap 1: Bereid de data voor om te coderen. Iedereen gaat dezelfde items coderen. Maak vooraf een inschatting van de tijd die het kost om alles te coderen. Zorg ervoor dat er genoeg items zijn om een goed beeld te hebben van de betrouwbaarheid van het codeerproces, zonder dat het te veel tijd kost. Hoeveel data kun je bijvoorbeeld coderen in 2-3 uur?

Stap 2: individueel coderen. Iedereen codeert de data individueel, zonder daarbij te overleggen (op welke manier dan ook). Maak wel aantekeningen bij het coderen, bijvoorbeeld met een extra kolom in het codeerformulier waarin je al je observaties en vragen kwijt kunt. Zo weet je zeker dat je de items terug kunt vinden waarover je getwijfeld hebt tijdens de pilotstudie.

(Stap 2.5: herstart) Als je helemaal vastloopt tijdens het individueel coderen, neem dan wel contact op met je groepsgenoten. Bedenk wat wel/niet werkt aan het codeerproces, en begin opnieuw. Het heeft geen zin om de intercodeursbetrouwbaarheid te gaan berekenen als een deel van de codeurs niet eens aan coderen toekomt.

Stap 3: bereken de intercodeursbetrouwbaarheid. In het volgende hoofdstuk leer je meer over betrouwbaarheid en statistiek, maar voor nu is het goed om te weten dat er verschillende

statistieken zijn die je kunt berekenen om te kijken in hoeverre er overeenstemming is tussen de verschillende codeurs. Je kunt ook in je favoriete spreadsheetprogramma kijken waar de verschillen nou eigenlijk zitten tussen de verschillende codeurs. (Dit kan automatisch via een formule waarbij je kijkt of de waarden in twee verschillende kolommen gelijk zijn aan elkaar.)

Stap 4: herzie het codeboek. Verduidelijk de categorieën in het codeboek, en geef extra instructies en voorbeelden om aan te geven hoe codeurs met verschillende situaties om moeten gaan. Zo wordt de betrouwbaarheid van het codeerproces een stuk hoger. (Bedenk wel wat dit mogelijk doet met de validiteit van jullie studie!)

Stap 5: schrijf alles op. Als je alles netjes hebt bijgehouden, kun je in je verslag aangeven welke knelpunten er waren bij het ontwikkelen van het codeboek. Beschrijf ook de intercodeursbetrouwbaarheid tijdens de pilotstudie.[5] De beste tijd om iets op te schrijven is wanneer je er mee bezig bent. Herinneringen vervagen snel, en je onthoudt altijd minder dan je denkt te gaan onthouden.

5.7 ONAFHANKELIJKHEID VAN DE CODEURS?

Bij het doen van onderzoek zijn er altijd spanningen tussen verschillende idealen. Een van die idealen is om inhoudsanalyses uit te laten voeren door *onafhankelijke codeurs*. Onafhankelijkheid kan in de context van inhoudsanalyse twee dingen betekenen:

1) Codeurs zijn onafhankelijk van elkaar bij het coderen van de data. Zo hebben we twee onafhankelijke metingen die we later met elkaar kunnen vergelijken.
2) Codeurs zijn onafhankelijk van de auteurs, zodat ze niet op de hoogte zijn van de hypothesen die zijn opgesteld over de resultaten van de inhoudsanalyse. Zo worden ze ook niet beïnvloed door de verwachtingen van de auteurs.

Beide doelen zijn lovenswaardig, maar de praktijk is weerbarstig. Vooral het tweede doel is vaak niet haalbaar, omdat het veel tijd en geld kost om iemand te trainen om data te coderen. Daarom zie je vaak dat auteurs de data zelf coderen. Dat heeft niet alleen maar nadelen. Een groot voordeel van coderende auteurs is dat ze de literatuur goed kennen, en de data kunnen coderen met alle relevante kennis (welke eigenschappen zijn van belang?) in het achterhoofd.[6]

5 In 'echte' publicaties zie je dit meestal niet; daar wordt het codeboek doorontwikkeld totdat de betrouwbaarheid hoog genoeg is. De uiteindelijke betrouwbaarheid aan het eind van de ontwikkelfase wordt dan benoemd. Soms is dit de enige betrouwbaarheidsscore die je ziet, maar soms zie je ook nog een betrouwbaarheidsscore die gerapporteerd wordt over de hoofdcodeerfase.

6 Dat ontslaat auteurs niet van de plicht om een goed codeboek op te stellen dat ook duidelijk is voor anderen; het idee van wetenschappelijk onderzoek is dat het transparant en herhaalbaar moet zijn, dus dat anderen ook moeten kunnen begrijpen hoe de data is gecodeerd en met dezelfde data en hetzelfde codeboek tot dezelfde resultaten zouden moeten kunnen komen.

Als je niet werkt met onafhankelijke codeurs is het belangrijk om de invloed van voorkennis te minimaliseren. Hieronder bespreken we twee manieren om dat te doen.

Gegevens afschermen
Als de codeurs niet onafhankelijk zijn, is het soms wel mogelijk om bepaalde gegevens af te schermen tijdens het codeerproces. Een voorbeeld hiervan is de studie van Mui et al. (2017, overigens wel met onafhankelijke codeurs), waarin gezichtsuitdrukkingen van participanten werden gecodeerd. In deze studie waren de codeurs niet alleen onbekend met de hypothesen, maar ze wisten ook niet in welke experimentele conditie de participanten waren ingedeeld. Daardoor konden ze ook niet (bewust of onbewust) een verband leggen tussen de conditie en de afhankelijke variabele.

Objectiviteit van het codeerproces maximaliseren
Subjectieve beslissingen zijn eenvoudiger te beinvloeden dan beslissingen die genomen kunnen worden op basis van objectief waarneembare criteria. Daarom is het verstandig om kritisch te kijken naar je codeboek, en te bedenken welke hoe subjectief de verschillende categorieën zijn. Het artikel van Mui et al. (2017) is wederom een goed voorbeeld: bij het coderen van de gezichtsuitdrukkingen van de participanten wilden de auteurs van iedere glimlach weten hoe oprecht die uitdrukking was. Dat zou je in theorie kunnen meten door codeurs te vragen om op een zevenpuntsschaal aan te geven hoe vals of oprecht de glimlach overkomt. Maar zo'n persoonlijk oordeel is niet de meest betrouwbare manier om de oprechtheid van een glimlach te meten. In plaats daarvan hebben de auteurs gebruik gemaakt van een bestaand referentiekader: het *Facial Action Coding System* (FACS; Ekman & Friesen (1978)), waarbij gekeken wordt naar spiersamentrekkingen in het gezicht die gepaard gaan met vals en oprecht glimlachen. Daarmee is het codeerproces gebaseerd op aanwijsbare (manifeste) eigenschappen, waardoor objectiever vast te stellen is welke categorie (oprecht of niet) van toepassing is.

5.8 TOT SLOT

Een codeboek is nooit echt helemaal af. Tijdens het ontwikkelen van het codeboek kun je er door middel van pilotstudies voor proberen te zorgen dat de instructies redelijk dichtgetimmerd zijn, maar er blijven altijd twijfelgevallen. Wester (2006) raadt daarom aan om ook na de trainingsfase te zorgen voor een systeem van "memo's of tussentijdse codeurbijeenkomsten" waarbij de codeurs elkaar op de hoogte houden van de uitdagingen die ze tegengekomen zijn en afspreken hoe ze daarmee om moeten gaan, zodat alle data eenduidig gecodeerd wordt.

HOOFDSTUK 6
Betrouwbaarheid en statistiek

6.1 INLEIDING

Betrouwbaarheid verwijst naar de precisie van een meetinstrument, en het vermogen van dat meetinstrument om bij herhaling dezelfde resultaten te produceren. Bij een inhoudsanalyse is dat instrument geen apparaat maar een proces: het coderen van de data. In dit hoofdstuk leer je meer over de betrouwbaarheid van inhoudsanalyses, en over de manieren om die betrouwbaarheid te kwantificeren. Daarnaast gaan we kort in op andere statistieken die je kunt gebruiken om een inhoudsanalyse te rapporteren.

6.1.1 INTERCODEURSBETROUWBAARHEID

Wanneer je de literatuur erop naslaat, valt op dat er veel verschillende termen zijn voor de betrouwbaarheid van een inhoudsanalyse. Bijvoorbeeld:

- Intercodeursbetrouwbaarheid
- Intercoder/interrater/interjudge reliability
- Inter annotator agreement

De verschillen in terminologie zeggen vooral iets over het vakgebied waar de auteurs vandaan komen. Computerwetenschappers zeggen bijvoorbeeld eerder "annoteren" dan "coderen" en dus zie je in artikelen vaak iets over *inter-annotator agreement*. Maar bovenstaande termen verwijzen allemaal naar hetzelfde idee: de betrouwbaarheid waarmee menselijke codeurs een verzameling data kunnen labelen of categoriseren. Dat gebeurt meestal door verschillende codeurs dezelfde data te laten coderen, en de resultaten vervolgens met elkaar te vergelijken. Zo kunnen we zien in hoeverre de codeurs hetzelfde doen. Wij gebruiken in dit boek de term *intercodeursbetrouwbaarheid* voor dit idee. Wester (2006) maakt hierbij nog een verder onderscheid tussen *codeerovereenstemming* en *accuraatheid*. Wanneer we het gedrag van twee

codeurs met elkaar vergelijken, dan is het fijn om een hoge overeenstemming te zien tussen hun oordelen. Maar die overeenstemming betekent nog niet dat de data *juist* gecodeerd is. Daarbij is het volgens Wester (2006) van belang om de oordelen van een codeur te vergelijken met een *normcodering* van de onderzoeker of een deskundige. Wanneer dat zo is, kunnen we zeggen dat de codeurs accuraat zijn.

6.1.2 INTRACODEURSBETROUWBAARHEID

Naast de verschillen en overeenkomsten tussen verschillende codeurs, kunnen we ook kijken naar de consistentie van individuele codeurs: worden vergelijkbare items op een vergelijkbare manier gecodeerd? Dit noemen we *intracodeursbetrouwbaarheid*.[1,2] Vaak wordt hiernaar gekeken binnen langere onderzoeksprojecten: codeer je aan het eind van het project nog steeds hetzelfde als aan het begin? Bij kleine projecten is intracodeursbetrouwbaarheid soms minder zinvol, omdat codeurs dan nog precies kunnen onthouden hoe ze eerdere beslissingen hebben genomen (en dan test je dus geheugen in plaats van consistentie).

6.2 BETROUWBAARHEID METEN

Er zijn verschillende statistische formules bedacht om iets te zeggen over de betrouwbaarheid van een inhoudsanalyse. Deze formules hebben in ieder geval gemeen dat ze data nodig hebben van ten minste twee codeurs, die dezelfde items *onafhankelijk van elkaar* hebben gecodeerd. Door de resultaten van de codeurs met elkaar te vergelijken via een statistische formule, krijg je een *betrouwbaarheidsscore*. De hoogte van die score zegt iets over de betrouwbaarheid van het codeerproces. De meest gebruikte statistieken om de betrouwbaarheid van het codeerproces te meten zijn:

- Percentage agreement
- Cohen's kappa
- Fleiss' kappa
- Krippendorff's Alpha

Er zijn verschillende andere statistieken die gebruikt worden om betrouwbaarheid te meten. Zie hiervoor bijvoorbeeld het overzicht van Hayes & Krippendorff (2007) of het handboek van Gwet (2014).

---

1 Wester (2006) gebruikt hiervoor de term *stabiliteit*, maar de term *intra-coder reliability* wordt internationaal vaker gebruikt.

2 Een manier om het verschil tussen *intercodeursbetrouwbaarheid* en *intracodeursbetrouwbaarheid* te onthouden is om de woorden verder te analyseren. De woorden *inter* en *intra* komen uit het Latijn en betekenen respectievelijk 'tussen' en 'binnen.' Deze woorden worden ook elders in het Nederlands gebruikt. Zo is een *intercity* een trein die *tussen* meerdere steden rijdt, *intranet* een computernetwerk dat *binnen* organisaties gebruikt wordt, en *internet* een verbinding *tussen* alle computers.

6.2.1 PERCENTAGE AGREEMENT

De eenvoudigste manier om de overeenkomst tussen twee codeurs te berekenen is om te kijken hoe vaak ze het eens zijn, en om dan het percentage te berekenen. Als twee codeurs samen 140 items dubbel coderen, en ze zijn het 70 keer eens, dan is er een overeenstemming van 50%. Het **voordeel** van deze methode is dat het heel eenvoudig is om een score te berekenen, en dat de score ook direct inzicht geeft: je weet meteen wat de score betekent. Het grote **nadeel** van deze methode is dat er geen kanscorrectie plaatsvindt.

Kanscorrectie

Kanscorrectie verwijst naar het idee dat je ook kijkt naar de kans dat codeurs het per toeval eens zijn, en die informatie meeneemt in het berekenen van de betrouwbaarheidsscore. Je kent dit idee misschien al van *Multiple Choice*-tentamens: als er vier opties zijn (*a*,*b*,*c*, en *d*), dan is de kans dat je het antwoord toevallig goed hebt 25%. Dus als je een willekeurige tentamenzaal binnenloopt bij een vak dat je nooit gevolgd hebt, kun je op basis van kans alleen al 25% van de vragen goed beantwoorden (bijvoorbeeld door gewoon overal *c* in te vullen). Hierdoor kun je zonder kanscorrectie een hoger cijfer krijgen dan je eigenlijk verdient.

Om tentamencijfers te corrigeren wordt er bijvoorbeeld gewerkt met een drempelwaarde die eerst behaald moet worden voordat vragen echt punten beginnen op te leveren.[3] Stel je voor dat er 40 tentamenvragen zijn met vier keuzes per vraag. In dat geval kun je er in theorie al 10 goed gokken. Als een docent vindt dat studenten in ieder geval de helft van de antwoorden goed moet kunnen beantwoorden, dan kan de docent ervoor kiezen om pas een voldoende te geven wanneer studenten de helft van de resterende 30 vragen goed hebben. Met andere woorden, pas vanaf 10 + (50% * 30) = 25 goede antwoorden. Daarmee is die docent strenger dan wanneer er gerekend zou worden met 50% van 40 vragen = 20 goede antwoorden, maar het tentamen met kanscorrectie geeft (in theorie) wel een beter beeld van de daadwerkelijke kennis van de studenten.

Kanscorrectie bij inhoudsanalyse is vergelijkbaar met kanscorrectie bij tentamens. Stel je voor dat je een item wil coderen met twee niveaus, bijvoorbeeld of er wel/geen sarcasme gebruikt wordt in een uitspraak, dan zijn er vier mogelijke uitkomsten die we kunnen illustreren als in Tabel 6.1. In de helft van de gevallen (groen gemarkeerd) zijn de codeurs het met elkaar eens, en in de andere helft (ongemarkeerd) niet. Je zou misschien denken dat de kans dat de codeurs het eens zijn daarmee dus ook 50% is, maar dan houden we geen rekening met de data: de kans dat codeurs het met elkaar eens zijn is afhankelijk van de mate waarin ze allebei voor de verschillende opties hebben gekozen.

3 Er zijn verschillende formules die worden gebruikt voor kanscorrectie. Het voorbeeld gebruikt een relatief eenvoudige methode om het idee van kanscorrectie te kunnen illustreren.

		Codeur 2	
		Wel sarcasme	**Niet sarcasme**
Codeur 1	**Wel sarcasme**	Optie 1	Optie 2
	Niet sarcasme	Optie 3	Optie 4

TABEL 6.1 *Illustratie van de overeenstemming tussen twee codeurs, bij het coderen van een variabele (sarcasme) met twee opties (wel en niet van toepassing). De cellen komen overeen met de keuzes die de twee codeurs gemaakt hebben. De groene cellen (Opties 1 en 4) bevatten de gevallen waarbij de codeurs het eens zijn.*

6.2.2 COHEN'S KAPPA

Cohen (1960)'s kappa (Cohen, 1960) is een veelgebruikte score om de intercodeursbetrouwbaarheid te berekenen, waarbij er gecorrigeerd wordt voor de gokkans. De gebruikte formule is als volgt:

$$\kappa = \frac{\text{Geobserveerde overeenkomst - Verwachte overeenkomst}}{\text{1 - Verwachte overeenkomst}}$$

De geobserveerde overeenkomst is de proportie van items die door beide codeurs hetzelfde is beoordeeld, en de verwachte overeenkomst komt overeen met de proportie die je zou verwachten op basis van het gedrag van de codeurs (waarbij het gedrag is: hoe vaak ze voor de verschillende categorieën hebben gekozen). Waar *percentage agreement* een percentage oplevert, krijg je bij Cohen's Kappa altijd een score tussen 0 (overeenkomst op kansniveau) en 1 (perfecte overeenkomst). Kappa kan ook negatief zijn, per toeval of wanneer codeurs het met elkaar oneens zijn. Je kunt Cohen's kappa alleen gebruiken om twee codeurs met elkaar te vergelijken, en alleen voor *nominale variabelen*. (Dus wel het categoriseren van items, maar geen Likert-scores om de items te beoordelen.)

Interpretatie van Cohen's kappa

Wanneer je een score hebt berekend, zul je je waarschijnlijk afvragen wat die score eigenlijk betekent. Bij een score van 1 is de conclusie duidelijk: Hoera! De codeurs zijn het altijd eens met elkaar. Bij een score van 0 is er duidelijk werk aan de winkel. Maar hoe zit het met de waarden daartussen? Wat betekent een score van 0.4 of 0.7 eigenlijk? Veel onderzoekers verwijzen naar de schaal van Landis & Koch (1977), die gegeven wordt in Tabel 6.2:

Score	Oordeel
<0.00	Slecht
0.00-0.20	Licht
0.21-0.40	Redelijk
0.41-0.60	Gemiddeld
0.61-0.80	Substantieel
0.81-1.00	Bijna perfect

TABEL 6.2 *De schaal van Landis & Koch (1977) die vaak gebruikt wordt om Cohen's kappa te interpreteren. Volgens deze schaal heb je voor je onderzoek een score van 0.6 of hoger nodig.*

De consensus in de literatuur lijkt te zijn dat je in ieder geval een kappa-score van 0.6 moet hebben om voorzichtige conclusies te kunnen trekken uit de data, en het liefst een score van 0.8 of hoger. Daarbij moeten we wel zeggen dat de interpretatie van kappa niet alleen afhangt van de score zelf, maar ook van de variabele. Sommige variabelen zijn bijvoorbeeld veel eenvoudiger te coderen dan andere variabelen. Voor duidelijke en objectieve variabelen (bijvoorbeeld: of een bericht een verontschuldiging bevat) ligt de standaard hoger dan voor variabelen die inherent subjectiever zijn (bijvoorbeeld: of een bericht sarcastisch is of niet). Uiteindelijk moeten we ons voor alle scores onder de 1.00 afvragen hoe het komt dat de codeurs het voor sommige items oneens met elkaar waren.

Hoge overeenstemming, lage kappa

Het kan gebeuren dat er een hoog overeenstemmingspercentage is, maar toch een lage kappa-score. De oorzaak daarvan ligt vaak bij de verdeling van de data. Als we bijvoorbeeld kijken naar het coderen van sarcasme versus geen sarcasme, dan kan het zo zijn dat er relatief weinig sarcastische uitspraken in de data voorkomen. Als dat het geval is, en de codeurs zijn het er niet altijd over eens of een sarcastische uitspraak wel echt zo bedoeld is, dan heeft onenigheid tussen de codeurs een verschillend effect op de genoemde statistieken:

- Onenigheid over hoogfrequente categorieën zorgt slechts voor een lichte afname van het overeenstemmingspercentage. Het gaat immers maar over een klein aantal, en het overeenstemmingspercentage kijkt naar het grotere plaatje.
- Onenigheid over laagfrequente categorieën zorgt voor een sterke afname van de kappa-score. Immers: het gaat erom dat codeurs de twee categorieën goed van elkaar kunnen onderscheiden, en het herkennen van echte gevallen van sarcasme gaat misschien niet vaak fout in absolute getallen, maar wel relatief ten opzichte van de totale hoeveelheid sarcastische uitspraken.

Tekortkomingen van Cohen's kappa

Hoewel Cohen's kappa een stap in de goede richting is, zijn er meerdere tekortkomingen van de kappa-score:

- Cohen's Kappa is alleen maar te gebruiken voor twee codeurs. Wanneer je meerdere codeurs hebt kun je Kappa wel gebruiken om de codeurs onderling te vergelijken, maar niet om een algemene betrouwbaarheidsscore te berekenen.[4]
- Zoals gezegd is Kappa alleen maar te gebruiken voor nominale variabelen.
- Niet alle onenigheden zijn even erg (sommige categorieën liggen dichter bij elkaar dan andere; denk aan een ordinale schaal), maar Kappa houdt daar geen rekening mee.

6.2.3 KRIPPENDORFF'S ALPHA

De derde veelgebruikte maat voor intercodeursbetrouwbaarheid is Krippendorff's alpha. Ook deze maat corrigeert voor gokkans, maar in tegenstelling tot Cohen's kappa kan Krippendorff's alpha gebruikt worden met meerdere codeurs, en ook bij verschillende soorten variabelen: je kunt alpha niet alleen berekenen voor nominale data, maar ook voor ordinale, interval, en ratiovariabelen. Bovendien houdt Krippendorff's alpha ook rekening met de grootte van de verschillen wanneer codeurs bijvoorbeeld verschillende scores geven op een Likert-schaal. Net als bij Cohen's kappa vallen de mogelijke alpha-scores tussen -1 (tegenovergestelde mening) en +1 (exact dezelfde mening). Gezien de vele voordelen van Krippendorff's alpha, is dit voor veel projecten waarschijnlijk de meest geschikte maat voor intercodeursbetrouwbaarheid.

Krippendorff's alpha berekenen

Krippendorff's alpha is iets lastiger te berekenen dan Cohen's kappa, waardoor Krippendorff's alpha wat minder wijdverspreid is dan Cohen's Kappa. Inmiddels wordt Krippendorff's alpha wel ondersteund in de meeste programma's: als je Krippendorff's alpha wil berekenen, kan dat in:

- SPSS: aan de hand van de KALPHA macro (Hayes & Krippendorff (2007); beschikbaar via de website van Andrew Hayes.)
- JASP: via de Reliability-module, onder het kopje "rater agreement"
- JAMOVI: via de Seolmatrix-module, vanaf Jamovi versie 2.4.8.[5]

4 Overigens kan Fleiss' kappa wel gebruikt worden om een algemene score te berekenen voor meerdere codeurs.

5 Op het moment van schrijven moet je niet de "solid" versie van Jamovi downloaden (2.3.28, die wordt aanbevolen voor de meeste gebruikers), maar de "current" versie (2.4.14) met de laatste features.

Belangrijk: Krippendorff's alpha is niet hetzelfde als Cronbach's alpha. Je kunt Cronbach's alpha dus niet gebruiken om een betrouwbaarheidsscore voor een inhoudsanalyse te berekenen.[6]

6.3 WELKE SCORE KIES JE?

Welke betrouwbaarheidsscore je kiest, hangt af van twee factoren:

- **Algemene voorwaarden:** Met nominale data kun je alle kanten op, maar met interval-scores ben je aangewezen op Krippendorff's alpha. Hetzelfde geldt voor het aantal codeurs: met twee codeurs zijn er verschillende opties, maar met meer dan twee codeurs worden de mogelijkheden beperkt.
- **Conventie**: Wat doen andere onderzoekers? Afhankelijk van het vakgebied waarin je werkt, zijn er verschillende voorkeuren voor verschillende betrouwbaarheidsscores. Als alle literatuur in jouw deelgebied gebruikmaakt van Cohen's Kappa, dan is het een goed idee om in ieder geval ook een kappa-score te berekenen.

Vaak is het een goed idee om Kappa of Alpha te presenteren met daarnaast altijd de procentuele overeenkomst tussen de codeurs. Op deze manier geef je meer inzicht in je resultaten dan wanneer je alleen de scores voor Cohen's Kappa of Krippendorff's Alpha rapporteert.

6.4 FACTOREN DIE DE BETROUWBAARHEID VAN JE STUDIE BEÏNVLOEDEN

Intercodeursbetrouwbaarheidsscores reduceren de betrouwbaarheid van je studie tot één getal. Als je meerdere variabelen codeert, kun je ook nog een betrouwbaarheidsscore per variabele berekenen. Maar dat blijft een beperkte indicatie van de betrouwbaarheid van je studie. Er zijn namelijk veel verschillende factoren die de betrouwbaarheid van je studie beïnvloeden. Deels zijn dit keuzes die je zelf in de hand hebt, en deels liggen die factoren besloten in het domein dat je bestudeert.

6.4.1 VERSCHILLENDE FACTOREN

Van Enschot et al. (2024) beschrijven drie verschillende factoren die de betrouwbaarheid van inhoudsanalyses beïnvloeden. Deze factoren worden hieronder omschreven.

6 Lombard et al. (2002) leggen uit dat Cronbach's alpha, net als andere correlatie-scores, kijken naar covariatie (als variabele A verandert, verandert variabele B dan ook?). Maar betrouwbaarheid gaat er juist om dat je exact dezelfde categorieën codeert. Om dat punt nog verder uit te leggen: het is mogelijk om een perfecte correlatie te hebben tussen de resultaten van twee codeurs terwijl ze nooit hetzelfde antwoorden. Bijvoorbeeld met een ordinale variabele, A: [1,2,3,4], B: [2,3,4,5].

Mate van controle

In hoofdstuk 3 heb je gezien dat er verschillende manieren zijn om data te verzamelen. Het grootste deel van dit boek is gericht op het verzamelen en coderen van boodschappen die 'in het wild' voorkomen, maar we hebben bijvoorbeeld ook besproken hoe je data zou kunnen verzamelen via een experiment. In de ervaring van Van Enschot et al. (2024) is het vaak eenvoudiger om experimenteel verzamelde data te coderen. Zij geven de studie van Arts et al. (2011) als voorbeeld. In deze studie schreven participanten korte zinnen die verwezen naar objecten die op een afbeelding te zien waren. In zo'n gecontroleerde setting heb je volledige kennis van de situatie, waardoor je precies weet wat de participanten bedoelen. Natuurlijk voorkomende data is minder voorspelbaar, en bevat vaak onverwachte elementen waar je bij het ontwikkelen van het codeboek misschien nog niet over had nagedacht.[7]

Oorsprong van de categorieën

In hoofdstuk 5 hebben we twee manieren besproken om een codeboek op te stellen: (1) een bestaand codeboek gebruiken en (waar nodig) aanpassen aan jouw onderzoek, of (2) zelf een codeboek ontwikkelen. Wanneer je een bestaand codeboek gebruikt dat gebaseerd is op eerdere literatuur, dan is er een duidelijk kader om de data te interpreteren. Wanneer je zelf een codeboek ontwikkelt, dan is er nog veel werk nodig om te komen tot goedomlijnde definities. Van Enschot et al. (2024) stellen dat we lage intercodeursbetrouwbaarheidsscores kunnen vermijden als we alleen variabelen zouden bestuderen "met dusdanig goed gedefinieerde categorieën dat er nauwelijks interpretatieverschillen kunnen ontstaan" (eigen vertaling). Maar binnen de communicatiewetenschap zijn er veel belangrijke variabelen waarvoor er nog geen doorontwikkeld codeboek bestaat. Dat betekent dat we soms niet anders kunnen dan zelf een codeboek te ontwikkelen. Daarmee sta je dan aan het begin van een proces dat kan leiden tot goedomlijnde categorieën die later betrouwbaar te coderen zijn.

De aard van de categorieën

Wanneer je een variabele codeert, kun je kiezen uit twee of meer categorieën. Van Enschot et al. (2024) beschrijven de situatie als een spectrum tussen 'gesloten' en 'open' variabelen. Gesloten variabelen zijn gedefinieerd in termen van twee tegenovergestelde categorieën (*ja/nee, aanwezig/afwezig*). Zulke variabelen zijn relatief eenvoudig om te coderen. Open variabelen zijn lastiger te coderen. Een extreem voorbeeld is het coderen van kleuren: er zijn oneindig veel manieren om kleuren te omschrijven, waardoor het lastig is voor twee codeurs om zonder overleg tot overeenstemming te komen over de juiste term.[8] Daarnaast is het ook

7 Bij dagboekstudies is de vraag wat er precies verzameld wordt. In de studie van Qutteina et al. (2019) verzamelen participanten bijvoorbeeld berichten van sociale media. Dat is dus ook natuurlijk voorkomende data, met dezelfde uitdagingen als data die je zelf verzamelt.

8 Van Enschot et al. (2024) noemen hier zelf het voorbeeld van taalintensivering.

lastig wanneer de gebruikte categorieën elkaar niet uitsluiten (Van Enschot et al. (2024) noemen hier het voorbeeld van *letterlijk* versus *figuurlijk* taalgebruik).

6.4.2 MANIEREN OM INTERCODEURSBETROUWBAARHEID TE VERBETEREN

Van Enschot et al. (2024) schetsen meerdere manieren om de intercodeursbetrouwbaarheid te verbeteren. Deze manieren kunnen we opsplitsen in twee groepen: aanpassing van het codeboek, en aanpassing van het codeerproces. Deze twee groepen worden hieronder besproken.

Codeboek

Zorg dat categorieën elkaar uitsluiten

Intercodeursbetrouwbaarheid is het hoogst wanneer er weinig onzekerheid is bij de codeurs. De twijfel neemt toe wanneer de categorieën binnen een variabele allemaal op elkaar lijken, of zelfs overlappen.

Verklein het aantal categorieën

Het aantal categorieën binnen een variabele heeft een negatief effect op de intercodeursbetrouwbaarheid; met meer categorieën is de kans groot dat de gemeten betrouwbaarheid afneemt. Dat heeft twee redenen:

1) Met meer categorieën is de kans groter dat codeurs fouten maken. Ze halen bijvoorbeeld categorieën door elkaar, of zien ze over het hoofd.
2) Wanneer je alle data heel fijnmazig categoriseert, dan is de gemiddelde frequentie per groep lager dan wanneer je een kleinere verzameling van ruimere categorieën hanteert. Dat maakt het lastiger om verschillen tussen groepen statistisch te analyseren. Daarnaast levert alle onenigheid tussen codeurs ook grotere afwijkingen op in de intercodeursbetrouwbaarheidsscore.

Door een kleinere groep categorieën aan te houden, maak je het eenvoudiger om de variabele te coderen, en heb je minder last van lage frequenties bij de statistische analyse van je data. Dus: vraag je af of alle categorieën strikt noodzakelijk zijn. Natuurlijk levert dit wel minder informatie op dan een fijnmazigere analyse, dus denk hier goed over na.

Codeerproces

Zorg voor duidelijke instructies

Van Enschot et al. (2024) beargumenteren dat de instructies die je geeft aan de codeurs een belangrijke rol spelen in hun uiteindelijke gedrag. Je kunt een uitgebreid codeboek maken en de codeurs daarbij instrueren hoe ze te werk moeten gaan, maar uiteindelijk loop je

daarbij het risico dat de codeurs de data allemaal op hun eigen manier gaan coderen. Om dat te voorkomen, zien Van Enschot et al. (2024) stroomdiagrammen (*flowcharts*) als een veelbelovende optie. Een stroomdiagram beschrijft stap-voor-stap de hele beslissingsprocedure om iedere observatie op de goede manier te kunnen categoriseren. Voorbeelden die geciteerd worden door Van Enschot et al. (2024) zijn: Burgers et al. (2011), Burgers et al. (2016), en Scholman et al. (2016).

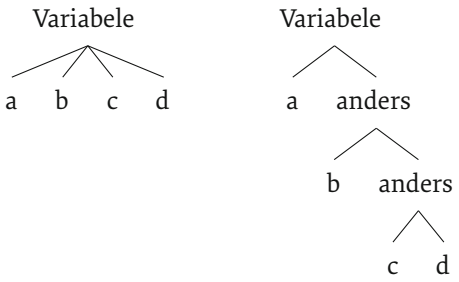
Zorg voor een goede trainingsprocedure

In navolging van Bayerl & Paul (2011), Neuendorf (2017), en vele anderen, benadrukken Van Enschot et al. (2024) het belang van training. Hoe meer tijd je besteedt aan het trainen van de codeurs, des te groter is de kans dat ze de data op dezelfde manier zullen coderen. In Hoofdstuk 5 werd ook al aandacht besteed aan de training van codeurs, onder de noemer 'uitproberen.' Voor het onderzoeksproject behorend bij dit boek is het grote voordeel dat je groepsgenoten hebt met wie je samen het codeboek ontwikkelt, en met wie je samen dus ook bekendheid op kunt bouwen met de literatuur en het codeerproces. Deze luxe hebben niet alle onderzoekers; soms moet je werken met mensen die verder weinig weten van het onderwerp, en die dus goed geïnstrueerd moeten worden.

In de hoeveelheid training zijn allerlei gradaties mogelijk, van het lezen van een codeboek (maar geen verdere oefening) tot jarenlange ervaring in het bestuderen en analyseren van berichten in een bepaald domein. Zoals gezegd zorgt training voor een hogere intercodeursbetrouwbaarheid, maar Van Enschot et al. (2024) signaleren ook een valkuil: soms kan een uitgebreide training wel zorgen voor een hoge betrouwbaarheid, maar komt die betrouwbaarheid eigenlijk doordat de codeurs een 'trucje' hebben geleerd om op basis van oppervlakkige eigenschappen van de data dezelfde categorieën toe te kennen aan de observaties. Op dat moment gaat de hoge betrouwbaarheidsscore ten koste van de externe validiteit van de studie.

Werk stapsgewijs

(hierarchische data of veel categorieën.) Hoe complexer een codeertaak wordt, des te uitdagender wordt het voor de codeurs om die taak op exact dezelfde manier uit te voeren. Van Enschot et al. (2024) richten hun pijlen nogmaals op het aantal categorieën waaruit codeurs kunnen kiezen. Zoals gezegd is het makkelijker om een binaire variabele (*ja/nee, aanwezig/afwezig*) te coderen dan een variabele met ontzettend veel verschillende categorieën om uit te kiezen. Het advies is dan ook om dat aantal zoveel mogelijk te minimaliseren. Als dat niet lukt, dan kun je misschien het codeerproces aanpassen van een eenmalige keuze naar een reeks van binaire keuzes, zoals gevisualiseerd in Figuur 6.1. (Dit doet ook weer denken aan het idee van een stoomdiagram.)



FIGUUR 6.1 *Verschillende manieren om een variabele met vier categorieën te coderen: als eenmalige beslissing tussen vier categorieën (links), of als reeks van drie binaire keuzemomenten (rechts).*

6.4.3 MEERDERE PERSPECTIEVEN: EEN *FACT OF LIFE*

Hoeveel je ook je best doet om intercodeursbetrouwbaarheid te verhogen, uiteindelijk moeten we ook erkennen dat de wereld zich gewoon niet eenvoudig laat categoriseren. De onderliggende gedachte bij intercodeursbetrouwbaarheidsscores is dat er altijd één Ware manier is om iedere observatie te categoriseren, en die aanname is niet altijd terecht. Steeds meer onderzoekers zien onenigheid niet als ruis, maar als waardevol signaal dat iets zegt over de ambiguïteit van de data Fleisig et al. (2024). Een vroege bespreking van dit idee vinden we terug in het werk van Spooren & Degand (2010). Zij vragen zich hardop af: wat moet je dan als onderzoeker, wanneer je werkt met ambigue data? Hieronder bespreken we een deel van hun suggesties.[9]

Spooren & Degand (2010) raden in ieder geval aan om de context zoveel mogelijk mee te nemen bij het analyseren van tekstuele boodschappen. Taal is uitermate context-afhankelijk; dezelfde uitspraak kan in een andere context op een totaal andere manier worden geïnterpreteerd. Wanneer je data verzamelt is het belangrijk om hiermee rekening te houden: is het in mogelijk om het bericht te interpreteren in de oorspronkelijke context?

Naast het belang van context, bespreken Spooren & Degand (2010) ook de voor- en nadelen van enkel en dubbel coderen.

Enkel coderen

Wanneer je data volledig door één iemand laat coderen, is de data in ieder geval *consistent* gecodeerd. Dat wil zeggen dat we er (hopelijk) vanuit kunnen gaan dat de codeur steeds op dezelfde manier te werk is gegaan. Dat zorgt voor een coherente interpretatie van de data. Het probleem is dat er sprake kan zijn van een *bias* in de interpretatie van de codeur: misschien zijn bepaalde categorieën systematisch te vaak toegekend. Dit kan problematisch zijn

9 De auteurs bespreken ook nog de potentie van automatische oplossingen. Zie daarvoor het oorspronkelijke artikel.

voor de generaliseerbaarheid van de resultaten, maar dat hoeft niet altijd het geval te zijn. (Wanneer je twee groepen vergelijkt, dan is er in beide groepen sprake van diezelfde systematische bias.)

Dubbel coderen

Wanneer je data volledig dubbel codeert, heb je in ieder geval een goed beeld van de onenigheid tussen codeurs. Door alle gevallen van onenigheid te bespreken, kun je komen tot een dataset waarbij de oordelen van de codeurs het gevolg zijn van een weloverwogen proces, en waarbij je ook de oorzaak kunt bespreken van de verschillende onenigheden: *waarom* waren ze het nou oneens? Het grote nadeel van dubbel coderen is dat het erg arbeidsintensief is.

Wat in ieder geval duidelijk wordt uit de literatuur, is dat het belangrijk is om de beperkingen van je onderzoek te erkennen. Zoals Spooren & Degand (2010) zeggen: lage kappas zijn in sommige gevallen gewoon een *fact of life*. Dat betekent niet dat we ambigue data moeten vermijden, maar dat we goed moeten nadenken over de keuzes die we maken, en over de conclusies die we desondanks *wel* kunnen trekken.

6.5 WAT TE DOEN BIJ EEN LAGE BETROUWBAARHEIDSSCORE?

Wat kun je doen als je de intercodeursbetrouwbaarheid te laag is? Dat ligt eraan in welke fase van het onderzoek je zit. Na een pilotstudie kun je het codeboek bijvoorbeeld nog herzien om de instructies voor de codeurs te verbeteren, of om de taak te vereenvoudigen. Als de betrouwbaarheidsscore aan het eind van het onderzoek te laag is, dan zijn je opties beperkt (tenzij je alles opnieuw wil gaan coderen).

6.5.1 DUBBEL GECODEERDE DATA

Sommige onderzoekers kiezen ervoor om alle data dubbel te laten coderen om het codeerproces zo betrouwbaar mogelijk te laten zijn. Als alle data dubbel gecodeerd is door twee codeurs, kun je onenigheden tussen de twee codeurs oplossen via onderlinge discussies, of de data opnieuw laten beoordelen door een derde codeur. Met meer dan twee codeurs kun je er ook voor kiezen om de meerderheid te laten bepalen hoe de data gecodeerd moet worden.[10] Het voordeel van alles dubbel coderen is dat de resultaten zo betrouwbaar mogelijk zijn, maar het nadeel is dat dubbel coderen ook veel meer tijd kost.

10 Met deze benadering geeft de intercodeursbetrouwbaarheidsscore niet zozeer een indicatie van de kwaliteit van de uiteindelijke data (want na het berekenen van de scores wordt de gecodeerde data nog verbeterd), maar meer over de betrouwbaarheid van het initiële proces. Door alle onenigheden nog eens onder de loep te nemen, kun je jezelf ervan verzekeren dat de uiteindelijke data betekenisvol is.

6.5.2 WAT TE DOEN BIJ VERSCHILLEN TUSSEN CODEURS?

Ook als je niet het hele corpus dubbel gaat coderen, is aan het eind van je studie een deel van de data dubbel gecodeerd. Voor deze data zul je ook moeten bepalen wat je doet met onenigheid tussen de codeurs voordat je de resultaten van je studie kunt berekenen. Hier heb je weer een aantal opties:[11]

- Alle verschillen tussen codeurs bespreken met de hele groep. (Dit levert ook gelijk input voor de resultaten en discussie aan het eind van je studie: hoe zijn de verschillen tussen codeurs ontstaan, en (hoe) zouden die in de toekomst voorkomen kunnen worden?)
- Onenigheid tussen codeurs laten oplossen door een extra codeur die de knoop doorhakt. Dit gebeurt bijvoorbeeld wanneer een hoofdonderzoeker geholpen wordt door codeurs die minder achtergrondkennis hebben. De hoofdonderzoeker kan dan aan de hand van zijn of haar expertise de laatste knopen doorhakken.
- Bij systematische verschillen:
  - Specifieke delen van de data opnieuw coderen, bijvoorbeeld omdat een codeur altijd bij dezelfde soort items de verkeerde keuze maakt.[12]
  - Twee of meerdere categorieën samenvoegen, omdat het onderscheid bij nader inzien te fijnmazig was om op een betrouwbare manier te coderen.
- Categorieën weglaten als je merkt dat ze niet betrouwbaar te coderen zijn, en eventuele conclusies die je trekt op basis van de data daarmee ook onbetrouwbaar zouden zijn.
- (Gevorderd) Een *multiverse* analysis uitvoeren waarbij je niet één maar meerdere statistische analyses uitvoert om te laten zien welke gevolgen bepaalde methodologische keuzes hebben op de uitkomst van je studie (Steegen et al., 2016).

6.5.3 AMBIGUÏTEIT, CODEREN, EN MENINGSVERSCHILLEN

In de discussie over intercodeursbetrouwbaarheid speelt nog een fundamentele vraag: is er bij het coderen altijd maar één mogelijk antwoord, of zijn er ook situaties waarbij meerdere antwoorden verdedigbaar zijn? De vraag stellen is hem beantwoorden; de wereld is complex, en het is niet mogelijk om alles in regels te vatten. Ambiguïteit is een *fact of life*. Bij ambigue codeertaken heb je een aantal opties, met elk hun eigen voors en tegens:

---

11 In theorie kun je ook alleen de data meenemen waarover codeurs het eens waren. Dit gebeurt bijvoorbeeld in de analyse van Liebrecht (2015). Deze optie is niet altijd toepasbaar, omdat je hiermee de generaliseerbaarheid van je studie beperkt. Door een deel van de data weg te laten, introduceer je een *bias* in de data. (Er wordt geen willekeurige data weggelaten, maar data met bepaalde eigenschappen die leiden tot onenigheid tussen codeurs.)

12 Het is verstandig om bij deze optie na te denken over een manier om fouten automatisch te detecteren. Bijvoorbeeld door middel van een formule in Excel die bepaalde woorden herkent. (Zie hiervoor Bijlage G.)

- Alle data door één codeur laten beoordelen. Het is dan misschien zo dat de codeur keuzes waar iemand anders misschien voor een andere optie zou kiezen, maar omdat dezelfde codeur alle data beoordeelt is de data wel consequent gecodeerd.
- Alle ambigue gevallen laten markeren door de codeurs, en apart analyseren.

6.6 STEEKPROEFGROOTTE

Zoals elke statistiek is de betrouwbaarheid van intercodeursbetrouwbaarheidsscores afhankelijk van de steekproefgrootte. Stel je voor dat je met twee codeurs 10 items gaat dubbelcoderen, dan is de uiteindelijke betrouwbaarheidsscore erg afhankelijk van de items die toevallig zijn opgenomen in de steekproef. Ieder verschil tussen de codeurs heeft dan een grote invloed op de uiteindelijke score. Immers: met één foutje is dan gelijk 10% van de dubbel gecodeerde data verkeerd gecodeerd.

De vraag hoe groot de steekproef dan moet zijn is lastig te beantwoorden, omdat dit afhangt van verschillende factoren. Onder andere:

- **Hoe gevarieerd is het corpus dat je wil onderzoeken?** Als er veel variatie zit in de data die je wil coderen, en die variatie is mogelijk van invloed op het codeerproces, dan heb je een grotere steekproef nodig zodat er genoeg verschillende voorbeelden zijn om de betrouwbaarheid van het codeerproces te toetsen.
- **Hoe vaak komen de verschillende categorieën voor die je wil coderen?** Als je voor iedere categorie wil weten hoe betrouwbaar ze geïdentificeerd kunnen worden, moeten alle categorieën ook voorkomen in de steekproef die je dubbel codeert. Sterker nog: ze moeten meerdere keren voorkomen als je wil weten hoe betrouwbaar de identificatie van die categorieën is.
- **Hoe betrouwbaar moet de betrouwbaarheidsscore zijn?** Met een grotere steekproef kun je nauwkeuriger bepalen hoe betrouwbaar de codeurs zijn, en weet je zekerder dat de betrouwbaarheidsscore niet door toeval bepaald is.
- **Hoeveel tijd heb je om de data te coderen?** Praktische haalbaarheid is eigenlijk geen zuivere reden om voor een bepaalde steekproefgrootte te kiezen; als je geen tijd hebt voor een bepaald soort onderzoek, had je misschien maar een ander onderwerp moeten kiezen. Maar soms kun je niet anders.

Lombard et al. (2010) geven geen exacte methode om de steekproefgrootte te bepalen, maar zeggen dat de steekproef voor het berekenen van de intercodeursbetrouwbaarheid niet kleiner moet zijn dan 50 items (het absolute minimum) of 10% van de data. Daarnaast geven ze ook een bovengrens: je hebt zelden meer dan 300 items nodig. Andere onderzoekers (Krippendorff, 2018; Lacy & Riffe, 1996) geven statistische formules om de steekproefgrootte te bepalen. Een discussie van deze formules gaat echter te ver voor dit werkboek.

6.7 BETROUWBAARHEID VISUALISEREN

Stel je voor dat er een codeertaak is waarbij de codeurs foto's van dieren moeten categoriseren als hond, wolf, kat, of lynx. Je kunt de overeenkomsten en verschillen tussen twee codeurs visualiseren aan de hand van een tabel waarin de kolommen overeenkomen met de waarden die zijn toegekend door codeur A, en waarin de rijen overeenkomen met de waarden die zijn toegekend door codeur B. Een visualisatie zoals deze geeft precies aan welke categorieën met elkaar verward worden en welke niet. In Excel of in gespecialiseerde software kun je zo'n tabel ook omzetten in een *heatmap table* waarbij alle cellen gekleurd zijn en de intensiteit van de kleuren in de tabel aangeeft hoe vaak iedere combinatie van categorieën voorkomt. Tabel 6.3 geeft een voorbeeld.[13]

		Codeur A			
		Hond	Wolf	Kat	Lynx
Codeur B	Hond	50	3	0	0
	Wolf	5	35	0	0
	Kat	0	0	52	7
	Lynx	0	0	6	29

TABEL 6.3: *Frequentietabel met overeenkomsten tussen twee codeurs: Codeur A en Codeur B.*

We kunnen Tabel 6.3 lezen als volgt: algemeen zien we een hoge mate van overeenstemming tussen Codeurs A en B, getuige het feit dat er hoge waarden in de diagonaal staan en lage waarden daarbuiten. We zien ook dat er onenigheid is bij het onderscheid tussen honden en wolven en het onderscheid tussen katten en lynxen (want daar zien we getallen groter dan 0), maar dat katten en honden niet door elkaar gehaald worden (want in die cellen staat een 0).

6.8 WAT TE DOEN MET DUBBEL GECODEERDE DATA?

Wanneer je de intercodeursbetrouwbaarheid berekend hebt, ben je bijna klaar om de resultaten van je studie te gaan analyseren. Voordat je met je favoriete statistiekprogramma aan de slag kunt gaan, moet je eerst bedenken wat je gaat doen met de dubbel gecodeerde data. Voor de meeste statistische analyses kun je per observatie (in ons geval meestal: per boodschap) en per categorie (bijvoorbeeld: aanwezigheid van sarcasme) maar één waarde hebben

13 In Engelstalige versie van Excel kun je deze optie vinden onder "Conditional Formatting" en dan de optie "Colour scales". In de Nederlandse versie heet het menu "voorwaardelijke opmaak" en de optie "kleurenschalen."

die toegekend wordt aan die observatie. Je zult dan de dubbel gecodeerde data op de een of andere manier moeten verwerken om aan die voorwaarde te voldoen.

Optie 1: één codeur selecteren
De eerste optie is om één van de codeurs te selecteren en alleen diens antwoorden te gebruiken. De andere antwoorden negeer je gewoon bij het uitvoeren van de statistische analyse. Deze optie zou je bijvoorbeeld kunnen gebruiken als één van de codeurs alle data heeft gecodeerd, en de tweede codeur alleen een subset van de data heeft bekeken om de intercodeursbetrouwbaarheidsscores te kunnen berekenen. Het voordeel hiervan is dat alle data op precies dezelfde manier is gecodeerd. Het nadeel is dat je mogelijke foutjes van de hoofdcodeur niet corrigeert.

Optie 2: samenvoegen van de dubbel gecodeerde data
De tweede optie is om de dubbel gecodeerde data samen te voegen. Alle items waarover de codeurs het eens zijn kun je gewoon bewaren zonder er verder naar te kijken. Voor de overige items zijn er drie verschillende opties:

1) De items waarover de codeurs het oneens zijn, kun je onderling bespreken om tot een consensus te komen. Het voordeel is dat dit gelijk nieuwe inzichten oplevert over verschillende interpretaties van de data, en dat het overleg de validiteit van de data verhoogt. Het nadeel is dat zo'n overleg tijdrovend kan zijn.
2) Bij een grotere groep codeurs kun je de meerderheid laten beslissen. Het voordeel is dat dit vaak snel gaat, en bij een oneven aantal codeurs altijd succesvol is. Het nadeel is dat de meerderheid niet altijd gelijk heeft.
3) Met twee codeurs kun je ook een derde codeur inschakelen om de knoop door te hakken. Dit werkt vooral goed als de derde codeur veel achtergrondkennis heeft om de juiste beslissing te nemen.

6.9 STATISTIEK

Wanneer alle data gecodeerd is, en je de intercodeursbetrouwbaarheid berekend hebt, is het tijd om de relevante statistieken voor je studie te gaan berekenen.

6.9.1 DESCRIPTIEVE STATISTIEK

Allereerst is het belangrijk om te beschrijven hoe de gecodeerde data eruit ziet. Dat kun je vaak het beste doen in een frequentietabel waarin je de groepen die je vergelijkt naast elkaar zet, met de categorieën die je gecodeerd hebt onder elkaar. Zo zien lezers in één oogopslag hoe vaak de verschillende categorieën voorkwamen in de groepen die je bestudeerd hebt. Tabel 6.4 geeft een voorbeeld.

Categorie	Groep 1	Groep 2	Totaal
A	8	27	35
B	19	3	22
C	6	7	13

TABEL 6.4 *Frequentietabel waarbij het aantal voorkomens van categorieën A, B, en C tussen groepen 1 en 2 vergeleken wordt.*

Bij descriptieve statistiek is het erg belangrijk om na te denken over de getallen die je rapporteert. Je kunt bijvoorbeeld kiezen tussen:

- Het rapporteren van **absolute aantallen** versus **percentages**: met absolute getallen zien lezers exact wat je hebt gevonden, maar absolute getallen kunnen misleidend zijn als je twee groepen die je vergelijkt niet even groot zijn. In dat geval werken percentages soms beter. Je kunt er natuurlijk ook voor kiezen om allebei te rapporteren. Let bij percentages wel op dat het duidelijk is voor de lezers wat de referentie is voor de percentages die je geeft. (Bijvoorbeeld: 10 procent *waarvan*?)
- Het rapporteren van het **gemiddelde** versus de **mediaan**: soms wil je ook samenvatten hoe vaak iets 'normaal gesproken' voorkomt per item. Bijvoorbeeld: hoe vaak zie je mensen elkaar onderbreken tijdens talkshow-afleveringen? In zulke gevallen is het gemiddelde vaak een goede keuze. Maar soms wordt het gemiddelde sterk beïnvloed door extreme waarden (*outliers*). Wanneer dat gebeurt kan het verstandiger zijn om de mediaan te rapporteren (het getal dat in het midden ligt als je alle waarden op grootte sorteert).

Afhankelijk van de data die je rapporteert kan het verstandig zijn om naast (of zelfs in plaats van) een frequentietabel ook andere soorten visualisaties te gebruiken. Daarover lees je later meer in het hoofdstuk over rapportage.

6.9.2 TOETSENDE STATISTIEK

Deze cursus veronderstelt basiskennis van statistiek, dus we zullen niet al te diep ingaan op de verschillende soorten analyses die je zou kunnen uitvoeren met gecodeerde data. Voor een inhoudsanalyse is vooral van belang dat je een analyse kiest die past bij de data die je gecodeerd hebt. Tabel 6.5 kan je daarbij helpen. Let daarbij wel op: afhankelijk van hoe je de data analyseert, kan het zijn dat je een andere analyse moet kiezen.

Toets	Onafhankelijke variabele	Afhankelijke variabele	Commentaar
Chi-kwadraat	Nominaal	Nominaal	Symmetrische statistiek
Correlatie	Ordinaal, interval, ratio	Ordinaal, interval, ratio	Symmetrische statistiek
Regressie	Nominaal, interval, ratio	Interval, ratio	
t-toets	Interval, nominaal	Interval, ratio	1 factor met max. 2 niveaus
ANOVA	Interval, nominaal	Interval, ratio	1 factor (meerdere niveaus) of meerdere factoren
Mann-Whitney	Interval, nominaal	Interval, ratio	1 factor met max. 2 niveaus
Kruskal-Wallis	Interval, nominaal	Interval, ratio	1 factor (meerdere niveaus) of meerdere factoren

TABEL 6.5 *Verschillende toetsen met de soorten variabelen waarmee ze gebruikt kunnen worden. Chi-kwadraat en Correlatie zijn symmetrische toetsen die vooral kijken naar de associatie tussen twee variabelen; het maakt niet uit welke variabele je als afhankelijk/onafhankelijk invult. De Mann-Whitney en Kruskal-Wallis toetsen kunnen worden gebruikt als alternatief voor de t-toets en ANOVA wannneer de data niet normaal verdeeld zijn.*

Voorbeeld

Stel je voor dat je het commentaar op twee socialemediaplatformen hebt gecodeerd, om te kijken in hoeverre gebruikers van de platformen behulpzaam zijn. Voor het onderzoek heb je 100 verschillende posts geselecteerd op ieder platform, en van die posts heb je het commentaar gecodeerd. Een van de variabelen is het geven van constructieve feedback (binair gecodeerd als: aanwezig/afwezig). Je kunt deze data op minstens twee manieren analyseren. Bijvoorbeeld:

- **Optie 1**: Je gooit alle commentaren op twee grote hopen: al het commentaar op platform A, en al het commentaar op platform B. Wanneer je de twee platformen met elkaar vergelijkt, toets je of de algemene proportie wel/niet constructieve reacties verschilt tussen de platformen. De eenheid van analyse is hier: comments, en de variabelen die je vergelijkt zijn nominaal (wel/niet constructief). Dit zou je kunnen doen met een Chi-kwadraattoets.
- **Optie 2**: Je aggregeert het commentaar op post-niveau: voor iedere post op ieder platform bereken je het percentage constructieve reacties. Wanneer je de twee platformen met elkaar vergelijkt, toets je of het gemiddelde percentage constructief commentaar voor alle posts verschilt tussen de platformen. De eenheid van analyse is hier dus: posts, en de variabelen die je vergelijkt zijn ratio (percentage constructieve reacties). Dit zou je kunnen doen met een t-toets.

Het voordeel van de eerste optie is dat je alle commentaren gebruikt in je analyse. Door de data te aggregeren, 'verdwijnen' de individuele reacties in een percentage. Je moet bij Optie 2 vooraf goed nadenken hoeveel reacties voldoende zijn om een betrouwbaar beeld te krijgen van de reacties voor een post. Als je met 100 reacties een goed beeld krijgt, hoef je geen 1000 reacties per post te coderen; het percentage wordt dan wel iets betrouwbaarder, maar het levert misschien meer op om 100 reacties voor 10 posts te coderen, dan om 1000 reacties op 1 post te coderen. Het voordeel van de tweede optie is dat je eenvoudiger kunt zien of er uitschieters zijn: posts met bijzonder veel/weinig constructief commentaar, die de algemene verhouding beïnvloeden. Door alles op één hoop te gooien, verlies je dit soort informatie.[14]

14 Er zijn ook geavanceerdere statistische technieken waarmee je de data zou kunnen analyseren, bijvoorbeeld met een *linear mixed model*, maar een diepgaande bespreking van statistiek gaat te ver voor deze cursus.

HOOFDSTUK 7
Mensen en computers

7.1 INLEIDING

In dit hoofdstuk gaan we kijken naar verschillende manieren waarop de computer ons kan helpen bij het uitvoeren van een inhoudsanalyse. Eerst blikken we kort terug op het automatiseren van de dataverzameling, waarna we dieper ingaan op het (semi-)automatisch coderen van data. Zoals we zullen zien, is het nuttig om programmeerkennis te hebben, maar absoluut niet noodzakelijk. Door verschillende programma's op een slimme manier in te zetten kun je al veel tijd besparen. Tegelijkertijd moet je de inhoud ook niet uit het oog verliezen; computers kunnen veel, maar niet alles. Uiteindelijk blijf je zelf als onderzoeker verantwoordelijk voor de kwaliteit van de verzamelde data, en de analyse daarvan.

7.2 DATAVERZAMELING

Voordat je handmatig data verzamelt van websites, sociale media, of uit verzamelingen van digitale documenten, is het goed om je af te vragen of dat niet automatisch kan. In hoofdstuk 3 hebben we al besproken dat je automatisch data kunt verzamelen door middel van bestaande programma's die gemaakt zijn om data te verzamelen, of door zelf een oplossing te programmeren. Daarbij hebben we vooral de nadruk gelegd op de verzameling (het downloaden van informatie) en organisatie van de metadata (opslaan in een Excelbestand). Afhankelijk van de hoeveelheid werk die het kost om de data handmatig te verzamelen, kan het de moeite waard zijn om de dataverzameling te automatiseren.[1]

1 Voor programmeurs is dit soms ook een valkuil: dat iets te programmeren valt, wil niet zeggen dat je het moet programmeren. Het kan ook langer duren om een oplossing te programmeren dan om zelf

In dit hoofdstuk leggen we de nadruk op de inhoud van de data. Sommige inhoudelijke eigenschappen worden ook vaak gebruikt tijdens de dataverzameling, bijvoorbeeld:

- Als je data van sociale media wil filteren op taal, bijvoorbeeld wanneer je alleen maar geïnteresseerd bent in Nederlandstalige Tweets.
- Als je foto's wil selecteren op bepaalde inhoudelijke eigenschappen, bijvoorbeeld wanneer je alleen maar foto's wil verzamelen waarin mensen afgebeeld zijn.
- Als je geïnteresseerd bent in een bepaald onderwerp, bijvoorbeeld als je alleen maar nieuwsberichten of posts van sociale media wil verzamelen die gaan over een bepaald onderwerp.

Deze eigenschappen zijn allemaal automatisch te herkennen. Automatisering maakt het op deze manier mogelijk om bepaalde onderzoeksvragen te kunnen bestuderen die je anders helemaal niet zou kunnen onderzoeken. Momenteel zien we dat oplossingen op basis van *Artificial Intelligence* (Kunstmatige Intelligentie) steeds beter worden, en daarmee groeien ook de mogelijkheden voor ons als communicatiewetenschappers om precies die data te vinden waar we in geïnteresseerd zijn.

Hoewel computers een goede inschatting kunnen geven van de eigenschappen van verschillende soorten boodschappen, valt het te betwijfelen of dit soort eigenschappen eigenlijk wel objectief te bepalen zijn; er zijn altijd twijfelgevallen die misschien wel, misschien niet tot de selectie zouden moeten behoren. Bijvoorbeeld: op sociale media gebruiken mensen niet altijd Standaardnederlands. Er wordt straattaal gebruikt, en ook woorden en zinsneden uit andere talen. Daardoor worden berichten van sociale media niet altijd door de computer herkend als Nederlands, ook al worden ze gewoon gebruikt in gesprekken tussen Nederlanders, en begrijpen de meeste mensen wel wat er bedoeld wordt. Waar het herkennen van "Nederlandstalige berichten" op het eerste gezicht niet zo moeilijk lijkt te zijn, is het bij nader inzien een erg complexe taak.

De uitdagingen bij het selecteren van 'de juiste data' maken selectieve dataverzameling ook een vorm van inhoudsanalyse. Voor bovenstaande eigenschappen gelden dezelfde overwegingen als die we hieronder zullen beschrijven: blijf kritisch op de validiteit van je onderzoeksmethode, en controleer goed of er geen systematische *bias* optreedt in de selectie. (Bijvoorbeeld wanneer altijd dezelfde soort items ten onrechte uitgesloten wordt.) Zoals bij iedere onderzoeksmethode geldt dat we de voor- en nadelen van die methode goed tegen elkaar moeten afwegen.

even handmatig aan de slag te gaan. Je moet dus ook een inschatting maken: is dit de tijdsinvestering waard?

7.3 AUTOMATISCH CODEREN

Met de term *automatisch coderen* verwijzen we in dit hoofdstuk naar de automatische verwerking van de verschillende soorten boodschappen die je zou kunnen bestuderen via een inhoudsanalyse, waarbij de computer extra informatie toevoegt over die boodschappen. Met dit hoofdstuk willen we je een eerste intuïtie geven over de dingen die je wel of niet met een computer zou kunnen doen. Daarbij proberen we de technische details zoveel mogelijk te vermijden. Tot op zekere hoogte kunnen we verschillende soorten programma's gewoon benaderen als een soort *black box* waar je data in stopt, en resultaten uit krijgt. Door kritisch te zijn op de manier waarop die programma's zich gedragen (en met een beetje gezond verstand), kunnen we op een weloverwogen manier bepalen wanneer het gepast is om een automatische inhoudsanalyse uit te voeren.

7.3.1 EEN EERSTE INTUÏTIE

Het belangrijkste uitgangspunt voor om automatische codeermethoden is dat computers niet kunnen toveren. Als iets te goed klinkt om waar te zijn, dan is het dat vaak ook.[2] Daarbij kunnen we drie generalisaties maken:

Regel 1: Als een mens het niet kan (ook niet met veel moeite), kan een computer het ook niet.
Meestal is automatisering gewoon een manier om menselijke handelingen te versnellen. Natuurlijk kunnen computers sommige dingen wel nauwkeuriger of betrouwbaarder doen, maar de meeste variabelen die je automatisch zou kunnen coderen, kun je eigenlijk ook wel met de hand coderen. Het duurt dan alleen langer. Als iemand je zegt dat een computer iets kan wat een mens niet kan, dan is het goed om kritisch te kijken naar het beschikbare bewijs: waar is die claim op gebaseerd? En hoe generaliseerbaar is die claim eigenlijk?

Regel 2: Hoe abstracter (of latenter) de analyse, des te moeilijker wordt het voor de computer.
Mensen zijn altijd op zoek naar betekenis. Als we een tekst of een afbeelding zien, dan proberen we automatisch te interpreteren wat er aan de hand is, en wat er met die boodschap

2 Een veelvoorkomend probleem is *wensdenken* (Engels: *wishful thinking*). Vanuit het idee dat het fantastisch zou zijn als een computer een bepaalde taak zou kunnen uitvoeren, wordt er dan gewerkt aan een manier om de computer die taak uit te laten voeren. Maar daarbij wordt een belangrijke vraag overgeslagen: is het überhaupt mogelijk om die taak uit te voeren op basis van de beschikbare informatie? Bekende historische voorbeelden zijn Fysionomie en Frenologie: de ideeën dat je aan iemands gelaat of schedel kunt herkennen wat voor persoon het is. Bij Fysiognomie is het idee dat jouw uiterlijk (bijvoorbeeld je kaaklijn of de vorm van je neus) iets zegt over je persoonlijkheid, en bij Frenologie is het idee dat vaardigheden en karaktertrekken aan de vorm van de schedel af te lezen zijn – daar komt ook het woord *talenknobbel* vandaan. Hoewel beide ideeën als pseudowetenschap worden beschouwd, zien we ze helaas ook terug in de AI-wereld, waar onderzoekers bijvoorbeeld hebben geprobeerd om te voorspellen op basis van foto's of iemand crimineel is of niet (Arcas et al., 2023). Dat is fundamenteel onmogelijk, want criminaliteit wordt bepaald door *sociale* factoren.

bedoeld wordt. Die bedoeling is meestal niet letterlijk aanwezig in een tekst, en valt meestal niet aan te wijzen in een afbeelding; het is de invulling die wij eraan geven vanuit onze ervaring en achtergrondkennis. In de context van inhoudsanalyses wordt er vaak gesproken over het onderscheid tussen manifeste (direct waarneembare) informatie, en latente (onderliggende, meer verborgen) informatie. Computers zijn erg goed in het herkennen van de eerste categorie, maar de tweede categorie is op zijn minst uitdagend. Je kunt een computer dus goed oppervlakkige fenomenen laten coderen, zoals het meten van de gemiddelde zinslengte of het aantal woorden in een tekst. Ook te doen, met een foutje hier en daar, zijn zaken als het herkennen van taalfouten, het ontleden van zinnen, en het herkennen van verschillende entiteiten in een tekst. (Hierover later meer.) Computers hebben veel meer moeite met zaken zoals humor en sarcasme,[3] omdat het hierbij vaak gaat om niet-letterlijke betekenis die geïnterpreteerd moet worden.

Regel 3: Technologie is nooit objectief.
Een veelvoorkomende misvatting is dat computers objectief zijn. Hoewel een computer altijd de gegeven instructies zal uitvoeren, ongeacht de gebruiker, moeten we niet denken dat het gebruiken van computers daarmee ook neutraal is. Dat heeft ten minste twee redenen:

1) Om een computerprogramma te schrijven, doen programmeurs altijd aannames over hoe de wereld in elkaar zit. Op die manier zit er altijd een bepaald perspectief op de wereld ingebakken in de programma's die je gebruikt.[4]
2) Door te kiezen voor een automatische oplossing, beperk je het aantal mogelijke oplossingen voor de vraag die je wil beantwoorden. Immers: je kunt alleen maar werken met variabelen die kwantificeerbaar en berekenbaar zijn, en het is vaak niet of maar beperkt mogelijk om uitzonderingen te maken voor speciale gevallen.

Dat betekent niet dat je geen computers mag gebruiken, maar het betekent wel dat je voorzichtig moet zijn met labels zoals "objectiviteit" en je moet blijven afvragen welke aannames verborgen zitten achter de keuze om data automatisch te analyseren.

3 Meer algemeen gaat het om niet-geconventionaliseerde pragmatische fenomenen.
4 Dat perspectief kan nog verder versterkt worden wanneer programma's gebruik maken van *Machine Learning*. De computer leert dan zelf een taak uit te voeren aan de hand van voorbeelddata, en door die data kunnen bepaalde vooroordelen in het programma terechtkomen.

7.3.2 VEELGEBRUIKTE OPLOSSINGEN

Hieronder bespreken we een aantal veelgebruikte oplossingen om teksten, afbeeldingen, en geluidsopnames automatisch te coderen. Dit is geen uitputtende lijst (het vakgebied ontwikkelt zich veel te snel om een volledig overzicht te kunnen geven), maar hopelijk zetten onderstaande benaderingen je aan het denken over manieren om jouw data automatisch te kunnen coderen.

Algemene (niet-inhoudelijke) statistieken

Een van de eenvoudigste automatische analyses is het geven van verschillende statistieken over de vorm van de tekst. Bijvoorbeeld de minimale, maximale, en gemiddelde zinslengte (en standaarddeviatie daarvan) in een tekst, of de meestgebruikte woorden in een tekst. Het is vaak wel de moeite waard om zulke statistieken te verkennen omdat ze eenvoudig te berekenen zijn en direct inzicht bieden in de verdeling van je data. Je kunt overwegen om deze statistieken te noemen in je methodesectie, bij de beschrijving van het corpus.

Leesbaarheidsscores

Naast de algemene statistieken zijn er ook meer ingewikkelde analyses mogelijk over de leesbaarheid van teksten. In het Engelse taalgebied is de Flesch-Kincaid score hier bijvoorbeeld veel voor gebruikt (Kincaid et al., 1975). Deze score is gebaseerd op het aantal woorden per zin (langere zinnen zijn moeilijker), en het aantal lettergrepen per woord (langere woorden zijn moeilijker). Voor het Nederlands is het programma T-Scan ontwikkeld om de complexiteit van teksten te meten (Pander Maat & Dekker, 2016).[5] Recent is er ook een nieuwe leesbaarheidsscore verschenen voor het Nederlands: *Leesbaarheidsinstrument voor Nederlandse Teksten* (LiNT; Maat et al., 2023). Deze score is gebaseerd op de woordfrequenties (hoe zeldzaam of veelgebruikt zijn de woorden in een tekst?), concreetheidsscores (hoe abstract/concreet zijn de gebruikte woorden?) en twee andere variabelen die de complexiteit van de zinsbouw meten.

Woordenlijsten

De klassieke oplossing om automatisch teksten te coderen is om te werken met woordenlijsten die geassocieerd zijn met een bepaalde categorie. (Soms wordt dit in het Engels een *dictionary-based approach* genoemd.) Bijvoorbeeld: als je verontschuldigingen zou willen coderen in online discussies, dan zou je relevante berichten automatisch kunnen herkennen aan het gebruik van woorden als *sorry,* of *excuses* en frases zoals *onze verontschuldigingen.* (Bijlage G beschrijft hoe je dit moet doen in Excel en andere spreadsheetprogramma's.) Het grote voor-

5 Zie Kraf et al. (2011) voor een toegankelijk overzicht van eerdere methoden om leesbaarheid van Nederlandse teksten te meten.

deel van woordenlijsten is dat het codeerproces snel, transparant en efficiënt verloopt. Maar er kleven natuurlijk ook nadelen aan het gebruik van woordenlijsten:

- Woordenlijsten hebben altijd maar een beperkte dekking. Bijvoorbeeld: als iemand zijn excuses aanbiedt op een manier die niet in je woordenlijst staat, dan wordt dat niet herkend als verontschuldiging. Kleinnijenhuis & Atteveldt (2006) noemen twee algemene uitdagingen voor computers die werken met woordenlijsten: *synomiemen* en *hyponiemen*. Zelf als een bepaald *kenobject* in principe wel in de lijst te vinden is, kan het zijn dat er synoniemen (andere woorden met dezelfde betekenis) ontbreken, of dat er een meer algemene term (hyponiem) gebruikt wordt.
- Betekenis is context-afhankelijk. Bijvoorbeeld: vaak gebruiken mensen het woord 'sorry' om zich te verontschuldigen, maar soms wordt het woord ook sarcastisch gebruikt. In dat geval is er geen sprake van een oprechte verontschuldiging, ondanks dat het woord 'sorry' in het bericht voorkomt. Een ander probleem dat door Kleinnijenhuis & Atteveldt (2006) genoemd wordt is het voorkomen van *homoniemen*: woorden die er hetzelfde uitzien maar iets anders betekenen (denk aan het woord *bank* dat gebruikt wordt voor een meubel maar ook voor een financiële instelling).

Andere beperkingen worden besproken door Mohammad (2020), specifiek voor sentiment-lexicons, maar veel argumenten gelden voor alle soorten woordenlijsten.

LIWC

LIWC is een programma met gevalideerde woordenlijsten voor verschillende psychologische dimensies. Het idee is niet zozeer dat het voorkomen van een woord direct gekoppeld is aan een categorie, maar dat de relatieve frequentie waarmee bepaalde woorden gebruikt worden iets zegt over de mentale toestand van de auteur van een tekst (Tausczik & Pennebaker, 2009). Er is inmiddels een indrukwekkende serie aan studies gepubliceerd waarin zulke verbanden aangetoond worden. (Tausczik & Pennebaker (2009) geven hiervan een impressie.)

Zoals de auteurs zelf ook aangeven is de techniek die LIWC gebruikt een grove benadering. Natuurlijk is er wel een correlatie tussen je mentale toestand en je taalgebruik, maar het is niet zo dat we op individueel niveau heel accuraat kunnen meten hoe iemand zich precies voelt aan de hand van zijn of haar taalgebruik; de gemeten effecten worden geobserveerd op groepsniveau. Deels komt dit weer door de context-afhankelijkheid van talige uitingen (die zorgt voor 'ruis' in de metingen), maar daarnaast kun je je afvragen in hoeverre informatie over onze diepste gevoelens echt te vinden is in ons taalgebruik (zie vuistregels 1 en 2 die hierboven omschreven zijn).[6] Dat doet niets af aan de eerdergenoemde studies, maar betekent

6 Hiernaast speelt er ook nog een ethisch vraagstuk: als je met een programma precies zou kunnen meten hoe iemand zich voelt aan de hand van geschreven teksten, is het dan oké om die teksten zonder toestemming te analyseren? Of komen we dan teveel aan de persoonlijke levenssfeer van de auteur?

wel dat je voorzichtig moet zijn met het trekken van conclusies op basis van woordfrequenties.

Sentiment-analyse

Er zijn verschillende programma's die ontwikkeld zijn om te meten of een tekst positief of negatief is. Dat noemen we sentiment-analyse. Oorspronkelijk werkten deze programma's aan de hand van woordenlijsten, waarbij gewoon gemeten werd hoeveel positieve of negatieve termen er gebruikt werden in een bericht. Door de jaren heen zijn deze programma's steeds geavanceerder geworden, vooral door het gebruik van *Machine Learning*. Automatische sentiment-analyse werkt nu aan de hand van programma's die zijn *getraind* op basis van grote verzamelingen voorbeelddata: zinnen en teksten waarvan codeurs hebben aangegeven of de tekst positief of negatief is.[7] Met deze voorbeelddata leert de computer om te voorspellen welk sentiment geassocieerd is met een bepaalde tekst, en in welke mate dat sentiment dan aanwezig is.

Hoewel de technologie steeds beter wordt, zijn er nog steeds beperkingen in het gebruik van sentiment-analyse. Hier bespreken we er twee:

1) Computers zijn minder goed in het begrijpen van niet-letterlijke betekenis. Humor en sarcasme worden dus niet altijd begrepen, omdat de juiste interpretatie sterk afhankelijk is van de context.
2) Wanneer je weet hoe positief of negatief mensen zijn in hun tekst, weet je nog niet *waarover* ze zich positief of negatief uitlaten. (Dit probleem wordt in het Engels *stance detection* genoemd.)

Automatische beeldherkenning

Naast programma's voor automatische tekstanalyse bestaan er ook oplossingen voor automatische beeldherkenning. De computer kan dan bijvoorbeeld herkennen hoeveel mensen er staan afgebeeld in een foto, of ze lachen of verdrietig kijken, wat voor andere objecten er afgebeeld zijn in de foto, en in wat voor omgeving de foto is genomen. Ook hier geldt dat de technologie zich enorm snel ontwikkelt, dus het is lastig om specifieke software aan te raden. Wel moet je hier vaak voor kunnen programmeren om automatisch grote verzamelingen afbeeldingen te coderen.[8]

7 Soms wordt hier ook bestaande data voor gebruikt, zoals recensies van films. De computer wordt dan getraind om de scores (het aantal sterren dat mensen hebben toegekend aan een film) te voorspellen op basis van de recensie.

8 Daarnaast heeft software voor automatische beeldherkenning vaak een sterke computer nodig, dus niet iedere computer is geschikt om automatisch afbeeldingen te coderen.

ChatGPT en andere taalmodellen

Voor sentiment-analyse en automatische beeldherkenning geldt dat er specifieke programma's zijn die gemaakt zijn om één taak op te lossen. Sinds de introductie van ChatGPT op 30 november 2022 kan iedereen gebruik maken van grote taalmodellen (Engels: *Large Language Models*, of *LLMs*) die kunnen helpen met allerlei verschillende taken. Zo kun je ChatGPT vragen om algemene vragen te beantwoorden, teksten samen te vatten, of zelfs om een gedicht te schrijven. Naast ChatGPT zijn er ook andere grote taalmodellen die vergelijkbare taken kunnen uitvoeren. Er zijn commerciële oplossingen zoals Claude, Google's Gemini, en Microsoft's Copilot. Daarnaast zijn er ook verschillende *open source* beschikbare taalmodellen (hoewel het per model verschilt hoe *open* het model nu echt is; zie Liesenfeld et al. (2023) voor een bespreking). Hoewel veel huidige taalmodellen *unimodaal* zijn, en zich dus alleen richten op tekst, zijn er ook *multimodale* taalmodellen die vragen kunnen beantwoorden over afbeeldingen of video's. Omdat computers nooit moe worden, en grote verzamelingen data kunnen coderen in korte tijd, lijkt het een aantrekkelijke optie om taalmodellen te gebruiken om automatische inhoudsanalyses uit te voeren.

Een van de grootste uitdagingen bij het uitvoeren van een inhoudsanalyse is om de studie zo op te zetten dat de resultaten valide en betrouwbaar zijn. Daarom hebben we in dit werkboek ook zoveel aandacht besteed aan het ontwikkelen van codeboeken en het meten van intercodeursbetrouwbaarheid. Bij grote taalmodellen speelt precies dezelfde uitdaging: hoe kunnen we ervoor zorgen dat het model valide en betrouwbare resultaten geeft?

Taalmodellen hebben duidelijke instructies nodig. Waar de instructies bij menselijke codeurs de vorm aannemen van een codeboek en trainingsessies (waarbij codeurs oefenen met het codeboek en de resultaten bespreken met de hoofdonderzoeker), krijgen taalmodellen geschreven instructies in de vorm van een zogenaamde *prompt*. Daarin wordt aangegeven hoe de algemene taak eruit ziet, wat er precies gecodeerd moet worden, hoe er precies gecodeerd moet worden, en hoe de output van het model eruit moet zien. *Prompting* is geen exacte wetenschap; het is niet altijd duidelijk hoe je nu precies een succesvolle prompt moet formuleren, en de prestaties van taalmodellen kunnen sterk variëren, afhankelijk van hoe de prompt precies geformuleerd is (Mizrahi et al., 2024).

Taalmodellen moeten getest worden. We kunnen niet blind vertrouwen op de prestaties van een willekeurig taalmodel. Alle modellen maken fouten, en het is aan jou als onderzoeker om te bepalen of het model dat je wil gebruiken betrouwbaar genoeg is. Daarvoor kun je ten minste twee benaderingen kiezen:

1) **Behandel de evaluatie van het taalmodel als een pilotstudie**: laat het model een verzameling boodschappen coderen, en laat een groep menselijke codeurs hetzelfde doen. Daarna kun je de resultaten vergelijken door te kijken naar de verschillen en overeenkomsten tussen de twee condities. Daarbij kun je op zoek gaan naar patronen: zit er enige regelmaat in de fouten die gemaakt worden? Het verschil tussen mensen en computers is dat je aan mensen kunt vragen om te verantwoorden waarom ze een bepaalde keuze

hebben gemaakt. Of tenminste: je kunt een taalmodel wel vragen stellen, maar taalmodellen zijn niet getraind om aan introspectie te doen; ze zijn getraind om overtuigende antwoorden te geven, zonder dat deze antwoorden waar hoeven te zijn.

2) **Behandel de evaluatie van het taalmodel als het testen van software**: het is goed om te kijken wat het model zegt over voorbeelddata, maar daarnaast is het ook goed om lastige voorbeelden voor te leggen aan het model. Deze voorbeelden mag je ook zelf verzinnen! Zo kun je controleren hoe het model omgaat met twijfelgevallen, situaties die sterk context-afhankelijk zijn (humor, sarcasme), en verwijzingen naar recente gebeurtenissen uit ons collectieve bewustzijn. (Taalmodellen worden veelal getraind met oudere data, dus ze zijn vaak niet blootgesteld aan gebeurtenissen die onlangs hebben plaatsgevonden.)

Verder is het belangrijk om te weten dat grote taalmodellen zeker niet altijd beter zijn dan kleinere, en dat specialistische programma's nog regelmatig beter presteren. Zo laten W. Zhang et al. (2024) bijvoorbeeld zien dat grote taalmodellen (anno 2024) wel in staat zijn om oppervlakkige sentiment-analyses (bijvoorbeeld: *is deze filmrecensie positief of negatief?*) uit te voeren, maar dat ze nog worstelen met uitdagendere taken.

Ten slotte nog een waarschuwing: **Taalmodellen zijn niet altijd veilig om te gebruiken**. Uiteindelijk zijn alle online taalmodellen gewoon programma's op de computer van een andere organisatie. Alle gegevens die je invoert, kunnen ook weer gebruikt worden door die organisatie. (Bijvoorbeeld om het model verder te trainen.) Daarom moet je voorzichtig zijn met de data die je geeft aan het model, omdat vertrouwelijke informatie anders openbaar gemaakt zou kunnen worden. Afhankelijk van je computer (je hebt een krachtige processor nodig) kun je er ook voor kiezen om alleen lokale taalmodellen te gebruiken. Je kunt bijvoorbeeld het programma LM Studio gebruiken om te kijken wat je met openbaar beschikbare modellen kunt doen.

7.3.3 SPECIFICITEIT

Bestaande automatische oplossingen zijn vaak erg specifiek; ze zijn gericht op één specifiek genre of domein. De term **genre** verwijst naar de tekstsoort, bijvoorbeeld berichten op sociale media, artikelen in kranten of tijdschriften, literatuur, enzovoorts. De term **domein** verwijst vaak naar het onderwerp van die teksten, zoals nieuws, vrije tijd, of medische teksten. Als jouw probleem afwijkt op één of beide onderdelen, dan wordt het nut en de betrouwbaarheid van het automatische systeem lager. Hieronder staan twee voorbeelden:

- Het programma VADER (Hutto & Gilbert, 2014) is specifiek ontwikkeld om sentiment-analyse uit te voeren voor korte Engelstalige berichten op sociale media. Het werkt redelijk goed voor vergelijkbare teksten, maar het is onduidelijk hoe goed het werkt voor andere documenten. Aangezien het taalgebruik op sociale media steeds verandert, kun je je ook

afvragen of positieve en negatieve woorden van meer dan 10 jaar geleden ook nog steeds dezelfde associaties met zich meedragen.
- Het programma Stanza (Qi et al., 2020) is ontwikkeld voor verschillende taken, waaronder *Named Entity Recognition* (herkennen van de namen van personen, locaties, en organisaties in teksten). Het programma is ontwikkeld aan de hand van verschillende soorten teksten, maar met een sterke nadruk op nieuws en tekst van Wikipedia. De auteurs hebben later een apart model ontwikkeld dat specifiek gericht is op biomedische teksten (Y. Zhang et al., 2021).

Een andere eigenschap van automatische oplossingen is dat ze soms gericht zijn op een specifieke verzameling concepten of entiteiten.[9] Een voorbeeld daarvan is software voor automatische beeldherkenning. Nanne et al. (2020) bespreken drie van zulke programma's, die tussen de 78 en 3.577 verschillende soorten entiteiten herkennen. Die ondergrens van 78 categorieën heeft geen diepere betekenis, maar komt voort uit een beslissing die ontwikkelaars van de MS COCO dataset (T.-Y. Lin et al., 2014) ooit hebben gemaakt, omdat ze op zoek waren naar een diverse verzameling van 80 objecten om te testen hoe goed computers zijn in het herkennen van afbeeldingen.

Het kan verleidelijk zijn om te kijken naar wat er allemaal al automatisch kan (bijvoorbeeld: welke objecten kunnen bestaande programma's voor mij herkennen in een verzameling van afbeeldingen?) en daar jouw onderzoek op aan te passen. Je kunt bijvoorbeeld besluiten om de onderzoeksgroepen alleen maar te vergelijken in termen van eigenschappen die automatisch te detecteren zijn met bestaande software. Op dat moment is het belangrijk je af te vragen *waarom* je het onderzoek eigenlijk wil doen; verlies je hiermee de relevantie van je onderzoek, of beantwoord je nog steeds een relevante vraag?

7.3.4 AUTOMATISCH VERSUS SEMI-AUTOMATISCH

Hierboven hebben we verschillende automatische oplossingen besproken, samen met hun voor- en nadelen. Maar de kwaliteit van je onderzoek staat of valt met de manier waarop je die automatische oplossingen gebruikt. We kunnen een onderscheid maken tussen volledig automatische en semi-automatische oplossingen.

Volledig automatische oplossingen

Volledig automatische oplossingen zijn (relatief) betrouwbaar. Je kunt deze oplossingen inzetten om menselijke codeurs te vervangen. Hierbij kun je onder andere denken aan:
- Zinslengte (aantal karakters, woorden)
- Voorkomen van specifieke woorden of karakters (inclusief emoji).

9 Een entiteit is een bestaand ding.

- Andere oplossingen die je getest hebt voor jouw data.[10]

Semi-automatische oplossingen

Als er geen volledig automatische oplossingen voorhanden zijn, kun je nog proberen om te werken met een *semi-automatische* aanpak. Daarbij laat je eerst de computer door alle data heengaan, en dan controleer je handmatig of alles klopt. Het idee hierachter is dat corrigeren sneller gaat dan controleren, als de computer *meestal* gelijk heeft. Een vuistregel die meestal klopt wordt ook wel een *heuristiek* genoemd.[11] We kunnen bij het coderen ten minste twee soorten heuristieken onderscheiden:

1) Woorden die meestal duiden op een bepaalde categorie X, maar niet altijd. (Bijvoorbeeld: als iemand *sorry* zegt, dan is het meestal een verontschuldiging.) In dat geval moeten we alle data (zowel positieve als negatieve gevallen) controleren, maar als het goed is hoef je weinig te corrigeren.
2) Woorden die meestal voorkomen bij een bepaalde categorie X, maar niet altijd. Als een bepaald woord of een bepaalde uitdrukking eigenlijk altijd wijst op een bepaalde categorie, kun je al die gevallen eenvoudig coderen. Het is dan vooral zaak om de negatieve gevallen te controleren of ze misschien ook behoren tot de relevante categorie.

In plaats van een semi-automatische benadering op basis van woordenlijsten, kun je natuurlijk ook andere technieken gebruiken. Bijvoorbeeld automatische taalherkenning (welke taal wordt er gesproken in deze boodschap?) of sentiment-analyse. Als de software maar betrouwbaar genoeg is, kan een semi-automatische benadering veel tijdswinst opleveren.

7.3.5 DEKKING

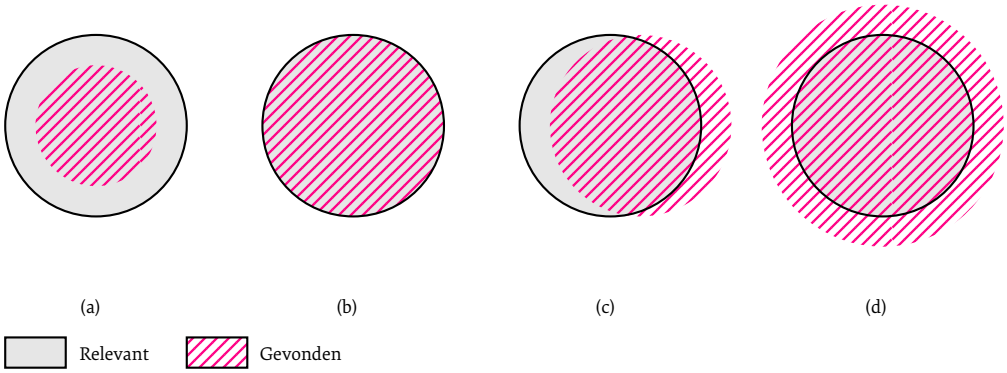
Bij automatische inhoudsanalyse is het belangrijk om na te denken over de *dekking* (Engels: *coverage*) van het algoritme dat je gebruikt: welke voorbeelden worden er wel of niet gevonden? Figuur 7.1 illustreert de verschillende mogelijkheden: afhankelijk van de manier waarop je de data codeert, kun je ook andere items (juist of onjuist) identificeren als behorend tot een bepaalde categorie. Bijvoorbeeld: als je alle geschreven berichten wil coderen die een verontschuldiging bevatten, dan kun je automatisch alle berichten die het woord “sorry”

10 Resultaten die behaald zijn met andere genres en domeinen zijn geen garantie voor succes. Maar als je gecontroleerd hebt of het programma werkt met specifieke voorbeelden en verzonnen data (zie ook de eerdere discussie over taalmodellen), dan kun je er redelijk zeker van zijn dat de oplossing goed werkt.

11 Misschien ken je deze term ook al uit het werk van Daniel Kahneman (2011). In zijn werk legt hij uit hoe mensen beslissingen kunnen nemen op twee manieren, namelijk: snel en intuïtief (dit noemt Kahneman *Systeem 1*) of langzamer en bewuster (*Systeem 2*). Waar Systeem 2 gebruik maakt van een uitgebreid redeneerproces, werkt Systeem 1 op basis van emotie en heuristieken, zoals die geïdentificeerd zijn in eerder werk (bijvoorbeeld: Tversky & Kahneman (1974)).

bevatten als zodanig coderen. Waarschijnlijk zul je daarmee veel relevante berichten vinden, maar niet zeker niet alle. Je kunt je hierbij twee vragen stellen:

1) Hoeveel **relevante** berichten zou je goed/fout hebben gecodeerd met deze methode?
2) Hoeveel **irrelevante** berichten zou je goed/fout hebben gecodeerd met deze methode?



FIGUUR 7.1 *Visualisatie van de verschilende mogelijkheden bij het automatisch coderen van verschillende categorieën. Van alle relevante items (lichtgrijs) is het mogelijk dat een computer (a) een deel van de relevante items codeert (roze gearceerd), (b) precies alle relevante items codeert (en geen andere items), (b) een deel van de relevante en een deel van de irrelevante items codeert, of (d) alle relevante en sommige irrelevante items codeert.*

Als je de meeste relevante items waarschijnlijk wel automatisch kunt identificeren, maar de automatische methode overgeneraliseert, dan kun je achteraf de automatisch toegewezen labels handmatig corrigeren. Dit is vaak het beste scenario, omdat je meer relevante items kunt coderen met minder werk. In de omgekeerde situatie (als het systeem maar weinig relevante items kan detecteren) is dat een stuk lastiger; je zult dan alsnog het hele corpus moeten doorlopen. De vraag is dan hoe nuttig het nog is om automatisch te coderen (maar zie 7.3.6 hieronder).

Bias

Een andere vraag die je kunt stellen omtrent de dekking van automatische oplossingen, is of bepaalde soorten uitingen stelselmatig verkeerd worden gecodeerd. Als dat zo is, moet je je afvragen of die systematische *bias* invloed kan hebben op je onderzoeksresultaten. Zouden de fouten van het systeem bijvoorbeeld gelijkmatig verdeeld zijn over de groepen die je vergelijkt, of zou het kunnen dat het systeem veel vaker fouten maakt bij één van die groepen?

7.3.6 VOLGORDE

Naast het automatisch vooraf coderen, kun je automatische oplossingen ook gebruiken om je werk achteraf te controleren. Je bent dan eigenlijk dubbel aan het coderen, samen met de computer. Door middel van verschillende heuristieken kun je dan controleren of je in ieder geval geen automatisch herkenbare voorbeelden hebt gemist. Het voordeel hiervan is dat het codeerproces hiermee betrouwbaarder wordt, maar het nadeel is dat je nog steeds zelf het hele corpus aan het coderen bent.

7.4 MODULARITEIT

Modulariteit betekent dat je verschillende onderdelen van je project splitst in losse eenheden, of *modules*. Dat heeft een aantal voordelen. Wanneer je verschillende taken scheidt van elkaar:

- Krijg je een beter overzicht van de taken die nog verricht moeten worden binnen het project.
- Kunnen verschillende groepsleden tegelijkertijd werken aan verschillende onderdelen.
- Is het achteraf ook duidelijker hoe je verschillende onderdelen aangepakt hebt.
- Kun je in toekomstige projecten eenvoudiger onderdelen uit eerdere studies hergebruiken.

Hieronder bespreken we onderdelen die je beter van elkaar gescheiden kunt houden.

7.4.1 SCHEIDING VAN DATAVERZAMELING EN DATA-ANALYSE

Als je aan het programmeren bent, kan het erg verleidelijk zijn om zowel de dataverzameling als het coderen in één keer uit te voeren. Dat lijkt erg efficiënt, want bij het verzamelen heeft de computer alle data toch al geladen, dus dan kun je alles net zo goed meteen coderen. Maar in de praktijk levert deze benadering vaak veel ellende op, zoals het verkeerd of helemaal niet opslaan van de oorspronkelijke berichten. En wanneer je de analyse toch nog wil aanpassen, betekent dit anders (wanneer je alles in één keer doet) dat je ook de data helemaal opnieuw moet downloaden.[12] Dus sla de ruwe data tussentijds op voordat je begint met het coderen. Daarmee voorkom je dat er iets misgaat met de data. Verdere voordelen van modulariteit:

- **Transparantie**: Door de data eerst op te slaan en een apart script te maken voor het coderen wordt het onderzoek transparanter (want je kunt precies zien hoe de ruwe data eruit ziet) en is het makkelijker om het werk onderling te verdelen (want iedereen kan zelfstandig aan de slag met de ruwe data).

12 Opnieuw data moeten downloaden kan betekenen dat de onderzoeksdata verandert of zelfs helemaal niet meer beschikbaar is. Bovendien is het ook minder belastend voor de websites waar je data van verzamelt om de data maar één keer te downloaden.

- **Stabiliteit**: Bovendien heb je dan ook de garantie dat de data stabiel blijft tijdens het onderzoeksproject; programma's die automatisch data verzamelen zijn vaak afhankelijk van het tijdstip waarop je die programma's uitvoert. (Bijvoorbeeld: de meest populaire berichten van vandaag kunnen compleet anders zijn dan de meest populaire berichten van morgen.)[13]
- **Taakverdeling**: Als je met elkaar van tevoren een vast data-format afspreekt, kun je alvast beginnen met het ontwikkelen van de codeer-scripts voordat de dataverzameling compleet is.

Dus: sla onderzoeksdata altijd meteen op, inclusief alle relevante metadata. Daarna kun je verder gaan met het coderen.

7.4.2 SCHEIDING VAN VERSCHILLENDE ANALYSES

Naast de scheiding van dataverzameling en het coderen, is het ook belangrijk om verschillende analyses van elkaar te onderscheiden. Daarbij zijn er twee sub-onderscheiden die we kunnen maken:

1) **Scheiding van verschillende variabelen**. Als je verschillende variabelen automatisch codeert, is het nuttig om de code voor het berekenen van de verschillende variabelen niet te veel te vervlechten met elkaar. Door de berekeningen gescheiden te houden van elkaar, kun je precies laten zien hoe iedere losse variabele berekend wordt.
2) **Scheiding van coderen en resultaten**. Als je verschillende variabelen automatisch codeert, kan het verleidelijk zijn om in hetzelfde programma ook meteen descriptieve (bijvoorbeeld: gemiddelde, standaarddeviatie) en toetsende statistieken (t-test, chikwadraattoets) te berekenen. Doe dat niet. Het is beter om de resultaten van het automatisch coderen eerst op te slaan, en vervolgens het opgeslagen bestand te analyseren. Zo zorg je voor extra transparantie en documentatie van het proces. Daarnaast kun je de taken ook eenvoudiger verdelen: als je afspraken maakt over de bestandsindeling, kunnen verschillende teamleden apart van elkaar werken aan het coderen en het berekenen van de resultaten.

13 Een ander probleem is dat platforms hun API (programmeer-interface) soms aanpassen, waardoor andere programma's opeens niet meer werken.

7.5 TOT SLOT: PROGRAMMEREN IN EXCEL

Als je twijfelt aan je programmeervaardigheden, dan kun je nog steeds een (semi-)automatische inhoudsanalyse uitvoeren. Excel en andere spreadsheetprogramma's hebben namelijk allerlei ingebouwde functies om data te analyseren. Daarnaast zijn er ook verschillende *plugins* beschikbaar die de functionaliteit van deze programma's nog verder uitbreiden. Voor het coderen van data in Excel heb je wel wat voorkennis nodig, die we bespreken in Bijlage G.

HOOFDSTUK 8
Ethiek

8.1 INLEIDING

In dit hoofdstuk gaan we dieper in op de ethische vragen rondom het uitvoeren van een inhoudsanalyse, bijvoorbeeld:

- Mag je zomaar data verzamelen van sociale media?
- Moet je eigenlijk ethische goedkeuring aanvragen voor het analyseren van publieke data?
- Wat betekent dit voor de privacy van de auteurs?
- Welke ethische implicaties zijn er bij samenwerkingen met externe organisaties?

Daarnaast zullen we ook kijken naar de bredere kaders die er bestaan om op een verantwoordelijke manier onderzoek te doen. Waarschijnlijk heb je al eerder iets gelezen over de Nederlandse gedragscode voor wetenschappelijke identiteit, of nagedacht over de rol van de onderzoeker: kun je bijvoorbeeld echt onafhankelijk onderzoek doen? Daarover lees je hieronder meer.

8.2 ETHISCHE GOEDKEURING

Wanneer je als onderzoeker aan de universiteit een studie wil doen, moet je vaak een aanvraag indienen bij de ethische commissie. Zo'n commissie wordt in het Engels vaak *Institutional Review Board* (IRB) genoemd. Afhankelijk van de situatie en de verantwoordelijkheden van de ethische commissie worden er ook andere namen gebruikt. In Tilburg is er bijvoorbeeld de *Research Ethics and Data Management Committee* (REDC) die niet alleen onderzoeksethiek

beoordeelt, maar ook het datamanagementplan (dat hier nauw mee verbonden is).[1] Voor medisch-wetenschappelijk onderzoek zijn er ook weer aparte commissies: Medisch-Ethische Toetsingscommissies (METC) en de Centrale Commissie Mensgebonden Onderzoek (CCMO).[2]

Een ethische aanvraag doe je voorafgaand aan de studie, zodat de commissie kan beoordelen of het voorstel voldoet aan de huidige ethische normen en aan de wet- en regelgeving omtrent dataverzameling en -opslag. Veel wetenschappelijke tijdschriften stellen als eis voor publicatie dat alle onderzoeken door een ethische commissie beoordeeld moeten zijn. Het is ook mogelijk dat een onderzoek door de regels van de ethische commissie vrijgesteld wordt van beoordeling. Vroeger gold dat vaak voor onderzoeken met bestaande data, maar tegenwoordig wordt daar genuanceerder naar gekeken.

Het beoordelen van onderzoeksvoorstellen is geen eenvoudige taak, want de beoordeling hangt vaak af van de context: de commissie weegt verschillende waarden (zoals *Privacy* en *Transparantie*) tegenover het doel en de voor- en nadelen van het onderzoek. Hetzelfde studie-ontwerp kan dus in het ene geval wel, en in het andere geval niet goedgekeurd worden. Onderzoeken door studenten (in het kader van onderwijs) worden meestal niet beoordeeld door de ethische commissie, maar door de docent.

8.3 BESTAANDE NASLAGWERKEN OVER ETHIEK

Aan eerdere literatuur over inhoudsanalyse kun je zien dat de aandacht voor onderzoeksethiek de laatste jaren een vlucht heeft genomen. Neuendorf (2017, pp. 130–131) besteedt in haar boek bijvoorbeeld slechts een halve pagina aan ethiek. Ze geeft aan dat goedkeuring door een ethische commissie vaak niet nodig is, omdat inhoudsanalyses meestal openbaar beschikbare data bestuderen. De enige uitzondering die ze hiervoor geeft is het coderen van data die door participanten in een experiment zijn gegenereerd. Daarnaast noemt Neuendorf een artikel van Nancy Signorielli waarin ethiek aan bod komt. Signorielli (2010) geeft aan dat eerdere naslagwerken[3] helemaal geen aandacht besteden aan ethiek. Zelf identificeert Signorielli (2010) drie discussiepunten:

1) Data-analyse en rapportage moet eerlijk en transparant zijn. Dit zien we ook nadrukkelijk terug in de Nederlandse gedragscode voor wetenschappelijk onderzoek die we hieronder zullen bespreken.
2) We moeten rekening houden met het mentale welzijn van codeurs; zeker wanneer zij boodschappen moeten coderen die dat welzijn kunnen schaden. In ieder geval is het

1 Lees hier meer over de REDC: https://www.tilburguniversity.edu/nl/onderzoek/ethics-review-boards/humanities

2 Zie: https://www.ccmo.nl/metcs

3 Signorielli (2010) noemt de boeken van Krippendorff (uit 1980 en 2004), Riffe et al. (uit 1998), en ook een eerdere editie van het naslagwerk van Neuendorf (uit 2002).

belangrijk om een debriefing te geven, waarin uitgelegd wordt wat het doel is van het onderzoek.[4]

3) Informed consent is noodzakelijk wanneer we data verzamelen uit privégesprekken. Signorielli (2010) erkent wel dat er een grijs gebied is voor online chatgesprekken waarvoor je moet inloggen, maar benoemt vooral praktische problemen, zoals het idee dat mensen ander gedrag vertonen wanneer ze weten dat ze bestudeerd worden.[5]

Tegenwoordig worden er in de literatuur uitvoerige discussies gevoerd over de verschillende ethische vragen die spelen bij het uitvoeren van inhoudsanalyses, en de *Association of Internet Research* heeft ethiek een prominente plaats gegeven op haar website: https://aoir.org/ethics/

8.4 DE NEDERLANDSE GEDRAGSCODE WETENSCHAPPELIJKE INTEGRITEIT

Alle wetenschappers in Nederland zijn gebonden aan de Nederlandse gedragscode voor wetenschappelijke integriteit (KNAW et al., 2018). Je kunt de gedragscode online lezen; het is een erg toegankelijk document. Als je studeert of onderzoek doet aan de universiteit, word je geacht het document gelezen en begrepen te hebben. In de appendix staan de relevante definities uit de gedragscode op één pagina, zodat je deze eenvoudig kunt printen. Naast de algemene Nederlandse gedragscode hebben universiteiten zelf vaak ook eigen gedragscodes.[6] Deze gedragscodes bouwen voort op de nationale gedragscode,[7] en geven bijvoorbeeld extra richtlijnen om de sociale veiligheid op de universiteit te garanderen. Hieronder worden de vijf principes uit de Nederlandse gedragscode verder besproken, in de context van het ontwerpen en uitvoeren van een inhoudsanalyse.

4 Dit probleem is de laatste jaren nog prominenter geworden dankzij het gebruik van crowdsourcing: online marktplaatsen waar mensen anoniem laagbetaald werk uitvoeren. Zie Gray & Suri (2019) voor een uitgebreide discussie.

5 Dit wordt ook wel het *Hawthorne effect* genoemd. Zie McCambridge et al. (2014) voor een systematische review van studies over dit effect.

6 Je kunt de gedragscode van Tilburg University hier lezen: https://www.tilburguniversity.edu/nl/over/gedrag-integriteit/gedragsregels. Als je in Tilburg werkt of studeert word je ook geacht deze gedragscode te kennen.

7 In het geval van conflicten tussen de verschillende gedragscodes blijft de nationale gedragscode leidend.

8.4.1 EERLIJKHEID

De eerste waarde uit de Nederlandse gedragscode wordt als volgt gedefinieerd:[8]

> "**Eerlijkheid** *houdt onder andere in dat men:*
> - *geen ongefundeerde claims doet,*
> - *over het onderzoeksproces correct rapporteert,*
> - *data of bronnen niet verzint of vervalst,*
> - *alternatieve visies en tegenargumenten serieus neemt,*
> - *open is over onzekerheidsmarges, en*
> - *resultaten niet gunstiger dan wel ongunstiger voorstelt dan ze zijn.*"

Dit is een basisvoorwaarde om wetenschappelijke discussies te houden. Oneerlijkheid schaadt het vertrouwen dat we als maatschappij hebben in het onderzoek dat aan de universiteit gedaan wordt.[9] De definitie geeft ook een brede opvatting van het begrip *Eerlijkheid*: dat betekent niet alleen dat je niet mag liegen, maar ook dat je verschillende argumenten en resultaten zorgvuldig moet wegen om tot een weloverwogen conclusie te komen. Het is dus ook niet oké om alleen maar artikelen op te zoeken die jouw zienswijze ondersteunen, en die te presenteren als hét standpunt van de wetenschap. (Dit zien we hieronder nogmaals benadrukt bij het principe *Onafhankelijkheid*.)

8.4.2 ZORGVULDIGHEID

De tweede waarde uit de Nederlandse gedragscode wordt als volgt gedefinieerd:

> "**Zorgvuldigheid** *houdt onder andere in dat men wetenschappelijke methoden gebruikt en optimale precisie betracht bij het ontwerp, de uitvoering, verslaglegging en disseminatie van het onderzoek.*"

Deze waarde toont het belang van onderzoeksmethodologie, en het maken van doordachte keuzes in iedere fase van het onderzoek. Optimale precisie betrachten betekent ook dat er een proces moet zijn om te controleren dat alles inderdaad klopt. Dus bijvoorbeeld:

8 Alle citaten uit de zijn steeds rechtstreeks ontleend uit de Nederlandse gedragscode wetenschappelijke integriteit (KNAW et al., 2018). De formattering is aangepast ter verbetering van de presentatie in de context van dit werkboek.

9 Diederik Stapel is een bekend voorbeeld van iemand die zich niet aan deze regel heeft gehouden. Hij publiceerde verschillende artikelen met verzonnen data, en bracht daarmee schade toe aan de wetenschap, zijn mede-auteurs, en zichzelf. Nadat bekend werd dat de artikelen van Stapel niet betrouwbaar waren, is er een uitgebreid onderzoek ingesteld door Commissies Levelt, Noort, en Drenth (2012).

- Dat je overlegt met elkaar (en eventueel een derde partij) om je ideeën te toetsen.
- Dat je niet zomaar de eerste versie van een verslag inlevert, maar elkaars schrijfwerk controleert en herziet om eventuele fouten te verbeteren.

Sommige onderzoekers, zoals Lakens (2023), beargumenteren ook dat er naast *ethische* commissies ook *methodologische* commissies zouden moeten komen om onderzoeksvoorstellen te toetsen op de kwaliteit van het ontwerp. Zorgvuldigheid in latere fases wordt niet stelselmatig gecontroleerd, hoewel artikelen wel beoordeeld worden door collega's voordat ze kunnen worden gepubliceerd.

8.4.3 TRANSPARANTIE

De volgende waarde, *Transparantie*, wordt als volgt gedefinieerd door de Nederlandse gedragscode:

> "**Transparantie** *houdt onder andere in dat het voor anderen helder is:*
> - *op welke data men zich heeft gebaseerd,*
> - *hoe deze zijn verkregen,*
> - *welke resultaten men heeft bereikt en langs welke weg, en*
> - *wat de rol van externe belanghebbenden is geweest.*
>
> *Als delen van het onderzoek of van de data niet toegankelijk worden gemaakt, dient de onderzoeker goed gemotiveerd aan te geven waarom dat niet mogelijk is.*
>
> *De wijze van uitvoering en fasering van het onderzoeksproces moet tenminste voor vakgenoten te volgen zijn. Dit betekent in ieder geval dat de argumentatie helder moet zijn en dat de stappen in het onderzoeksproces controleerbaar moeten zijn.*"

Hoewel deze waarde gedeeld wordt door alle wetenschappers, zijn er wel verschillen tussen individuele onderzoekers en ook verschillende implicaties voor verschillende soorten onderzoek.

Hoe maak je je onderzoek transparant? Onderzoek kun je transparant maken door in alle fasen van het onderzoek bij te houden wat je doet en wat je in iedere volgende stap wil gaan doen, en door alle bijbehorende data en programmeercode met de onderzoeksgemeenschap te delen. Nosek et al. (2015) geven een overzicht van de mogelijkheden die je hier hebt als onderzoeker, en van de transparantie-eisen die wetenschappelijke tijdschriften zouden kunnen stellen aan auteurs.

Een van de transparantie-eisen die Nosek et al. (2015) noemen is het *preregistreren* van je onderzoek. Dat betekent dat je voorafgaand aan je studie een onderzoeksplan schrijft met specifieke, toetsbare hypothesen, en een heldere omschrijving van het ontwerp van je studie.

Daarbij specificeer je ook het analyseplan, waarbij je aangeeft hoe de data verwerkt zal worden, welke exclusiecriteria gehanteerd zullen worden, en welke toetsen er gebruikt zullen worden. Dat plan kun je indienen bij een tijdschrift, of bij een speciale website waar alle onderzoeksplannen samen worden opgeslagen met de exacte datum en het tijdstip waarop je die plannen hebt ingediend.[10] Wanneer de onderzoeksresultaten binnen zijn, voer je het onderzoek precies zo uit als je het gepland hebt. Anderen kunnen dan achteraf zien dat je niet zomaar wat bent gaan vissen in je data om maar significante resultaten te vinden.[11] Preregistratie is populair omdat het wordt gezien als een middel om de betrouwbaarheid van wetenschappelijk onderzoek te verhogen.[12] Daarnaast biedt preregistratie ook veel praktische voordelen: als je voorafgaand aan je onderzoek een uitgewerkt plan maakt van wat je wil gaan doen, dan helpt dit om latere problemen te voorkomen (want je hebt al over alles nagedacht) en het zorgt er ook voor dat je een goede basis hebt voor je onderzoeksverslag.

Hoe open moet de wetenschap zijn? Openheid heeft ook grenzen. In Hoofdstuk 2 bespraken we al het probleem van *Dual use*: sommige studieresultaten kunnen ook door anderen misbruikt worden. Het is dan belangrijk om voorzichtig te zijn met het volledig openbaar maken van de data en programmeercode die je gebruikt hebt. Daarnaast kan het openbaar maken van je data ook gevolgen hebben voor verschillende belanghebbenden (Engels: *stakeholders*), zoals de auteurs van de boodschappen die je hebt geanalyseerd, of de mensen die in die boodschappen omschreven worden. Als onderzoeker heb je de verantwoordelijkheid om belanghebbenden te beschermen tegen potentiële negatieve gevolgen van je onderzoek. Dat kan soms betekenen dat je jouw onderzoeksdata niet, of slechts beperkt kunt publiceren.

Transparantie komt in de volgende twee hoofdstukken (*H9: Datamanagement* en *H10: Rapportage*) uitgebreid aan bod.

8.4.4 ONAFHANKELIJKHEID

Het vierde principe uit de Nederlandse gedragscode wordt als volgt gedefinieerd:

> "**Onafhankelijkheid** *houdt onder andere in dat men zich in de keuze van de methode, bij de beoordeling van de data en in de weging van alternatieve verklaringen, maar ook bij het beoordelen van onderzoek of onderzoeksvoorstellen van anderen, niet laat leiden door buiten-wetenschappelijke overwegingen (bijvoorbeeld overwegingen van commerciële of politieke aard). Aldus gefor-*

10 Bekende preregistratiewebsites zijn http://aspredicted.org en http://osf.io.

11 Het is bij een preregistratie echter wel mogelijk om af te wijken van je plan, maar dan moet je duidelijk in je verslag aangeven waar en waarom je bent afgeweken van je originele plan. Afhankelijk van de aanpassingen en de redenen voor die aanpassingen kan dit wel implicaties hebben voor de wetenschappelijke status van je resultaten.

12 Waar preregistratie eerst alleen werd voorgesteld als een hulpmiddel voor kwantitatief onderzoek, gaan er ook stemmen op om kwalitatief onderzoek te preregistreren (bijvoorbeeld L. Haven & Van Grootel (2019).

muleerd omvat onafhankelijkheid ook onpartijdigheid. Onafhankelijkheid is in elk geval vereist bij de opzet en uitvoering van en rapportage over het onderzoek; bij de keuze van het onderzoeksobject en van de onderzoeksvraag is onafhankelijkheid niet altijd nodig"

Hier zien we een duidelijke tweedeling tussen het bepalen van het onderwerp en het uitvoeren en rapporteren van het onderzoek. In de uitvoering en rapportage is het belangrijk dat je onafhankelijk en onpartijdig bent, zodat de maatschappij op jouw deskundige oordeel kan vertrouwen. Maar zelf ben je natuurlijk ook onderdeel van de maatschappij, en maak je je misschien ook zorgen over bepaalde ontwikkelingen (bijvoorbeeld over de klimaatcrisis en hoe we daarmee omgaan, of ongelijkheid binnen de maatschappij). Het is logisch dat die zorgen jou dan motiveren om een bepaald onderwerp verder te bestuderen, ook gezien de verantwoordelijkheid (zie hieronder) die je als onderzoeker hebt naar de maatschappij. Het blijft natuurlijk wel belangrijk om eerlijk te zijn over jouw onderzoeksresultaten, en die niet te verdraaien om het probleem groter te doen lijken dan het is.

Conflicterende belangen

Naast verschillende (en soms conflicterende) waarden gaat ethiek ook over machtsverhoudingen. Het kan bijvoorbeeld gebeuren dat je graag bepaalde data wil analyseren, maar dat je daarvoor wel toestemming nodig hebt van een commerciële organisatie. (Bijvoorbeeld omdat de data alleen intern beschikbaar zijn.) Op dat moment moet je een overeenkomst sluiten waarin vastgelegd wordt hoe je met de data omgaat, en met welk doel je de data analyseert. Maar geen enkele organisatie wil natuurlijk in een kwaad daglicht komen te staan, en je bent van de organisatie afhankelijk om de data te krijgen. Er is dan een scheve machtsverhouding tussen jou en de organisatie. Het is in dat soort situaties belangrijk om waakzaam te zijn en jouw onafhankelijkheid niet op te geven. Wel kun je kijken wat jullie gezamenlijke belangen zijn, en hoe je de identiteit van de organisatie zou kunnen beschermen. (Bijvoorbeeld door een algemenere omschrijving te kiezen: 'een kledingzaak in Nederland' in plaats van 'Henks hippe hoedenwinkel.')

8.4.5 VERANTWOORDELIJKHEID

Het laatste principe uit de Nederlandse gedragscode voor wetenschappelijke integriteit is *Verantwoordelijkheid.* Dat wordt als volgt gedefinieerd:

"**Verantwoordelijkheid** *houdt onder andere in dat men zich rekenschap geeft van het feit dat men als onderzoeker niet in isolement opereert, en daarom binnen de grenzen van het redelijke rekening houdt met de legitieme belangen van bij het onderzoek betrokken personen en dieren, van eventuele opdrachtgevers en financiers, en van de omgeving. Verantwoordelijkheid houdt ook in dat men onderzoek doet dat wetenschappelijk en/of maatschappelijk relevant is.*"

Hierboven hebben we al gezien dat er wrijving kan ontstaan tussen *Transparantie* en *Verantwoordelijkheid* doordat de publicatie van jouw onderzoek kan leiden tot directe of indirecte schade aan verschillende belanghebbenden. Ook is er een zekere overlap tussen *Onafhankelijkheid* en *Verantwoordelijkheid* waar het gaat om de afweging tussen maatschappelijke belangen en de belangen van eventuele opdrachtgevers of financiers.

Hieronder bespreken we verschillende vragen die je jezelf en andere leden van jouw onderzoeksproject kunt stellen om verder na te kunnen denken over de manier waarop jullie je verhouden tot het onderzoek.

8.5 REFLEXIVITEIT

Hoewel je als onderzoeker streeft naar objectiviteit, is het onmogelijk om volledig neutraal te zijn. Jouw blik op de wereld wordt gekleurd door jouw kennis en ervaring, en die zijn per definitie niet neutraal. Dat maakt het belangrijk om je bewust te zijn van je identiteit en hoe die zich verhoudt tot jouw onderzoek. Hiervoor zijn twee kernbegrippen geïntroduceerd in de literatuur (zie Wilson et al. (2022) en Jamieson et al. (2023) voor een overzicht):

- **Positionaliteit** gaat om het expliciet maken van je identiteit en je houding en aannames ten opzichte van het onderwerp dat je onderzoekt.
- **Reflexiviteit** gaat om het kritisch bevragen van jouw positionaliteit en de effecten die jouw positionaliteit kan hebben op het onderzoek. Daarbij kijk je ook naar strategieën om hiermee om te gaan, zodat je eventuele negatieve effecten kunt ondervangen.

Deze begrippen vinden hun oorsprong in kwalitatief onderzoek, maar worden ook steeds gangbaarder in kwantitatieve studies. Jamieson et al. (2023) hebben een lijst met vragen opgesteld die je kunnen helpen om verder na te denken over de manier waarop je je tot je onderzoek verhoudt. Deze vragen vormen eigenlijk een verdere uitwerking van de verantwoordelijkheid die we als wetenschapper hebben naar de maatschappij in het algemeen, en de groepen die we bestuderen in het bijzonder. Hieronder bespreken we de vragen per onderzoeksfase.

8.5.1 HET ONTWERP VAN JE STUDIE

In Hoofdstuk 2 hebben we gekeken naar het kiezen van een onderwerp en een onderzoeksvraag. Daarbij lag de nadruk op de wetenschappelijke en maatschappelijke relevantie van het onderzoek. Jamieson et al. (2023) richten zich meer op de manier waarop jij je als onderzoeker tot de doelgroep verhoudt. Hieronder bespreken we een selectie van drie vragen.[13]

13 De volgende opsomming is vrij vertaald naar de originele vragen van Jamieson et al. (2023, Tabel 1).

Waarom wil ik deze groep [mensen] onderzoeken?
De eerste vraag dient om je beweegredenen helder te krijgen: wat is de *persoonlijke* relevantie van het onderzoek voor jou? Dat is niet iets wat je in een verslag hoeft te zetten, maar het helpt wel om je bewuster te worden van jouw perspectief op het onderzoek, en de mogelijke vooroordelen die dat perspectief met zich meebrengt.

In hoeverre ben ik onderdeel van de groep die ik wil bestuderen? Ben ik een 'insider' of een buitenstaander (of allebei)?
In Hoofdstuk 2 hebben we al kort gekeken naar de rol van ervaringsdeskundigen, en geconstateerd dat het kan helpen om bij leden van de relevante gemeenschap te controleren of er eigenlijk behoefte is aan jouw onderzoek, en of het perspectief dat je hanteert goed aansluit bij de manier waarop zij hun situatie ervaren. Als je zelf ervaringsdeskundige bent, is het belangrijk om te controleren of jouw ervaringen overeenkomen met die van anderen.[14] Er zijn verschillende manieren om ervaringsdeskundigen te betrekken bij je project, variërend van een eenmalig contact tot een langdurige samenwerking waarbij de ervaringsdeskundige ook co-auteur wordt van het uiteindelijke verslag. Jamieson et al. (2023) citeren verder het werk van Chavez (2008), waarin wordt beargumenteerd dat het onderscheid tussen *insiders* en *outsiders* niet altijd goed te maken is; Chavez (2008) spreekt daarom van een spectrum van betrokkenheid waarbij onderzoekers op verschillende manieren wel of niet betrokken kunnen zijn.

Wat kan ik teruggeven aan deze groep?
Kom je alleen iets halen, of kom je ook iets brengen? De verhouding tussen onderzoekers en de groepen die ze bestuderen is niet altijd even eerlijk: onderzoekers krijgen onderzoeksdata en (hopelijk) waardering voor hun publicaties, maar wat krijgen de mensen die onderzocht worden ervoor terug? En hebben ze daar dan ook wat aan?[15]

Om ervoor te zorgen dat de gemeenschap die je onderzoekt ook echt iets heeft aan jouw onderzoek kun je, als je zelf geen onderdeel bent van die gemeenschap, overwegen om samen te werken met mensen uit die gemeenschap.[16] Later in dit hoofdstuk komen we nog eens terug op dit onderwerp, wanneer we kijken naar de verschillende manieren waarop je iets terug zou kunnen doen voor online gemeenschappen bij het bestuderen van sociale media.

14 Een algemene valkuil bij het betrekken van ervaringsdeskundigen is dat het verleidelijk is om te generaliseren vanuit de ervaringen van een beperkt aantal mensen. Wees dus bedacht op mogelijke verschillen tussen verschillende leden van dezelfde gemeenschap.

15 Sommige onderzoekers, zoals Steven Bird (Bird, 2020; Bird & Yibarbuk, 2024), leggen hier een expliciet verband met het kolonialisme: in plaats van fysieke grondstoffen worden er nu gegevens geëxtraheerd uit verschillende gemeenschappen, zonder dat die gemeenschappen daar zeggenschap over hebben.

16 In sommige situaties kun je niet echt spreken van een gemeenschap, maar eerder van individuele ervaringsdeskundigen die je zou kunnen raadplegen.

8.5.2 DATAVERZAMELING

Wat betreft participanten is inhoudsanalyse een vreemde eend in de bijt. Jamieson en collega's bespreken bij de dataverzameling verschillende vragen over de relatie die je als onderzoeker hebt met je participanten, maar bij inhoudsanalyse werken we meestal met bestaand materiaal, en blijven de producenten van dat materiaal veelal buiten beeld. Als je niet echt contact hebt met de auteurs die de boodschappen genereren die je analyseert, is de relatie tussen onderzoekers en auteurs dan irrelevant? Misschien niet helemaal; in ieder geval is het een *keuze* om wel of niet contact op te nemen met de auteurs om ze op de hoogte te stellen van je onderzoek.[17]

De kern van reflexiviteit, volgens Lazard & McAvoy (2017 instemmend geciteerd door Jamieson et al.) is de vraag hoe het onderzoeksproces er precies uitziet, en hoe jouw positionaliteit dat proces beïnvloedt. Een belangrijke kwestie is ook hoe je precies bepaalt waar je de data vandaan haalt. Vaak kun je geen census nemen, en moet je toch kijken naar een specifiek deel van de populatie. Intuïtief weet je misschien al wel waar je relevante data zou kunnen vinden die ondersteuning zou kunnen geven voor jouw hypothesen. Maar ben je daarmee niet bevooroordeeld? Jamieson et al. suggereren dat preregistratie hier misschien zou kunnen helpen; door een onderzoeksvoorstel met een expliciete motivatie te formuleren, dwing je jezelf om te reflecteren op de waarom-vraag. (De auteurs benadrukken wel expliciet dat het uitvoeren van een preregistratie op zich niet betekent dat je reflexief bezig bent, maar dat het reflexieve proces er wel mee ondersteund kan worden.)

8.5.3 DATA-ANALYSE EN INTERPRETATIE

Hoe beïnvloeden onze ideeën en verwachtingen de manier waarop we data analyseren en de resultaten van die analyses interpreteren? Jamieson et al. (2023) citeren Lehner et al. (2008) die laten zien dat *confirmation bias* een grote rol speelt in de manier waarop mensen omgaan met tegenstrijdige aanwijzingen. Stel je voor dat je twee analyses uitvoert op een corpus, en de ene analyse ondersteunt jouw hypothese, en de andere analyse wijst de andere kant uit. Volgens Lehner et al. zijn mensen geneigd om meer waarde te hechten aan bewijs dat hun eigen zienswijze ondersteunt. Deze observatie geldt natuurlijk niet alleen voor analyses als geheel, maar ook voor de verschillende keuzes die je maakt tijdens het onderzoeksproces. Daarom is het goed, volgens Jamieson en collega's, om een uitgebreid logboek bij te houden waarbij je doorlopend opschrijft waarom je bepaalde keuzes maakt. Die keuzes kun je vervolgens ook bespreken met collega's om samen te verkennen welke andere keuzes er waren, en wat men vindt van de gemaakte keuzes.

17 Een alternatief kan misschien ook zijn om een representatieve sample van doelgroep te vragen wat ze vinden van jouw onderzoek, zodat je in ieder geval een beeld hebt van de mening van de doelgroep.

8.5.4 CONCLUSIE EN FRAMING

Jamieson et al. (2023) beargumenteren dat de manier waarop resultaten geïnterpreteerd worden, de conclusies die we daaruit trekken, en de manier waarop die conclusies gepresenteerd worden allemaal beïnvloed worden door onze ervaringen en ons wereldbeeld. Een goede manier om hiermee om te gaan is om te proberen onze positie expliciet bespreekbaar te maken. Jamieson et al. stellen hierbij de volgende discussievragen voor:[18]

- Is de manier waarop je met de verschillende resultaten omgaat beïnvloed door jouw positionaliteit? (Dus door jouw eigen gevoelens, behoeften, houding, of ervaring?)
- Wat levert dit onderzoek voor je op? En wat levert het op voor de doelgroep die je bestudeert? Zit er een spanning tussen die twee?

Deze vragen kun je in ieder geval samen met jouw mede-onderzoekers bespreken. Jamieson et al. erkennen dat het misschien wat ongemakkelijk is om al die overwegingen publiek te delen,[19] en stellen als alternatief voor dat je het onderzoek ook kan laten beoordelen door anderen (die misschien ook anders in het leven staan).

Een laatste punt dat Jamieson et al. (2023) aandragen gaat over de studies waar je naar verwijst. In de alinea over Eerlijkheid in dit hoofdstuk hebben we al besproken dat het belangrijk is dat je zowel verwijst naar studies die jouw argumenten ondersteunen als naar studies die juist tegenbewijs leveren. Eerder onderzoek door Harper (2020) suggereert dat dit niet altijd gebeurt. Jamieson et al. (2023) verwijzen ook naar studies van Fulvio et al. (2021) en Bertolero et al. (2020) die laten zien dat er disproportioneel vaak naar mannelijke en witte auteurs verwezen wordt (ten opzichte van vrouwelijke auteurs en mensen uit etnische minderheidsgroepen). Het is belangrijk om waakzaam te zijn bij het zoeken naar referenties, zodat je deze ongelijkheden niet versterkt.[20]

18 Deze vragen zijn wederom vertaald en (deels) geparafraseerd vanuit Tabel 1 van Jamieson et al. (2023, p. 10).

19 De auteurs geven zelf niet aan waarom dat ongemakkelijk zou zijn, maar je kunt je voorstellen dat het misschien niet prettig voelt om zulke persoonlijke informatie openbaar te delen. Daarnaast kan het delen van je persoonlijke opvattingen misschien ook averechts werken: hoewel reflexiviteit bedoeld is om na te denken over de manier waarop jouw positionaliteit het onderzoek zou kunnen beïnvloeden en ongewenste invloeden tegen te gaan, kan het delen van jouw persoonlijke opvattingen ook de suggestie van vooringenomenheid wekken.

20 Dit kan bijvoorbeeld gebeuren wanneer je alleen maar verwijst naar referenties die je hebt gevonden in andere artikelen. Op deze manier verarmt de literatuur (omdat er geen nieuwe verwijzingen meer worden aangedragen) en worden bestaande ongelijkheden vergroot (want je citeert alleen mensen die toch al geciteerd werden – dit wordt ook wel het Mattheuseffect genoemd; Merton (1968).

8.6 INHOUDSANALYSE EN IDENTITEIT

Inhoudsanalyses zijn uitermate geschikt om ongelijkheid in de maatschappij te bestuderen, bijvoorbeeld door te kijken hoe verschillende sociale groepen worden gepresenteerd of gerepresenteerd in de media. Of het nu gaat om de verhouding tussen mannen en vrouwen, seksuele geaardheid, of culturele minderheden, identiteit is een regelmatig terugkerend onderwerp in het maatschappelijke debat, en daardoor is dit onderwerp ook erg populair om te onderzoeken. Tegelijkertijd is identiteit ook een erg gevoelig onderwerp, omdat het gaat om de kern van ons bestaan. Wanneer je een onderzoek ziet waarin mensen ingedeeld worden in verschillende groepen, is het goed om extra alert te zijn en je af te vragen:

- Waar komt dit onderscheid vandaan?
- Op basis waarvan wordt dit onderscheid gemaakt?
- Door wie wordt het onderscheid gemaakt? (Gaat het om zelf-gerapporteerde eigenschappen, of wordt het onderscheid door een externe partij gemaakt?)
- Worden bepaalde groepen uitgesloten door mensen op deze manier te onderscheiden van elkaar?

Voordat je zelf zomaar mensen indeelt in verschillende groepen, is het belangrijk om je te verdiepen in het thema dat je bestudeert. Hieronder bespreken we twee thema's: (1) ras, huidskleur, en etniciteit, en (2) geslacht en gender. Dit zijn natuurlijk niet de enige eigenschappen die bepalend zijn voor je identiteit; seksuele geaardheid, religie, en politieke oriëntatie zijn bijvoorbeeld ook belangrijke factoren. Maar de genoemde thema's zijn duidelijke voorbeelden waarbij vaak gedacht wordt dat je iemands identiteit aan hun uiterlijk kunt aflezen —dat is niet het geval. Daarbij geven we meteen de *disclaimer* dat het vrijwel onmogelijk is om recht te doen aan de complexiteit van deze thema's in de beperkte ruimte die er is in dit werkboek. Zie de tekst hieronder dan ook vooral als een aanmoediging om de literatuur verder te verkennen.

8.6.1 RAS, HUIDSKLEUR, EN ETNICITEIT

Ras, huidskleur, en etniciteit zijn drie begrippen die gebruikt worden om verschillende bevolkingsgroepen van elkaar te onderscheiden. Van deze begrippen is huidskleur het meest zichtbaar: het lijkt bijvoorbeeld eenvoudig om Zwarte en Witte mensen van elkaar te onderscheiden, om te bestuderen hoe mensen van kleur gerepresenteerd worden in de media. Maar hoewel huidskleur aantoonbaar effect heeft op de manier waarop mensen elkaar behandelen (Devaraj et al., 2018; Hunter, 2007), is het belangrijk om complexe verschillen tussen bevolkingsgroepen niet zomaar te reduceren tot één factor.[21] Het is bijvoorbeeld niet altijd aan de

21 Los daarvan zijn er vaak ook verschillen binnen bevolkingsgroepen zelf. Die verschillen *binnen* bevolkingsgroepen kunnen soms groter zijn dan verschillen *tussen* bevolkingsgroepen.

buitenkant te zien of iemand deel uitmaakt van de Zwarte gemeenschap, of zich identificeert als Zwart. Identiteit is *cultureel bepaald* (Templeton, 2013; Ware et al., 2020).

In een toegankelijk overzichtsartikel bespreken Sen & Wasow (2016) de uitdagingen bij het bestuderen en operationaliseren van ras/etniciteit. Zij beargumenteren dat je vaak beter kunt kijken naar de losse componenten waaruit het complexe construct "ras/etniciteit" bestaat (onder andere: huidskleur, dialect, religie, voor/achternaam, en de buurt waar iemand vandaan komt), maar waarschuwen tegelijkertijd dat geen twee operationalisaties hetzelfde meten. Wanneer je een inhoudsanalyse wil uitvoeren om verschillende bevolkingsgroepen met elkaar te vergelijken, kun je het beste voorzichtig zijn met de claims die je maakt. Bestudeer je bijvoorbeeld de representatie van *moslims* in de media, of eigenlijk alleen maar de representatie van *mensen met een hoofddoekje*? En als je weet dat iemand onderdeel is van een bepaalde bevolkingsgroep, dan wil dat niet zeggen dat die persoon in alle situaties herkenbaar is als lid van die bevolkingsgroep.

8.6.2 GESLACHT EN GENDER

De termen *geslacht*, *sekse*, en *gender* worden soms door elkaar gebruikt om te verwijzen naar het onderscheid tussen mannen en vrouwen. Het verschil tussen deze drie termen is dat de eerste twee verwijzen naar puur biologische eigenschappen, terwijl de laatste term verwijst naar een combinatie van psychologische, maatschappelijke, en gedragsmatige eigenschappen die vaak geassocieerd worden met mannelijkheid of vrouwelijkheid. *Genderidentiteit* verwijst naar de manier waarop je jezelf ziet in termen van die eigenschappen: zie je jezelf als mannelijk, vrouwelijk, een combinatie van die twee, of juist als iemand die buiten die kaders valt? Gender en genderidentiteit zijn geen zichtbare eigenschappen. Wat je ziet is het gevolg van *genderexpressie*: gedragingen en uiterlijke kenmerken die overeenkomen met bepaalde *genderrollen*. Dat zijn houdingen en gedragingen die binnen de maatschappij als passend gezien worden bij een bepaald gender. De manier waarop mensen hun gender wel of niet tot uiting laten komen is afhankelijk van de situatie waarin ze zich bevinden.

Er is een rijke traditie van inhoudsanalyses om genderrollen te bestuderen. De eerste studies over verschillen in de weergave van mannen en vrouwen in de media verschenen al in de jaren '50 van de vorige eeuw (Saenger, 1955; Spiegelman et al., 1953). Rudy et al. (2010) onderscheiden vier verschillende doelstellingen die we kunnen herkennen in de literatuur:[22]

1) Om feministische claims over gender-gebaseerde ongelijkheden in de maatschappij te ondersteunen. Wanneer je zulke ongelijkheden aan kunt tonen, heb je een sterkere positie in discussies om ongelijkheid aan te pakken. Een voorbeeld is het onderzoek van Courtney & Whipple (1983) dat laat zien dat advertenties in de Verenigde Staten een ste-

22 De eerste zin in deze opsomming is telkens een vertaling van het oorspronkelijke artikel. Daarna volgt een verdere duiding aan de hand van voorbeelden die door Rudy et al. (2010) worden gegeven. Voor meer voorbeelden kun je het overzicht van Rudy et al. (2010, 2011) raadplegen.

reotype beeld schetsen van de maatschappij, waarin de vrouw ondergeschikt is aan de man.

2) Om de overeenkomst (of het gebrek daaraan) tussen de werkelijkheid en weergaven in de media te onderzoeken. Rudy et al. (2010) noemen bijvoorbeeld het onderzoek van Fouts & Burggraf (2000) waarin gekeken wordt naar de representatie van vrouwen in *sitcoms*. Zij laten zien dat vrouwen op televisie disproportioneel vaak dunner zijn dan gemiddeld.[23]
3) Om een empirische basis te geven aan theorie en onderzoek waarin gekeken wordt welke effecten verschillende berichten hebben op (verschillende groepen binnen) het publiek. Wat er daarbij gebeurt is dat onderzoekers bijvoorbeeld kijken naar de verschillende eigenschappen van reclames of televisieprogramma's, en vervolgens aan de hand van verschillende theorieën over media-effecten (bijvoorbeeld *social learning theory*; Bandura (1977)) voorspellen hoe die programma's het publiek zouden kunnen beïnvloeden.
4) Om een empirische basis te geven aan theorie en onderzoek waarin gekeken wordt welke effecten verschillende media-producenten hebben op de inhoud van de berichten. Met andere woorden: wat kunnen we leren over zenders op basis van de boodschappen die ze verspreiden? Dat kan gaan over culturele verschillen in de manier waarop mannen en vrouwen worden afgebeeld in advertenties (Gilly, 1988), maar ook over de bronnen die gebruikt worden door mannelijke en vrouwelijke journalisten (Armstrong, 2004).

Er is dus genoeg werk om je bij aan te sluiten en de genoemde doelstellingen verder te verkennen, maar enige voorzichtigheid is wel geboden. Larson (2017) geeft daarbij de volgende adviezen:

1) Zorg voor een duidelijke definitie van gender, die aansluit op hetgeen je wil onderzoeken. Maak voor jezelf en de lezers duidelijk wat je precies denkt te meten met deze variabele. Gaat het bijvoorbeeld om de manier waarop verschillende genderrollen worden gepresenteerd, of gaat het om de manier waarop auteurs zich identificeren?
2) Zorg ervoor dat je een goede reden hebt om gender te bestuderen. Waarom denk je bijvoorbeeld dat er een verschil zou kunnen zitten in berichten die geschreven worden door mannen of vrouwen? Ligt dat echt aan hun gender, of is er een andere variabele die dat verschil zou kunnen verklaren?[24]
3) Zorg ervoor dat het duidelijk is hoe je tot de categorisatie gekomen bent. Hoe je berichten, zenders, of ontvangers zou moeten categoriseren is afhankelijk van je onderzoeksvraag. Als je graag de rol van genderidentiteit wil onderzoeken, dan is het vaak verstandig om mensen te vragen naar hun identiteit. (Die is immers niet zichtbaar!) Maar als het gaat

23 Daarnaast tonen Fouts & Burggraf (2000) aan dat er vaker grappen worden gemaakt over het gewicht van vrouwen die dikker zijn dan gemiddeld, waarover ook weer vaker gelachen werd door het publiek.

24 Larson gebruikt zelfs fermere bewoordingen: vermijd het gebruik van gender als variabele tenzij het echt noodzakelijk is (en leg daarbij goed uit waarom dat zo is).

om de manier waarop verschillende genderrollen gepresenteerd worden, dan is dat misschien niet nodig.

Boven alles is het essentieel om respectvol om te blijven gaan met de belanghebbenden in jouw onderzoek. Denk bijvoorbeeld na over de gevolgen van de operationalisatie van gender: niemand vindt het fijn om weggezet te worden als "anders."[25]

8.7 SAMENWERKING MET EXTERNE ORGANISATIES

Samenwerkingen met externe organisaties kunnen zeer waardevol zijn. Als onderzoeker krijg je toegang tot informatie en andere middelen die je misschien anders nooit zou kunnen krijgen, en je kunt via samenwerkingen met praktijkpartners ook een positieve bijdrage leveren aan de maatschappij. Tegelijkertijd zijn er veel verschillende ethische vragen rondom de samenwerking met externe organisaties. Hierboven heb je al gelezen over de conflicterende belangen die een rol kunnen spelen bij zulke samenwerkingen, waardoor jouw onafhankelijkheid als onderzoeker in het geding zou kunnen komen, of waardoor je worstelt met vragen over jouw verantwoordelijkheid als onderzoeker tegenover de maatschappij. Daarnaast kun je je zorgen maken over *dual use*: hoe worden jouw bevindingen straks gebruikt? En wat zijn eigenlijk de neveneffecten van samenwerkingen met externe organisaties?

De laatste jaren is er ook steeds meer aandacht voor de manier waarop organisaties oppervlakkige groene maatregelen nemen om zich milieubewuster te doen lijken dan ze eigenlijk zijn. Dit wordt *greenwashing* genoemd (zie Freitas Netto et al. (2020; Santos et al., 2024) voor een overzicht). Samenwerkingen met bedrijven rondom duurzaamheidsthema's kunnen ook voor dit doeleinde worden ingezet. Daarnaast bestaan er ook meer structurele zorgen over de manier waarop samenwerkingen met bedrijven bepaalde *frames* in stand kunnen houden. Van Hintum (2024) citeert bijvoorbeeld ethicus Dirk Hilbers die aangeeft dat…

> *"…de energietransitie volledig wordt ingericht rondom technologische innovatie die de huidige energieconsumptie in stand houdt. Dit paradigma legitimeert enerzijds de samenwerking met de fossiele industrie, en anderzijds de gedachte dat de energietransitie nu eenmaal niet sneller en eerlijker kan gaan dan ons technologisch innoverend vermogen. Het is inmiddels pijnlijk duidelijk dat dit niet snel genoeg, en al helemaal niet rechtvaardig genoeg is."*
>
> (Van Hintum, 2024, p. 9)

25 Een nuttige bron om meer te leren over gender-inclusief taalgebruik en het hanteren van verschillende persoonlijke voornaamwoorden (Engels: *pronouns*) is: https://pronouns.org.

Zo bezien kunnen samenwerkingen met bedrijven (in dit geval: de grote spelers in de fossiele industrie) aandacht voor andere oplossingen (zoals het grondig verminderen van ons energieverbruik, in plaats van het vergroenen van energieproductie) in de weg kan staan.

Naast greenwashing zijn er ook nog talloze andere soorten van *–washing*, zoals *rainbow washing* (aandacht besteden aan de LGBTQ+-gemeenschap in marketingboodschappen om betrokkenheid te veinzen; zie Tiziana Schopper & Vogelgsang (2024)), of *ethics washing* (voornamelijk *big tech*-bedrijven die ethici aanstellen om te doen alsof ze verantwoord bezig zijn; zie Van Maanen (2020)). Het patroon is duidelijk: veel organisaties willen zich graag profileren als maatschappelijk betrokken, maar die betrokkenheid wordt niet altijd even oprecht beleden. Het lijkt daarmee verstandig om bij samenwerkingen niet alleen na te denken over de manier waarop de onderzoeks*resultaten* gebruikt worden (zoals gebruikelijk bij dual use-kwesties), maar om ook na te denken over de manier waarop *de samenwerking zelf* gebruikt zou kunnen worden.

8.8 SOCIALE MEDIA BESTUDEREN

Sociale media vormen een goudmijn voor communicatie-onderzoekers: nooit eerder hadden we toegang tot zoveel data van zoveel verschillende gemeenschappen die over zoveel verschillende onderwerpen met elkaar in gesprek gingen. Bovendien gaat het om *natuurlijke* data, die zonder inmenging van onderzoekers tot stand is gekomen. Dat alles leidt ertoe dat we beter kunnen begrijpen hoe verschillende groepen mensen met elkaar omgaan. Tegelijkertijd is het bestuderen van sociale media niet zonder risico's. Hieronder bespreken we verschillende ethische overwegingen omtrent het bestuderen van sociale media.[26]

8.8.1 PRIVACY

Onderzoek naar communicatie via sociale media leidt vaak tot discussies over privacy. Jarenlang was de norm dat het geen probleem is om sociale media te bestuderen omdat alle gegevens openbaar zijn, maar inmiddels is de consensus hierover veranderd. Hiervoor zijn drie verschillende redenen aan te wijzen.

Reden 1: verschillende schandalen rondom het gebruik van data uit sociale media. Er zijn verschillende schandalen geweest rondom onderzoekers en bedrijven die sociale media bestuderen. Hieronder bespreken we twee voorbeelden.

In 2016 publiceerde een Deense student een grote dataset met informatie uit datingprofielen, die hij automatisch had verzameld van de Amerikaanse website OKCupid. Deze informatie was enkel beschikbaar nadat hij ingelogd was, maar in principe kon iedereen een profiel aanmaken en dezelfde informatie te zien krijgen. Het grote verschil is dat de publica-

26 Een eerder overzicht van ethische overwegingen wordt gegeven door Zook et al. (2017).

tie van de dataset ervoor zorgde dat alle informatie opeens *doorzoekbaar* was op een manier die voorheen nog niet mogelijk was. Zo werden de seksuele voorkeuren en andere intieme gegevens van de gebruikers van OKCupid ineens openbaar gemaakt, met alle gevolgen van dien (McCook, 2016). Dit voorbeeld laat zien dat het verzamelen en publiceren van data een transformatieve handeling is, die beperkingen op de doorzoekbaarheid kan wegnemen, extra aandacht kan vestigen op bestaande data, en gegevens verder in de openbaarheid kan brengen.

In 2018 werd bekend dat het Britse bedrijf Cambridge Analytica data had verzameld via sociale media om een database aan te leggen met de persoonlijke gegevens van de gebruikers van Facebook. De gegevens werden verzameld via een app ("*This is your digital life*") waarbij gebruikers een 'persoonlijkheidsquiz' konden invullen, en de resultaten konden delen met hun vrienden. De app was oorspronkelijk gemaakt door een wetenschapper aan de Universiteit van Cambridge. De database van Cambridge Analytica werd vervolgens gebruikt om de Amerikaanse verkiezingen te manipuleren door gerichte advertenties te sturen naar verschillende gebruikers, op basis van hun politieke voorkeuren (Berghel, 2018). Dit voorbeeld laat zien dat gegevens die oorspronkelijk voor onderzoeksdoeleinden bedoeld zijn, ook voor andere, immorele toepassingen gebruikt kunnen worden. Later onderzoek heeft laten zien dat gebruikers veel moeite hebben om de verschillende privacy-instellingen op sociale media te begrijpen Hinds et al. (2020).

Reden 2: nieuwe wetgeving. De Algemene Verordening Gegevensbescherming (AVG) is sinds 25 mei 2018 van toepassing. Deze wet staat in het Engels ook wel bekend als de General Data Protection Regulation (GDPR), en schrijft voor hoe organisaties binnen de maatschappij om dienen te gaan met persoonsgegevens (zie voor details: Autoriteit Persoonsgegevens (n.d.; Voigt & Bussche, 2017)). Voor onderzoekers betekent deze wetgeving dat ze goed moeten controleren of er bij het opslaan van data wordt voldaan aan al deze richtlijnen.

Reden 3: veranderende opvattingen over privacy. Onze opvattingen over privacy zijn door de tijd veranderd van een individuele keuze, waarbij je zelf controle hebt over al je gegevens, naar een meer complexe opvatting waarbij mensen afhankelijk zijn van het platform dat ze gebruiken en de normen die gelden binnen de gemeenschap waarbinnen ze verkeren (Trepte, 2020). Het is daarom ook erg lastig om te voorkomen dat anderen informatie over jou delen op sociale media. Als je bijvoorbeeld jouw geboortedatum verborgen houdt op sociale media, maar anderen feliciteren jou met je verjaardag, dan is het nog steeds eenvoudig om te achterhalen wanneer je jarig bent.[27]

27 Hoeveel mensen over jou te weten kunnen komen via sociale media valt goed te illustreren met het fenomeen "Consensual Doxxing." Doxxing is het ongevraagd openbaar maken van persoonlijke informatie, vaak met het doel om iemand te intimideren. Maar influencer Kristen Sotakoun (@notkahnjunior

Privacy is in tijden van sociale media een constante onderhandeling tussen jouw eigen voorkeuren, de functionaliteit van diverse apps en websites, en onderlinge normen (die ook nog eens kunnen verschillen tussen de verschillende groepen waar je deel van uitmaakt). Dit idee wordt *networked privacy* genoemd (Marwick & boyd, 2014). Theorieën over *networked privacy* bouwen voort op het idee dat privacy contextueel is. Marwick & boyd (2014) verwijzen instemmend naar Nissenbaum (2009) die aangeeft dat de normen omtrent het delen van informatie afhangen van de aard van de informatie, de verhoudingen tussen zender, onderwerp, en ontvanger, en de manier waarop de boodschap gestuurd wordt. De auteurs vertellen bijvoorbeeld over jongeren die verontwaardigd zijn dat hun ouders reageren op hun status-updates wanneer die boodschap 'duidelijk niet voor hen bedoeld is.' Daarnaast is het bijvoorbeeld ook *not done* om te reageren op oudere status-updates; wanneer je een foto online zet is het prima als mensen daarop reageren, maar het wordt wel wat ongemakkelijk wanneer iemand opeens reageert op een bikinifoto van twee jaar geleden. De sociale verhoudingen zijn in dit voorbeeld misschien hetzelfde, maar de context is veranderd. Dit voorbeeld laat zien dat het niet alleen belangrijk is om te kijken naar de vraag of iets openbaar is, maar ook naar de vraag met welke *intentie* het bericht is geplaatst en voor wie het bericht is bedoeld.[28]

Onderzoek naar sociale media werkt per definitie decontextualiserend: je haalt berichten uit hun oorspronkelijke context, en bestudeert ze zonder bekend te zijn met de intentie van de auteur en de situatie waarin de berichten geplaatst zijn. Dat levert twee problemen op:

1) Als onderzoeker behoor je waarschijnlijk niet tot het beoogde publiek voor wie het bericht is geplaatst. Het is bovendien onduidelijk of gebruikers van sociale media redelijkerwijs hadden kunnen verwachten dat hun berichten gebruikt zouden kunnen worden voor onderzoeksdoeleinden. Een vragenlijststudie van Fiesler & Proferes (2018) onder Twitter-gebruikers liet zien dat 61% van hen zich niet bewust was dat hun openbare berichten gebruikt zouden kunnen worden voor wetenschappelijk onderzoek.
2) Doordat je berichten uit hun oorspronkelijke context haalt, wordt de kans groter dat de berichten verkeerd begrepen worden. Dezelfde studie van Fiesler & Proferes laat zien dat respondenten zich hierover zorgen maken.

Fiesler & Proferes bespreken ook verschillende factoren die ervoor zorgen dat mensen (of in ieder geval de respondenten uit hun studie) zich minder zorgen maken over het gebruik van berichten op sociale media. Bijvoorbeeld:

op TikTok) laat op verzoek (dus met instemming van de persoon om wie het gaat) zien hoe gemakkelijk het is om op basis van jouw account meer te leren over wie je bent. Via informatie die je zelf deelt, accounts met wie je bevriend bent, het gebruik van dezelfde gebruikersnaam op verschillende sociale media, en reacties van anderen op jouw posts laat ze zien dat bijna niemand echt anoniem is (Tolentino, 2022).

28 Als je meer wil lezen over privacy en sociale media, is het overzichtsartikel van Sabine Trepte (2020) een aanrader.

- Gebruikers maken zich minder zorgen over grote datasets waar veel verschillende berichten in zijn verzameld, dan over kleine datasets. Het idee hierachter is dat je dan meer opgaat in de massa dan als onderzoekers maar naar een kleine selectie kijken.
- Gebruikers maken zich minder zorgen om automatische analyses dan om onderzoekers die handmatig door de berichten heen gaan. Een verklaring hiervoor wordt niet gegeven in het artikel, maar algemeen lijkt het zo te zijn dat gebruikers het fijn vinden als onderzoekers een zekere mate van afstand bewaren, en dat het onderzoek niet te persoonlijk wordt.
- Gebruikers maken zich minder zorgen als hun berichten anoniem blijven. Dus in tegenstelling tot academische artikelen willen de auteurs van berichten op het toenmalige Twitter juist niet dat hun naam genoemd wordt. Dit hoofdstuk gaat verder niet in op praktische manieren om participanten te beschermen. Anonimisatie en pseudonimisatie bespreken we verder in *Hoofdstuk 9: Datamanagement*.

Op basis van deze informatie lijkt het verstandig om je onderzoek zo op te zetten dat de auteurs van berichten op sociale media zo anoniem mogelijk blijven, bijvoorbeeld door geen gebruikersnamen te verzamelen en de data zoveel mogelijk automatisch te analyseren (waarbij je dus in sommige gevallen de afweging moet maken tussen Validiteit en Privacy). Maar voordat je aan de slag gaat met de dataverzameling moet je eerst goed nadenken over toestemming.

8.8.2 TOESTEMMING

Toestemming is een fundamentele waarde binnen onze maatschappij. Voor wetenschappelijk onderzoek heb je meestal toestemming nodig van twee partijen:

1) Een ethische commissie die beoordeelt of jouw onderzoek op een verantwoordelijke manier omgaat met de participanten, en die afweegt of het doel van het onderzoek voldoende opweegt tegen eventuele nadelige gevolgen van het onderzoek.[29]
2) De deelnemers aan het onderzoek dienen ook in te stemmen met hun deelname, nadat ze voorgelicht zijn over de inhoud van de studie, eventuele nadelige gevolgen, en hun rechten als deelnemer (waaronder het recht om ten allen tijde hun toestemming in te trekken, zonder verdere nadelige gevolgen). Deze voorlichting maakt dat er sprake is van zogenaamde *informed consent*.

Hieronder bespreken we Toestemming vanuit drie verschillende perspectieven (de ethische commissie, gebruikers, en platforms) in de context van inhoudsanalyse.

29 Wanneer je een opdracht maakt in de context van een vak aan de universiteit, neemt de docent vaak de rol in van de ethische commissie. Zelf heb je natuurlijk ook de verantwoordelijkheid om na te denken over de ethische implicaties van je onderzoek, en dien je bij twijfel altijd aan de bel te trekken.

Toestemming van een ethische commissie

Zoals we al hebben besproken in de sectie over Privacy, zijn de normen omtrent het onderzoeken van sociale media sterk veranderd in het afgelopen decennium. Waar inhoudsanalyses eerst waren vrijgesteld van ethische goedkeuring, worden er nu steeds meer eisen gesteld waaraan onderzoekers moeten voldoen. Dat heeft deels te maken met de wettelijke plicht om persoonsgegevens te beschermen, maar het komt ook door voortschrijdend inzicht omtrent de implicaties van onderzoek naar sociale media. Toch laat een recente inhoudsanalyse van studies die data van Reddit gebruiken (Proferes et al., 2021) zien dat de meeste studies geen melding maken van ethische goedkeuring, of aangeven dat het onderzoek vrijgesteld is.

Toestemming van gebruikers

De vragenlijststudie onder (toenmalige) Twitter-gebruikers van Fiesler & Proferes (2018) heeft ook gekeken naar toestemming. Desgevraagd liet een ruime meerderheid (65%) weten dat onderzoekers volgens hen eerst toestemming zouden moeten vragen voordat ze data van sociale media gebruiken.[30] Een kleine meerderheid van de participanten (53%) gaf aan dat ze in dat geval wel akkoord zouden gaan, en een derde (33%) van de participanten zei dat ze dat wel zouden doen als er aan bepaalde voorwaarden voldaan werd (bijvoorbeeld: dat de berichten geanonimiseerd zouden worden).

Wanneer heb je toestemming nodig?

Op basis van de vragenlijst van Fiesler & Proferes lijkt het een goed idee om mensen toestemming te vragen voor het analyseren van hun berichten op sociale media. Het is in ieder geval raadzaam om dat te doen wanneer je een analyse uitvoert die gericht is op specifieke personen, of die sterk inbreuk maakt op de privacy van de mensen van wie je de berichten analyseert. Maar het is niet altijd noodzakelijk om toestemming te vragen. Hieronder volgt een korte reflectie.

Het publieke spectrum. Wanneer we kijken naar de verschillende soorten accounts die informatie delen op sociale media, lijkt er een spectrum te zijn van publieke personen en organisaties (de premier van Nederland, BN'ers, grote bedrijven zoals Schiphol en PostNL) naar particulieren zoals jij en ik. Publieke personen en organisaties zijn zich enorm bewust van hun openbare uitingen, en hebben vaak mensen of hele afdelingen in dienst die berichten plaatsen en het verkeer op sociale media in de gaten houden. Gezien hun mediabewustzijn en de invloed die zij hebben binnen de maatschappij, is het verantwoord (of zelfs noodzake-

30 Technisch gezien hebben gebruikers van sociale media vaak al toestemming gegeven voor het gebruik van hun data door andere partijen, via de gebruiksvoorwaarden waar je bij registratie verplicht mee moet instemmen. Maar de meeste mensen lezen die voorwaarden niet (dus ze zijn niet echt geinformeerd) en door het verplichtende karakter van de *Terms of service* kun je ook niet echt spreken van vrijwillige toestemming. (Als je de website wil gebruiken, dan kun je niet anders dan instemmen.)

lijk ter controle van de macht) voor wetenschappers om hun boodschappen te analyseren. Voor particulieren kunnen we eenzelfde argumentatie niet maken: zij zijn zich vaak minder bewust van de implicaties van hun online aanwezigheid, en hebben (in ieder geval op individueel niveau) ook minder invloed op de maatschappij. Dat maakt dat je goede redenen nodig zou hebben om hun berichten te bestuderen.

Tussen publiek en privé. Er zit nog veel ruimte tussen accounts van particulieren en die van publieke personen of grote bedrijven. Hoe zit het bijvoorbeeld met *influencers* en kleine organisaties zoals de bakker om de hoek? Zoals met zoveel ethische kwesties is dit een grijs gebied. We hebben meer context nodig om te beoordelen of het oké is om zonder toestemming data te verzamelen van accounts in het midden van het publieke spectrum. Daarbij kun je vragen stellen als:

- Gaat de analyse om specifieke accounts, of beoogt de studie te generaliseren over een grotere groep? Bij een grotere groep behoud je als onderzoeker wat meer afstand tussen jou en de persoon of organisatie wiens account je bestudeert, wat de studie ook eenvoudiger te verantwoorden maakt.
- Wat is het doel van het onderzoek? Een belangrijke rol van de wetenschap is om de impact van verschillende fenomenen op de maatschappij te onderzoeken. Daarom valt het bijvoorbeeld ook goed te verantwoorden wanneer je *influencer marketing* bestudeert bij accounts die gericht zijn op jonge kinderen. Misschien wordt de controle vanuit de wetenschap door de influencers zelf niet altijd gewaardeerd, maar het is belangrijk om te begrijpen hoe jonge kinderen beïnvloed worden op sociale media.
- Wat voor data ga je verzamelen, wat ga je er vervolgens mee doen, en welke informatie wil je bewaren? De antwoorden op deze vragen bepalen de mate van inbreuk die je maakt op de privacy van de gebruikers. Hoe verder die inbreuk gaat, des te meer heb je een goede verantwoording nodig voor het uitvoeren van het onderzoek zonder toestemming.

Als je geen toestemming kan of wil vragen

Het is niet altijd mogelijk om toestemming te vragen. Hieronder bespreken we enkele problemen en mogelijke alternatieven voor het vragen om toestemming aan alle gebruikers. We beginnen met de bezwaren:

- Het is niet altijd mogelijk om (discreet) om toestemming te vragen. Op verschillende platforms kun je geen privéberichten sturen zonder eerst bevriend met elkaar te zijn. Het alternatief is om in het openbaar om toestemming te vragen, maar daarmee de-anonimiseer je gebruikers (Tatman, 2018).
- Afhankelijk van de omvang van je dataset is het misschien ook niet haalbaar om iedereen om toestemming te vragen; het zou simpelweg te veel tijd kosten en ontzettend ingewikkeld zijn om de administratie bij te houden voor (potentieel) duizenden accounts.
- Vragen om toestemming kan er ook voor zorgen dat er een *sampling bias* kan optreden in je dataset. Als bepaalde groepen van gebruikers stelselmatig weigeren om hun data te

laten analyseren, dan is de dataverzameling niet meer representatief voor de gehele populatie van socialemediagebruikers.
- Het is ook niet altijd duidelijk wie je precies om toestemming moet vragen. Hoe ga je bijvoorbeeld om met *reposts* en *stitches* waarbij berichten van verschillende auteurs worden gecombineerd?

Bovengenoemde bezwaren kunnen ertoe leiden dat je geen toestemming kan of wil vragen aan de gebruikers zelf. Hoe moeten we dan verder? Allereerst is het belangrijk om zelf na te denken over de spanning die er ontstaat tussen de verschillende waarden die we besproken hebben in dit boek. Weegt het gebrek aan consent en de zwaarte van de inbreuk op de privacy van de gebruikers op tegen de baten van het onderzoek? Dat is een abstracte vraag, die je concreter kunt maken door bijvoorbeeld een opsomming te maken van de informatie die je over de gebruikers verzamelt. Laten we kijken naar een voorbeeld.

Voorbeeld: misinformatie op sociale media. Als je misinformatie en complottheorieën wil bestuderen op sociale media, is het onwaarschijnlijk dat de verspreiders van die berichten akkoord gaan met wetenschappers die het probleem in kaart willen brengen. Tegelijkertijd is er een groot maatschappelijk belang, vanwege de schade die misinformatie en complottheorieën kunnen aanbrengen aan onze maatschappij en het democratische bestel. Voor het onderzoek heb je voornamelijk de inhoud van de berichten nodig, en wellicht metadata zoals het aantal 'likes' en reacties. Hoewel er persoonlijke informatie kan staan in de berichten die je verzamelt (bijvoorbeeld in de discussie over vaccinaties),[31] kun je voorzorgsmaatregelen treffen om de berichten te anonimiseren en zo de potentiële inbreuk op de privacy te minimaliseren. Hier zou je waarschijnlijk wel onderzoek naar mogen doen, na goedkeuring door een ethische commissie.

In het bovenstaande voorbeeld zijn er eigenlijk geen alternatieven voor het vragen van toestemming aan alle gebruikers, maar dat is lang niet altijd zo. Zoals we ook eerder besproken hebben in Hoofdstuk 2, is het bij onderzoek naar een specifieke doelgroep goed om een aantal ervaringsdeskundigen naar hun mening te vragen. (Zeker als je zelf geen onderdeel bent van de doelgroep.) Je kunt dan jouw onderzoeksvoorstel voorleggen aan een panel van ervaringsdeskundigen, om te kijken hoe je zo veel mogelijk rekening kunt houden met de behoeften en mogelijke bezwaren van de doelgroep. In het geval van online gemeenschappen, zoals Reddit, kun je mogelijk ook de moderatoren van de relevante gemeenschap om advies vragen.

31 Natuurlijk is er hier een kans dat die 'persoonlijke' informatie is gefingeerd, maar we kunnen hier niet bij voorbaat vanuit gaan, en het valt lastig te controleren wat er precies waar of onwaar is aan berichten op sociale media. Daarom moeten we werken onder de aanname dat alle persoonlijk ogende informatie ook daadwerkelijk persoonlijk is.

Toestemming van platforms

Wanneer je berichten wil analyseren van sociale media, ben je afhankelijk van de platforms waar die berichten online worden gezet. Het is niet altijd duidelijk of je zomaar berichten van die platforms mag verzamelen en opslaan in een eigen corpus of dataset. Verschillende websites hebben verschillende regels omtrent dataverzameling, en landen verschillen ook in de wetgeving die van toepassing is op het (al dan niet geautomatiseerd) verzamelen van online gegevens. Laten we beginnen met de websites zelf.

Gebruiksvoorwaarden en API's. Op bijna alle grote website staan wel ergens gebruiksvoorwaarden (Engels: *Terms of service*) vermeld. Deze voorwaarden beschrijven wat je wel en niet mag doen met de gegevens die op de website staan. Soms bieden websites ook een API (*Application Programming Interface*) aan waarmee je die gegevens ook automatisch kunt verzamelen. Wanneer dat zo is, kun je je vaak registreren en toegang tot de API aanvragen. Daarna kun je aan de slag om de gewenste data te verzamelen. Vaak worden er wel grenzen gesteld aan de hoeveelheid data die je mag verzamelen, en wat je ermee mag doen. Dit is het ideale scenario, omdat er een voorgeschreven manier is om data te verzamelen, waarvan je weet dat het in orde is.

Naast het ideale scenario zijn er natuurlijk ook talloze andere situaties waar bijvoorbeeld het automatisch verzamelen van gegevens niet toegestaan is maar handmatig verzamelen wel, of waar überhaupt geen dataverzameling wordt toegestaan door de gebruiksvoorwaarden. Fiesler et al. (2020) bespreken verschillen en overeenkomsten tussen de gebruiksvoorwaarden op verschillende websites, alsook de implicaties die de *Terms of Service* hebben voor het doen van wetenschappelijk onderzoek.[32] Kort samengevat: het is ingewikkeld. Zoals vrijwel alle ethische vragen hangt het af van de precieze context: wat je wil verzamelen, hoeveel je wil verzamelen, en met welk doel je dat dan doet. Wetenschappelijk onderzoek heeft samen met de journalistiek een bijzondere status in onze maatschappij als controleurs van machtige organisaties en overheden, en om kritisch te blijven op maatschappelijke ontwikkelingen. Dat betekent dat een 'nee' van een organisatie niet noodzakelijkerwijs betekent dat je hun data niet mag bestuderen, maar het geeft ons ook geen vrijbrief om zomaar data te verzamelen; het onderzoek moet wel verantwoord zijn.

8.8.3 WELK PLATFORM KIES JE?

Ethische vragen omtrent sociale media gaan niet alleen maar om de mensen wier data we analyseren, maar ook om onze keuzes die betrekking hebben op de validiteit en betrouwbaarheid van ons onderzoek. Een van die keuzes gaat over het platform dat we willen onderzoeken: verzamel je data van X (voorheen Twitter), Instagram, TikTok, Mastodon, Facebook, Snapchat,

32 Opvallend genoeg wordt wetenschappelijk onderzoek vrijwel nooit genoemd in de gebruiksvoorwaarden, maar wordt het verzamelen van data vaak gewoon categorisch uitgesloten.

LinkedIn, Pinterest, of ...? De keuze die je hier maakt, beïnvloedt wat je kunt bestuderen, en de conclusies die je hierover kunt trekken. Immers: ieder platform heeft andere eigenschappen en omgangsvormen.

De Turks-Amerikaanse socioloog Zeynep Tufekci verwees ooit (2014) naar Twitter als "model-organisme" voor onderzoek naar sociale media. Waar biologen vaak specifieke organismen uitkiezen om in het lab te bestuderen vanwege de praktische voordelen die ze bieden, kozen communicatiewetenschappers vaak voor Twitter omdat de data toentertijd eenvoudig en op grote schaal te verkrijgen was. Praktische haalbaarheid is een valide reden om het ontwerp van je studie aan te passen, maar we moeten uitkijken dat het niet omslaat in gemakzucht; het is schadelijk voor de wetenschap als iedereen alleen maar 'makkelijk onderzoek' doet, omdat we dan het grotere verhaal uit het oog verliezen.

Wat zijn dan wel goede redenen om bepaalde platforms te bestuderen? Dat ligt aan je onderzoeksvraag en de onderliggende motivatie voor je onderzoek. Waar wil je eigenlijk iets over kunnen zeggen, en welke (sub)populaties zijn voor jou van belang? Het is in ieder geval goed om je bewust te zijn van de technische verschillen tussen platforms, maar ook van de demografische verschillen. Gottfried (2024) laat goed zien hoe groot die verschillen zijn bij Amerikaanse socialemediagebruikers (anno 2024); jongere Amerikanen zitten bijvoorbeeld veel vaker op Instagram, TikTok en BeReal dan oudere Amerikanen, en gezinsinkomen is een goede voorspeller voor het gebruik van LinkedIn (dat vaker gebruikt wordt door rijkere Amerikanen).

8.9 SCHADELIJKE INHOUD

In Hoofdstuk 2 hebben we al gesproken over maatschappelijke relevantie als motiverende factor van je onderzoek. Veel van deze onderwerpen zijn ook geassocieerd met controverse en misinformatie (bijvoorbeeld klimaatverandering of het Nederlandse vaccinatiebeleid) of met expliciete inhoud (bijvoorbeeld racisme en seksisme in de maatschappij). Dat levert volgens Kirk et al. (2022) twee problemen op:

1) De inhoud die gecodeerd wordt kan schadelijk zijn voor de codeurs. Bijvoorbeeld wanneer ze blootgesteld zijn aan schokkende beelden, of een grote hoeveelheid seksistische berichten.
2) Eventuele voorbeelden in het verslag van je onderzoek kunnen schadelijk zijn voor lezers, de personen die in de boodschappen genoemd worden, of de maatschappij.

Die problemen hebben implicaties voor de manier waarop je je onderzoek uitvoert en vervolgens rapporteert. Deze onderdelen worden hieronder apart besproken.

8.9.1 UITVOERING

Bij het uitvoeren van je onderzoek zul je rekening moeten houden met de mentale gevolgen van het verwerken van schadelijke informatie. Kirk et al. (2022) geven hiervoor het volgende stappenplan:

1) Maak van tevoren een inschatting van het soort boodschappen dat je zult tegenkomen tijdens het onderzoek. Zijn jullie als onderzoeksteam bereid om hieraan blootgesteld te worden?
2) Blijf met elkaar communiceren over je ervaringen. Het is goed om je ervaringen met elkaar te delen, en ervoor te zorgen dat het goed gaat met iedereen.
3) Denk na over manieren om de blootstelling aan schadelijke boodschappen te beperken. Bijvoorbeeld door het codeerproces zoveel mogelijk te automatiseren. Of als dat niet kan: foto's zwart-wit maken om de impact van de beelden te verkleinen, of bepaalde woorden automatisch te vervangen door een algemenere term ("Alle BEVOLKINGSGROEP moeten GESMURFT worden.")
4) Maak een overzicht van de beschikbare mentale ondersteuning, bijvoorbeeld adviezen over wat te doen bij constante negatieve gedachten.
5) Bespreek het onderzoek na met alle betrokkenen. Wat heeft het onderzoek opgeleverd? Hoe hebben jullie het proces ervaren? Wat ging er wel/niet goed? Hoe zouden jullie dit in de toekomst kunnen verbeteren?

8.9.2 RAPPORTAGE

Bij het rapporteren van je onderzoek zul je rekening moeten houden met de gevoeligheid van het onderwerp, om ervoor te zorgen dat je geen misinformatie verspreidt of je lezers ongevraagd beschadigt met schokkende teksten of beelden. Kirk et al. (2022) geven hiervoor de volgende tips:

1) Doe aan verwachtingsmanagement. Geef aan dat het verslag inhoud bevat die mogelijk als schokkend ervaren kan worden. Zorg ervoor dat kwetsende voorbeelden niet prominent op pagina 1 staan, maar dat lezers ook de mogelijkheid hebben om deze voorbeelden te vermijden.
2) Blijf respectvol. Denk na over de mogelijke gevolgen van de publicatie van jouw onderzoek voor de bevolkingsgroepen om wie jouw onderzoek gaat. Staan er kwetsende dingen in jouw verslag over deze groepen? Zijn dat vermijdbare uitingen?
3) Schep afstand. Zorg ervoor dat er een duidelijk onderscheid is tussen het onderzoek (wat jouw bevindingen zijn) en de meningen die gegeven worden in aanstootgevende boodschappen. De auteurs geven hier twee strategieën voor:
    - Je kunt enige vorm van censuur toepassen door kwetsende woorden vervangen door generieke labels (zoals SCHELDWOORD), of door delen van afbeeldingen te maskeren, indien de exacte inhoud van de kwetsende boodschappen niet relevant is.

- Een andere manier om afstand te scheppen is bijvoorbeeld om voorbeelden expliciet te markeren als KWETSEND of MISINFORMATIE, bijvoorbeeld via een watermerk op de foto's of schermafbeeldingen die je gebruikt.

Kirk et al. (2022) bespreken ook strategieën om wetenschappers te beschermen tegen de mogelijke gevolgen van de publicatie van hun onderzoek. (Bijvoorbeeld harde online kritiek en lastercampagnes van mensen die het niet met het onderzoek eens zijn.) Aangezien dit minder relevant is voor de onderzoeksprojecten uit deze cursus (die verder niet gepubliceerd worden, en dus geen verdere reacties uitlokken) laten we deze informatie hier achterwege.

8.10 RESULTATEN DELEN

Wanneer je klaar bent met het uitvoeren je studie, dan moeten de resultaten nog wel gerapporteerd worden. Ook daar komen meerdere ethische vragen bij kijken. Hieronder bespreken we er twee: (1) Hoe kun je op een zinvolle manier je publiek bereiken? En (2) hoe ver ga je in het communiceren van je conclusies?

8.10.1 HET PUBLIEK BEREIKEN

Traditioneel gezien is het publiceren van een verslag in een wetenschappelijk tijdschrift het eindpunt van je studie; na de openbaarmaking van de resultaten kunnen anderen de resultaten zelf lezen en beoordelen wat ze ervan vinden. Veel wetenschappers worden beoordeeld op de hoeveelheid en de kwaliteit van hun artikelen. Maar er zijn (minstens) twee problemen met wetenschappelijke publicaties:

1) Er zijn zoveel publicaties dat jouw resultaten verworden tot een speld in een hooiberg. Als mensen jouw resultaten niet opmerken, wat hebben we er dan aan?
2) Wetenschappelijk onderzoek wordt vaak gepubliceerd in een andere taal (Engels) en op zo'n manier gepresenteerd dat het lastig is om als leek te bepalen wat de bevindingen eigenlijk zijn. Met andere woorden: het onderzoek is niet toegankelijk.

Wanneer je onderzoek doet met een bepaalde doelgroep is het goed om na te denken over manieren waarop je de resultaten van jouw onderzoek met hen zou kunnen delen. De makkelijkste manier om dat te doen, om in ieder geval het eerste probleem op te lossen, is om gewoon het artikel op te sturen.[33] Een manier om beide problemen op te lossen is om de resultaten van het onderzoek ook via andere kanalen te communiceren, bijvoorbeeld via een

33 De inhoudsanalyse van Proferes et al. (2021) suggereert echter dat dit nog (te) weinig gebeurt. In hun sample van meer dan zevenhonderd studies die uitgevoerd waren met data van het socialemediaplatform Reddit vonden ze maar weinig voorbeelden waar de auteurs zelf hun bevindingen gedeeld hebben met de gemeenschappen die zij bestuderen.

blog, een ingezonden brief in de krant, een toegankelijke video of een podcast.[34] Andere voorbeelden zijn publieksevenementen zoals *Night University* en de *Kinderuniversiteit* die ieder jaar bij Tilburg University gehouden worden.

8.10.2 ACTIVISME?

Je kunt ook verder gaan dan alleen het vertalen en toegankelijk maken van jouw bevindingen voor een breder publiek. Er zijn ook verschillende voorbeelden van wetenschappers die zich naar aanleiding van hun onderzoek actief inzetten om de wereld te verbeteren. Tegenstanders van dit activisme vinden dat je als wetenschapper altijd objectief moet zijn, en je dus ook geen kant moet kiezen in het maatschappelijke debat. Binnen deze visie is er een duidelijke taakverdeling in de maatschappij: wetenschappers zorgen voor betrouwbare informatie, maar het is vervolgens aan de politiek om daar iets mee te doen. Voorstanders van activistisme zeggen dat zij zich juist wel moeten inzetten, omdat zij zich door hun overtuigende resultaten gedwongen voelen om anderen te overtuigen van de omvang van de problematiek en de noodzaak van verandering. Over precieze maatregelen kunnen we discussiëren, maar het is wel belangrijk dat er (snel) iets gebeurt.[35]

8.11 TOT SLOT

Eerder in dit boek hebben we al twee onderwerpen besproken die te maken hebben met ethiek: maatschappelijke relevantie en het betrekken van ervaringsdeskundigen (Hoofdstuk 2) en dataverzameling (Hoofdstuk 3). Na dit hoofdstuk zullen we nog een aantal ethische vragen behandelen:

Hoofdstuk 9: Datamanagement

- Hoe kun je de data op een verantwoordelijke manier opslaan en archiveren?
- In hoeverre kun je data van sociale media anonimiseren of pseudonimiseren?
- Mag je de data die je verzameld hebt ook weer delen met anderen?

Hoofdstuk 10: Rapportage

- Hoe verwijs je naar de data die je hebt verzameld? Moet je de auteur juist wel of niet citeren?

34 Wat de beste methode is om de resultaten te delen met de doelgroep kun je natuurlijk ook gewoon vragen aan een aantal ervaringsdeskundigen.

35 Zie Looman (2024) en Zom & Van der Deijl (2024) voor een toegankelijke discussie van deze onderwerpen in *Onderzoek*, het tijdschrift van NWO.

HOOFDSTUK 9
Datamanagement

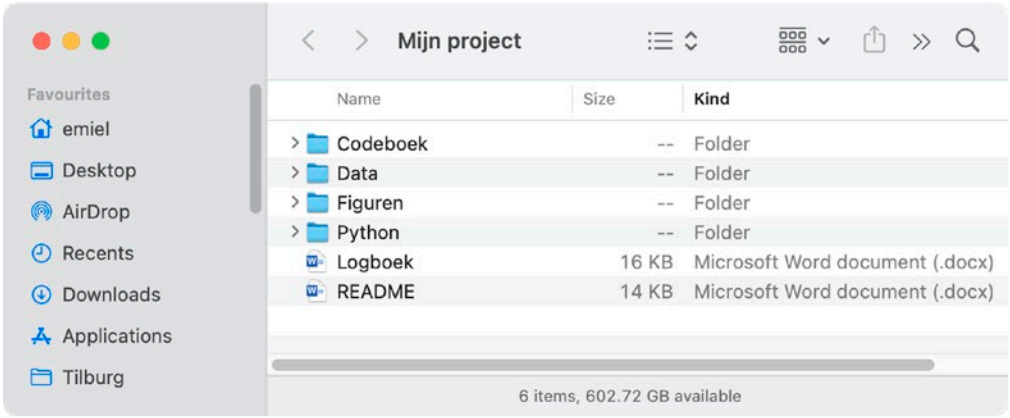
9.1 INLEIDING

In dit hoofdstuk kijken we naar manieren om onderzoeksdata zo goed mogelijk te beheren. Goed beheer betekent niet alleen dat je voorkomt dat je gegevens kwijt raakt, maar ook dat je nadenkt over de transparantie van het onderzoeksproces, de langetermijnopslag van je data, en jouw verantwoordelijkheden als beheerder van die data. Als je aan het begin van je project goed nadenkt over datamanagement, pluk je daar gedurende het hele onderzoek (en lang daarna) de vruchten van.[1]

9.2 DE PROJECTMAP

De organisatie van je project begint met de projectmap: één centrale map waarin alle relevante data, software, aantekeningen, en andere relevante bestanden opgeslagen zijn. Samen met je onderzoeksteam spreek je af om alle gegevens systematisch op te slaan in een goed georganiseerde map, zodat iedereen precies weet waar alles staat. Zo'n map ziet er dan bijvoorbeeld uit als in Figuur 9.1.

1 De aanbevelingen in dit hoofdstuk zijn deels gebaseerd op eerdere presentaties van Bryan (2015) en Navarro (2022).

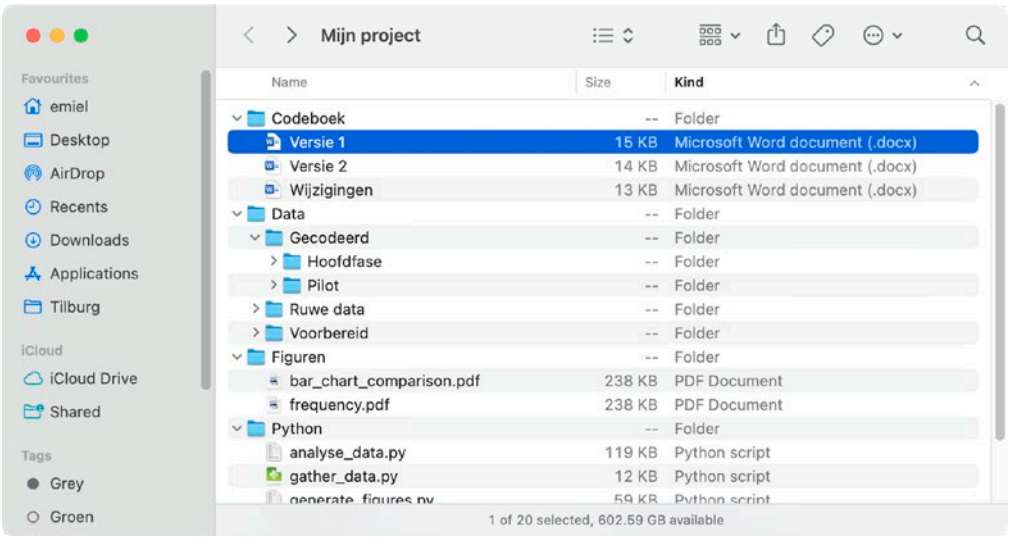


FIGUUR 9.1 *Voorbeeld van een projectmap.*

Binnen de projectmap staan verschillende mappen voor de losse onderdelen van het project (Codeboek, Data, Figuren, Python-scripts), samen met een logboek en een document met een algemene beschrijving van de inhoud en indeling van de map. Hierboven is dat document "README" genoemd, om aan te sluiten bij de internationale conventie om dat zo te doen. Dit bestand is straks, als het onderzoek is afgerond, het eerste bestand dat anderen zullen lezen als ze de map openen. Tijdens het onderzoek kun je hier informatie toevoegen over de bestanden die je in de map zet.

9.2.1 MAPPEN

Zoals gezegd bevat de projectmap meerdere mappen met bestanden die relevant zijn voor het onderzoek. Een map is niets meer dan een middel om verschillende bestanden bij elkaar op te slaan, zodat je het overzicht niet kwijtraakt. Een groot project kan uiteindelijk honderden verschillende bestanden opleveren, waardoor het een uitdaging kan worden om precies het goede bestand te vinden. Je weet dat alles goed georganiseerd is als anderen zonder verdere instructies begrijpen waar de verschillende bestanden opgeslagen zijn. Dat kun je bereiken door ervoor te zorgen dat de structuur van de projectmap overeenkomt met het verloop van het project. Laten we bijvoorbeeld kijken naar de onderstaande map in Figuur 9.2:



FIGUUR 9.2 *Dezelfde map, waarbij nu ook de inhoud van de submappen zichtbaar is.*

Het verloop van het project is zichtbaar in:

- **De verschillende versies van het codeboek**, die apart opgeslagen zijn.
- **De indeling van de Data-map**, waarbij ruwe data, voorbewerkte data, en gecodeerde data een eigen *sub-folder* hebben gekregen. Binnen de map met gecodeerde data zien we ook weer aparte mappen voor de pilot-studie en de hoofdfase van het onderzoek.
- **De losse Python-scripts**, die elk een deel van de dataverwerking voor hun rekening nemen.

Met deze (of een vergelijkbare) indeling weet je precies waar alles staat, en verlies je geen tijd aan het opzoeken van relevante bestanden.

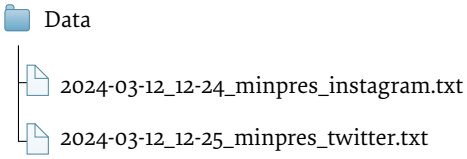
9.2.2 COMPUTERS EN BESTANDSNAMEN

Naast een goede organisatie van de projectmap is het ook belangrijk om na te denken over de naam van je bestanden. Daarbij kunnen we twee doelen onderscheiden: (1) De namen moeten leesbaar en begrijpelijk zijn voor mensen, en (2) De namen moeten eenvoudig te vinden en sorteren zijn op je computer.

Hoewel moderne computers allerlei tekens toestaan in bestandsnamen, is het verstandig om je te beperken tot letters, cijfers, het koppelteken (-) en de *underscore* (_). De laatste twee worden vaak gebruikt om spaties te vervangen, aangezien sommige programma's niet goed kunnen omgaan met spaties in bestandsnamen. Dit is vooral van belang bij bestanden die ook automatisch worden geanalyseerd.

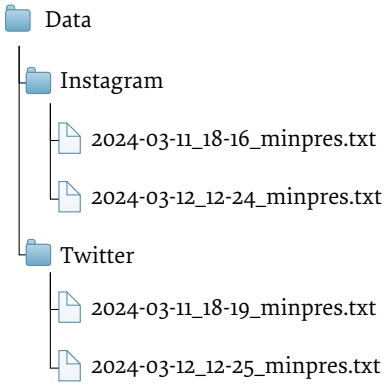
Structuur en inhoud van bestandsnamen

Bestandsnamen geven je niet alleen een idee van de inhoud van het bestand, maar ze bevatten ook vaak verschillende soorten informatie. Stel je voor dat je wil bestuderen hoe onze premier zich presenteert op verschillende sociale media via zijn officiële account. Daarbij zou je bestandsnamen kunnen gebruiken die er zo uitzien als in Figuur 9.3. Deze bestandsnamen hebben steeds dezelfde structuur: [DATUM][TIJDSTIP][ACCOUNT][PLATFORM], gescheiden door een underscore. Hierbij staat de datum ook dusdanig genoteerd dat het eenvoudig is om de verschillende bestanden op datum te sorteren.[2] Dankzij deze structuur is het eenvoudig om automatisch de juiste bestanden te selecteren.



FIGUUR 9.3 *Map met twee bestanden erin. De bestandsnamen gebruiken de volgende structuur: jaar-maand-dag-uur-minuten_accountnaam_platform.txt*

Er is ook een alternatieve oplossing voor hetzelfde probleem: mappen gebruiken voor de verschillende databronnen. Bijvoorbeeld: één map voor Twitter en één map voor Instagram, waarbij de bestandsnamen alleen nog maar datum, tijdstip, en account bevatten (maar geen platform). Dan krijgen we een structuur zoals in Figuur 9.4.



FIGUUR 9.4 *Map waarvan de submappen data bevatten van verschillende websites.*

2 Hiervoor is het gangbaar om de internationale ISO-standaard (ISO 8601) te gebruiken, die aangeeft hoe je data dient te noteren zodat iedereen ze begrijpt. https://www.iso.org/iso-8601-date-and-time-format.html

Waarom zou je voor een map kiezen in plaats van het uitbreiden van de bestandsnaam? Het grote voordeel van de tweede oplossing is dat het overzichtelijker kan zijn om alle bestanden uit dezelfde bron bij elkaar te zetten. (Zeker als je veel data verzamelt!) Voor de computer maakt dit onderscheid niet zoveel uit; als er maar een bepaalde systematiek achter zit die aan de map- en bestandsnamen is af te lezen. Maar voor jezelf en andere mensen is het goed om na te denken wat de meest intuïtieve manier is om je data te organiseren.

Lengte en volgorde
Je kunt bestandsnamen bijna zo lang maken als je wilt,[3] maar op een gegeven moment wordt de bestandsnaam soms afgekapt omdat de naam niet meer in het scherm past. Daarom is het slim om goed na te denken over de volgorde van de onderdelen van je bestandsnaam. Bijvoorbeeld:

- **analyse01_descriptieve-statistiek.py**
- **figuur01_vergelijking-tussen-belgie-en-nederland.py**

Deze namen worden op je scherm misschien ingekort tot "analyse01_descr…" en "figuur01_vergel…", maar het blijft altijd duidelijk waar het bestand bij hoort (data-analyse of het maken van figuren voor je verslag). Je kunt dan meer informatie krijgen door het bestand te selecteren en de rest van de naam te lezen.

9.2.3 WAT SLA JE OP?

Het is goed om aan het begin van het project te bedenken wat je allemaal wil opslaan. Natuurlijk is het belangrijk om je data en codeboek op te slaan tijdens je project, maar welke andere bestanden zijn relevant om te bewaren? Uiteindelijk is het belangrijk dat je transparant bent over het onderzoeksproces en dat het onderzoek in principe reproduceerbaar is.

Transparantie betekent dat je inzicht geeft in het verloop van je onderzoek, en dat het telkens duidelijk is waarom je bepaalde keuzes hebt gemaakt. Veel van de motivatie achter die keuzes zal ook naar voren komen in je uiteindelijke verslag, maar soms worden details vanwege de leesbaarheid of de paginalimiet weggelaten. Die details moeten dan wel ergens bewaard blijven. Hoe je dat doet, verschilt van onderzoek tot onderzoek. Sommige mensen maken notulen voor iedere bijeenkomst, terwijl anderen de details vastleggen in een README-bestand. Natuurlijk moeten ook afspraken buiten jullie bijeenkomsten bijgehouden worden, bijvoorbeeld beslissingen die genomen zijn via e-mail of WhatsApp. Als je die afspraken niet meteen ergens vastlegt, is de kans groot dat ze verloren gaan.

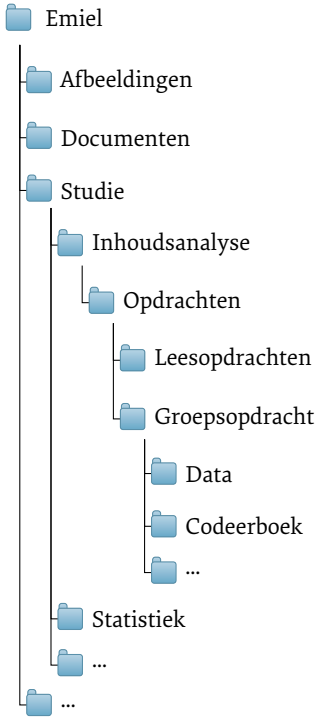
Reproduceerbaarheid betekent dat anderen jouw onderzoek eenvoudig kunnen herhalen en daarbij (ongeveer) dezelfde resultaten zouden kunnen verkrijgen. Daarbij kan het gaan

3 Windows heeft wel een beperking dat het volledige pad naar het bestand niet langer mag zijn dan 265 karakters, maar meestal is dit geen probleem.

om het hele onderzoek, of slechts een deel van het onderzoek (bijvoorbeeld alleen maar de dataverzameling, of alleen maar het coderen). Dat betekent dat je niet alleen de definitieve onderzoeksgegevens moet opslaan, maar ook de tussenliggende versies. Zo kunnen anderen ook controleren of het onderzoek goed is verlopen.

9.2.4 WAAR SLA JE ALLES OP?

Het is belangrijk om je bestanden ergens op te slaan waar je de bestanden makkelijk terug kunt vinden, en waar je ze ook niet zomaar per ongeluk kunt verwijderen. Op je eigen computer kun je jouw bestanden bijvoorbeeld zo organiseren als in Figuur 9.5.



FIGUUR 9.5 *Bestandsstructuur op de computer. De figuur is weergegeven als een boom, waarbij de inhoud van de mappen is aangeven door middel van vertakkingen. De map Studie bevat bijvoorbeeld de mappen Inhoudsanalyse en Statistiek.*

Back-ups

Naast lokale opslag is het ook goed om na te denken over back-ups. Er zijn verschillende programma's om automatisch back-ups te maken van je computer op een externe harde schijf, maar je kunt ook overwegen om gebruik te maken van online diensten zoals Dropbox, Google Drive, enzovoorts. Het is daarbij wel belangrijk om na te denken over de aard van de data. Als je werkt met vertrouwelijke data, moet je er zeker van kunnen zijn dat anderen niet zomaar

toegang hebben tot die data. Sommige onderzoekers aan de universiteit werken daarom ook niet met commerciële partijen samen, maar maken gebruik van beveiligde wetenschappelijke infrastructuur zoals de *SURF Research Drive*.

Samenwerken

Samenwerken met anderen brengt veel voordelen met zich mee, maar ook een aantal uitdagingen. Hoe zorg je er bijvoorbeeld voor dat iedereen werkt met de juiste versie van de data en andere bestanden die ontwikkeld zijn binnen het project? En hoe zorg je ervoor dat je alles op een veilige manier met elkaar deelt? Hiervoor zijn verschillende oplossingen, bijvoorbeeld:

- **Samenwerken in een online omgeving** zorgt ervoor dat iedereen toegang heeft tot de laatste versie van de bestanden waarmee je werkt. Het is dan vaak wel belangrijk dat je goede afspraken met elkaar maakt, zodat je bijvoorbeeld niet elkaars bestanden overschrijft. Het kan verstandig zijn om ook iemand verantwoordelijk te maken voor het maken van offline back-ups.
- **Gebruik versienummers voor belangrijke bestanden.** Als je na iedere wijziging van het bestand een nieuw versienummer hanteert, dan kan iedereen controleren of hun bestand nog up-to-date is.[4] Later in dit hoofdstuk zullen we dit onderwerp verder behandelen, bij de bespreking van versiebeheer.
- **Werk modulair**. Als je ervoor zorgt dat iedereen verantwoordelijk is voor een eigen onderdeel, dan kun je (in theorie) los van elkaar werken en de resultaten op een later moment bij elkaar zetten.

De exacte voorzorgsmaatregelen die je neemt zijn afhankelijk van de aard van het project. Uiteindelijk is het allerbelangrijkst om heldere afspraken te maken zodat iedereen weet wat er van elkaar verwacht wordt.

9.3 PRIVACY EN/VERSUS TRANSPARANTIE

In dit hoofdstuk hebben we datamanagement vooral besproken vanuit het perspectief van onderzoekers die transparant willen zijn over hun onderzoek, en die ervoor willen zorgen dat hun data goed opgeslagen is. Op deze manier kun je achteraf laten zien hoe je tot je conclusies bent gekomen, en je data en methoden delen met andere onderzoekers. Tegelijkertijd heb je als onderzoeker ook de verantwoordelijkheid om goed te zorgen voor andere belanghebbenden: vertrouwelijke informatie moet vertrouwelijk behandeld worden, gege-

4 Dit wordt in de programmeerwereld vaak gedaan aan de hand van semantic versioning. Programma's krijgen dan een versienummer, zoals 2.0.6, waarbij het eerste getal staat voor grote wijzigingen, het tweede getal staat voor kleinere toevoegingen, en het derde getal voor minimale wijzigingen (bijvoorbeeld het verbeteren van typfouten).

vens moeten geanonimiseerd worden voor zover dat mogelijk is, en de data moet worden opgeslagen op een veilige plaats. Deze twee doelen komen soms met elkaar in conflict, bijvoorbeeld wanneer het niet of maar zeer beperkt mogelijk is om de data te anonimiseren. Op dat moment kun je de onderzoeksdata niet zomaar met anderen delen.

9.3.1 PRIVACY

Inhoudsanalyses werken niet met traditionele participanten, maar met gegevens die gemaakt zijn door andere mensen, of verwijzingen naar andere mensen bevatten. Hoewel de onderzoeksdata vaak al openbaar zijn, is het nog steeds belangrijk om deze belanghebbenden te beschermen. Dat komt door het transformatieve karakter van het onderzoek: dankzij het verzamelen en coderen van de data wordt alles beter doorzoekbaar, en komen de belanghebbenden (potentieel) meer in de schijnwerpers te staan dan ze zelf hadden verwacht. Als onderzoeker heb je de verantwoordelijkheid om mensen te beschermen tegen potentieel ongewenste aandacht. Met andere woorden: we moeten zorgen voor de privacy van de belanghebbenden.

Anonimiseren en pseudonimiseren

Anonimiseren betekent dat je de identificerende gegevens van iemand *verwijdert*, zodat de data niet meer aan die persoon gekoppeld kunnen worden. Pseudonimiseren betekent dat je de identificerende gegevens van iemand *versleutelt*, waardoor de identiteit van de betrokken personen verborgen is, maar je verschillende soorten data van dezelfde persoon (bijvoorbeeld verschillende posts op sociale media) wel aan elkaar kunt koppelen. Om de bespreking hieronder beknopt te houden, zullen we voornamelijk anonimisatie bespreken.

Er zijn verschillende manieren om data te anonimiseren, bijvoorbeeld:

1) De auteursnaam verwijderen, of (bij pseudonimiseren) vervangen door een code.
2) Verwijzingen naar bestaande personen uit een tekst verwijderen.
3) Gezichten uit foto's of video's vervagen (*blurren*) zodat ze onherkenbaar zijn.
4) Stemmen van personen in geluidsopnames vervormen zodat ze onherkenbaar zijn.

Hoe meer je over anonimisatie nadenkt, des te lastiger het wordt. Kun je mensen bijvoorbeeld ook niet herkennen aan hun houding of de kleding die ze dragen? Of aan de manier waarop ze praten en de woorden die ze gebruiken? De uitdagingen die er spelen bij het anonimiseren, en de risico's die ontstaan wanneer de identiteit van de belanghebbenden openbaar wordt, bepalen ook in hoeverre je de data openbaar kunt maken (zie de paragraaf over Transparantie hieronder).

Er zijn verschillende momenten waarop we de data kunnen anonimiseren:

1) **Meteen na het verzamelen**. Hierdoor kun je ervoor zorgen dat codeurs ook niet blootgesteld worden aan de auteurs van de data of de personen die worden afgebeeld. Dat heeft het grote voordeel dat codeurs onbevooroordeeld zijn bij het coderen; ze weten immers

niet om wie het gaat. Het nadeel van anonimiseren is dat er ook een deel van de inhoud en context verloren gaat. Hier zul je dus een afweging moeten maken.

2) **Bij het publiceren van de data**. Hierdoor kun je ervoor zorgen dat de identiteit van de belanghebbenden niet publiek bekend wordt. Als anonimisatie ervoor zorgt dat de data volledig onbruikbaar wordt, dan is het goed om na te denken over alternatieve manieren om anderen inzicht te geven in het onderzoeksproces.
3) **Bij het schrijven van het verslag**. Het is belangrijk voor de rapportage om inzichtelijke voorbeelden te geven, maar tegelijkertijd komen die voorbeelden wel onder een vergrootglas te liggen. Dat maakt het extra belangrijk om je voorbeelden goed te kiezen, en ervoor te zorgen dat de belanghebbenden geen nadelen ondervinden van de publicatie van je onderzoek.

Technologie en anonimisatie

Anonimisatie valt tegenwoordig steeds beter te automatiseren. Zo kun je bijvoorbeeld automatische gezichtsherkenning inzetten om op grote schaal gezichten onherkenbaar te maken. Onderzoekers aan de universiteit van Heidelberg hebben hiervoor het programma deface ontwikkeld, dat werkt om video's en foto's te anonimiseren. Dat maakt het mogelijk om data te publiceren die anders te gevoelig zou zijn om te publiceren (en waarvoor volledig handmatig anonimiseren veel te arbeidsintensief zou zijn). Zoals we in het hoofdstuk over Mensen en Computers hebben gezien, blijft het wel belangrijk om je bewust te zijn van de beperkingen van technologie; er zijn altijd tekortkomingen en uitzonderingsgevallen. Het is verstandig om handmatig te controleren of automatische oplossingen inderdaad doen wat ze beloven.

De andere kant van automatisering is dat het ook mogelijk is om gegevens te de-anonimiseren. De eenvoudigste manier om dit te doen is om geanonimiseerde berichten in een online zoekmachine te stoppen, en om te kijken of je het originele bericht kunt vinden. Ayers et al. (2018) hebben gekeken of ze op deze manier originele Tweets konden vinden die werden geciteerd in onderzoeksverslagen. Dit was mogelijk in 84% van de gevallen, wat betekent dat de originele auteur voor deze berichten eenvoudig te achterhalen is. Natuurlijk is het de vraag of andere mensen dat ook daadwerkelijk zouden doen, maar het is goed je bewust te zijn van de risico's wanneer je gevoelige citaten zou willen delen met anderen.

Veilige data-opslag

Om te zorgen voor goede privacy, moet de data ook veilig opgeslagen zijn. Veilig betekent in dit geval dat alleen de leden van het onderzoeksteam toegang hebben tot de data, en dat anderen er niet bij kunnen. Vaak gebeurt dat in een beveiligde omgeving waar je een wachtwoord voor nodig hebt (het liefst met tweestapsverificatie, bijvoorbeeld met een wachtwoord en een code die naar je telefoon gestuurd wordt). Maar hoe goed de omgeving ook is, beveiliging is altijd maar zo sterk als de zwakste schakel. Als je bijvoorbeeld even wegloopt van je

computer om koffie te halen, zonder het scherm te vergrendelen, dan kan iedereen zo bij de gegevens die je op dat moment geopend hebt.

Veilig data versturen
Het kan gebeuren dat je de data per e-mail wil sturen naar anderen, of zou willen overzetten naar een andere computer met een USB-stick. In die gevallen is het verstandig om de bestanden te beveiligen met een wachtwoord. De eenvoudigste oplossing is om alle bestanden samen op te slaan in een gecomprimeerde (ook wel: *gezipte*) map, en dit bestand van een wachtwoord te voorzien.[5]

9.3.2 TRANSPARANTIE

Transparantie gaat over het inzichtelijk maken van het onderzoeksproces. Als onderdeel van het inzichtelijk maken van dat proces kun je ervoor kiezen om de data openbaar beschikbaar te maken. Er zijn daarbij meerdere gradaties van openheid:

- **De data openbaar maken**. Dit is de meest open optie: je publiceert eigenlijk gewoon alles wat je hebt. Naast de data publiceer je ook een (korte) handleiding om een overzicht te geven van de data en wat je ermee kunt doen. Deze optie is het meest geschikt voor datasets waarbij er geen privacyproblemen of problemen met *copyright* zijn.
- **Alleen de metadata openbaar maken**. Met deze optie blijft al het werk dat de codeurs hebben gedaan openbaar, maar de gegevens die je gecodeerd hebt zijn geen onderdeel van de data die je publiceert; er zijn slechts verwijzingen (URLs) naar de originele boodschappen die gecodeerd zijn. Daarmee zorg je ervoor dat anderen de dataset wel kunnen reconstrueren (door de originele boodschappen opnieuw te downloaden), maar dat de originele auteurs volledige controle houden over hun data. Het nadeel aan deze optie is dat de originele auteurs hun berichten kunnen verwijderen, maar dat is beter dan het alternatief (berichten blijven bewaren van mensen die eigenlijk niet willen dat hun gegevens online staan).
- **De data op verzoek ter beschikking stellen**. Met deze optie zorg je ervoor dat alle onderzoeksgegevens veilig gearchiveerd zijn, en dat anderen contact met je kunnen opnemen om de data te verkrijgen. Het voordeel hiervan is dat je volledige controle houdt over je gegevens en waar ze belanden. Het nadeel is dat deze methode extra drempels opwerpt voor andere onderzoekers om de data te krijgen. Bovendien is het onduidelijk hoe houdbaar deze methode is op de lange termijn. (Ben je over tien jaar nog steeds goed te bereiken, en kun je de data ook dan nog eenvoudig vinden?)[6]

5 Er zijn ook online diensten om bestanden op een veilige manier naar elkaar te sturen. Een voorbeeld hiervan is de SURF Filesender. Op het moment van schrijven is deze dienst helaas alleen beschikbaar voor onderzoekers.

6 Eerdere onderzoeken hebben laten zien dat auteurs, ondanks hun belofte om de data op verzoek te delen, lang niet altijd bereid of in staat zijn om hun onderzoeksgegevens ook daadwerkelijk te delen (zie o.a.

- **De data op verzoek ter beschikking stellen, na het ondertekenen van een geheimhoudingsverklaring**. Deze optie geeft je alle voor- en nadelen van de vorige optie. De geheimhoudingsverklaring geeft je verder nog een extra middel om de participanten te beschermen. Dit wordt vaak gedaan voor gevoelige data.

Als je klaar bent met je onderzoek, zijn al deze opties prima. Je kunt er zelfs voor kiezen om de data niet beschikbaar te maken, maar elke keuze vereist een motivatie: waarom heb je hiervoor gekozen? Naast de data moet je ook transparant zijn over het onderzoeksproces. Daarvoor is het belangrijk om na te denken over versiebeheer.

9.4 VERSIEBEHEER

Verantwoordelijk onderzoek doen betekent dat je zorgvuldig te werk moet gaan, zodat jouw tijd en de tijd van de participanten (indien van toepassing) niet voor niets is geweest. Binnen het kader van datamanagement is **versiebeheer** erg belangrijk om te voorkomen dat je jouw werk niet verliest, en later transparant kunt zijn over het proces. De term 'versiebeheer' verwijst naar het bijhouden van de verschillende veranderingen die optreden in de bestanden die je produceert tijdens het onderzoek. Als je die veranderingen goed bijhoudt, kun je altijd teruggaan naar een eerdere versie van je bestanden, en kunnen anderen reconstrueren wat je gedaan hebt.

De algemene stelregel voor versiebeheer is: **zorg voor een apart document bij iedere stap van het onderzoek**. Dat betekent in de praktijk dat je de meeste bestanden twee keer opslaat: een versie voor het archief, en een versie om mee verder te werken. Dus:

- Wanneer de dataverzameling klaar is, sla je de ruwe data twee keer op. De gearchiveerde data raak je tijdens de rest van het onderzoek niet meer aan, zodat je niets per ongeluk overschrijft. Daarnaast is er nog een versie van de data die je kunt coderen.
- Bij het ontwikkelen van het codeboek sla je ook meerdere versies op. Bijvoorbeeld: de eerste versie die een beeld geeft van het oorspronkelijke onderzoeksplan, de versie waarmee je de pilotstudie begint, en de verbeterde versie waarmee het hoofdonderzoek is uitgevoerd.[7]
- De gecodeerde data wordt ook meerdere keren opgeslagen. Vaak is er een los bestand voor iedere codeur, waarin de codeurs hun eigen deel van de data hebben gecodeerd. Daarnaast is er ook een samengevoegd bestand waarmee de intercodeursbetrouwbaarheid is berekend. Ten slotte is er vaak nog een versie waarmee de uiteindelijke statistieken zijn berekend (en waarin de onenigheden tussen de codeurs zijn opgelost).

(Hussey, 2023; Krawczyk & Reuben; 2012; Tedersoo et al., 2021)).

7 Overweeg dit ook te doen voor alle programmeercode die ontwikkeld wordt tijdens het project.

9.5 TOT SLOT

Met versiebeheer zijn we aan het eind gekomen van dit hoofdstuk. Een belangrijke drijfveer achter dit hoofdstuk is het idee dat we als onderzoekers transparant moeten zijn over onze werkwijze. Dat levert soms wat extra werk op, doordat alles goed gedocumenteerd moet zijn. Tegelijkertijd kan een goede documentatie ook voorkomen dat je data per ongeluk verliest. Daarnaast hebben we gezien dat transparantie en privacy soms met elkaar in conflict kunnen komen, maar dat er ook in die situaties verschillende opties zijn om inzicht te geven in je werkwijze, zonder de betrokkenen enige schade te berokkenen.

HOOFDSTUK 10
Rapportage

10.1 INLEIDING

Schrijfopdrachten aan de universiteit zijn een proeve van bekwaamheid. Met een verslag laat je zien dat je alle onderdelen van de cursus beheerst, en dat de leerdoelen van de cursus behaald zijn. Daarnaast gelden er nog algemene normen over academische schrijfvaardigheid waaraan je behoort te voldoen, zoals het hanteren van de gangbare stijlregels (bijvoorbeeld APA), gepast taalgebruik (niet te informeel, maar ook zeker niet te wollig), en tot slot: algemene kennis en een niveau van reflectie dat past bij de fase van jouw studie. Hieronder staan verschillende aandachtspunten voor ieder onderdeel van het verslag (Inleiding, Methode, Resultaten, Discussie/Conclusie; zie Bijlage E voor de richtlijnen voor het groepsrapport), gevolgd door nog wat algemene tips. Het hoofdstuk neemt aan dat je al eerder een vak hebt gevolgd over academisch schrijven, en dat je de basis beheerst. Met andere woorden: het is een aanvulling op de vaardigheden die je eerder geleerd hebt, toegespitst op het rapporteren van inhoudsanalyses.

10.2 HOE SCHRIJF JE EEN INLEIDING?

In een inleiding maak je duidelijk wat je onderzoekt, waarom je dat onderzoekt, en wat dat ons oplevert. Dat doe je op een aantrekkelijke manier, zodat lezers niet afhaken bij het lezen van je verhaal.[1] Maar hoe schrijf je een inleiding voor een inhoudsanalyse?

1 Als je een academische tekst schrijft, kun je er meestal vanuit gaan dat de lezers mensen zijn die een vergelijkbare studie hebben gevolgd, maar die niets weten van het specifieke onderwerp dat je onderzoekt.

10.2.1 DE AANLEIDING

Meestal begin je een inleiding met de aanleiding voor jouw studie. Die aanleiding geeft vaak ook aan wat de maatschappelijke relevantie is van je studie (zonder die expliciet als zodanig te benoemen). Daarnaast geef je ook duidelijk aan wat het onderwerp van je studie is. Jouw doel als schrijver is dat alles wat volgt op de eerste zin natuurlijk aanvoelt voor de lezer, zodat niets zomaar uit de lucht komt vallen, en belangrijke termen meteen uitgelegd worden. Er zijn verschillende manieren om je tekst te beginnen. Bijvoorbeeld:

Start met het probleem dat je wilt oplossen: probeer de kern van het probleem samen te vatten in één zin. Dat wordt de hoofdgedachte van de eerste alinea, die je daarna met feiten, argumenten en bronnen kunt onderbouwen. Na de omschrijving van het probleem zal de lezer zich afvragen hoe we dat dan op moeten lossen. Vervolgens is het aan jou om aan te geven hoe jouw onderzoek bijdraagt aan een oplossing. (Meestal zijn problemen te groot om in één keer op te lossen.)

Start met een anekdote: dit is eigenlijk een variatie op het eerste punt. Wanneer je begint met een anekdote schets je eerst een beeld van een situatie die illustratief is voor het punt dat je wil maken. Daarna kun je vanuit die anekdote het probleem verder uitwerken. Bijvoorbeeld:

> Op 13 januari 2024 zette de Amerikaanse rapper Nicki Minaj een bericht op sociale media om haar nieuwste project te promoten. Ter verfraaiing van haar bericht gebruikte ze automatisch gegenereerde afbeeldingen die waren gemaakt met behulp van kunstmatige intelligentie. Deze afbeeldingen leidden tot een verhitte discussie op sociale media; het gebruik van 'AI-generated content' zou het creatieve imago van de rapper ondermijnen, omdat zij geen echte ontwerpers had ingezet in haar reclamecampagne maar een automatische oplossing die er volgens sommigen voor zorgt dat kunstenaars steeds meer uitgebuit worden (Kaur, 2024). Generatieve AI (programma's die met behulp van kunstmatige intelligentie verschillende soorten media kunnen produceren) kan duidelijk sterke reacties oproepen, maar het is nog niet duidelijk hoe wijdverspreid het negatieve sentiment omtrent generatieve AI is. Daarom brengt deze studie in kaart hoe er in het algemeen gereageerd wordt op het gebruik van generatieve AI op sociale media.

In dit voorbeeld wordt in de laatste zin ook al een doel geschetst van het onderzoek. Dat kan door expliciet te beschrijven wat het doel is van de studie (een situatie in kaart brengen), of door een algemene vraag te stellen die je later in de inleiding verder kan verfijnen. Zo kun je al vroeg in de inleiding afbakenen waar je onderzoek over gaat, en de lezer een referentiekader geven voor het vervolg van de inleiding.

Clichés

Een valkuil bij het schrijven van de eerste alinea van een tekst, is om te vervallen in algemeenheden of clichés zoals "steeds meer mensen gebruiken sociale media om op de hoogte te blijven van recente ontwikkelingen" of "het internet is inmiddels niet meer weg te denken uit het dagelijks leven." Zulke platgetreden paden zijn voor lezers niet zo interessant, en met het gebruik van clichés mis je dus de kans om de aandacht van de lezer te grijpen met een tekst die ze zal bijblijven. Bovendien zijn clichés ook vaak gedateerd: ze worden al zo lang gebruikt dat je je kunt afvragen of er inderdaad "steeds meer mensen" zijn die sociale media gebruiken.

Als we clichés moeten vermijden, waarom lijken veel teksten toch te beginnen met een cliché? De eerste factor is tijd: veel artikelen die je voor je studie leest zijn al wat ouder. Tijdens het schrijven van die artikelen was het misschien nog wel zo dat sociale media in opkomst waren, en was het een goed idee om die ontwikkeling te beschrijven voor lezers die nog niet zo bekend waren met sociale media. Inmiddels beginnen artikelen niet meer met zo'n samenvatting omdat sociale media gemeengoed zijn geworden. De tweede factor is gemak: het is niet eenvoudig om te beginnen met een lege pagina, en dan voelt het fijn om te beginnen met iets vertrouwds. Als dat je helpt om te schrijven, is het niet erg om voor de eerste versie van je verslag te beginnen met een cliché. Later kun je de tekst altijd herschrijven, maar dan staat er in ieder geval alvast iets op papier.

10.2.2 HUIDIGE STAND VAN ZAKEN

Nadat de lezer een beeld heeft gekregen van het onderwerp van je verslag, kun je kort weergeven wat de stand van zaken is in de wetenschappelijke literatuur: wat weten we al wel over dit onderwerp, en wat weten we nog niet? Daarbij is het belangrijk dat je ook aangeeft waarom het belangrijk is om dat laatste toch te gaan onderzoeken.

Relevantie

De grootste uitdaging bij het weergeven van de huidige stand van zaken is dat je al snel te veel literatuur bespreekt. Welke literatuur is relevant om te bespreken? Deze vraag kun je beantwoorden aan de hand van het doel van het stuk dat je schrijft. In een inleiding is het belangrijk om de belangrijkste begrippen uit de onderzoeksvraag te definiëren, en om te laten zien op welke studies jouw onderzoek voortbouwt.

10.2.3 DOEL VAN DE STUDIE

Misschien heb je in de eerste alinea al een algemeen doel geformuleerd dat je hoopt te bereiken. Na het bespreken van de huidige stand van zaken in de literatuur, is het tijd om het doel van je studie verder te concretiseren. Dat doe je aan de hand van een onderzoeksvraag. In hoofdstuk 2 hebben we eerder al besproken wat geschikte onderzoeksvragen zijn voor een inhoudsanalyse.

Sommige inleidingen houden abrupt op na het stellen van de onderzoeksvraag, maar vaak is het beter om expliciet te maken hoe je de onderzoeksvraag gaat beantwoorden, en wat dat concreet oplevert. Hieronder staat een fictief voorbeeld:

> "Om de toegankelijkheid van visualisaties in communicatie-tijdschriften in kaart te brengen, hebben we de volgende onderzoeksvraag opgesteld: hoe toegankelijk zijn staaf- en lijngrafieken in het Journal of Communication, in de periode 2014-2024? Deze onderzoeksvraag beantwoorden we aan de hand van een inhoudsanalyse waarbij we handmatig analyseren of gepubliceerde grafieken voldoen aan bestaande richtlijnen omtrent toegankelijkheid (bijvoorbeeld: zijn de gebruikte kleurcontrasten ook geschikt voor kleurenblinde lezers?). Dit onderzoek helpt ons om beter te begrijpen welke obstakels er zijn voor slechtziende lezers om toegang te krijgen tot de wetenschappelijke literatuur, en laat ook zien in hoeverre de situatie in de afgelopen tien jaar verbeterd is, en hoeveel ruimte er ook nog is voor verbetering in de toekomst."

Bovenstaand voorbeeld laat zien dat je met twee extra zinnen na de onderzoeksvraag een veel duidelijker beeld kunt geven van de inhoud en meerwaarde van je onderzoek. Daarmee kun je lezers overhalen om ook de rest van je onderzoek te lezen.

10.2.4 VERDERE INHOUD

Voorbeelden
Wanneer je data codeert, is het slim om ook een voorbeeld van die data te laten zien in je verslag. Analyseer je bijvoorbeeld reclames op websites? Laat dan een reclame zien, en leg aan de hand van dat voorbeeld uit waar jullie in het onderzoek naar kijken. Op deze manier kun je alle kernbegrippen concreet maken, en begrijpen lezers beter waar de studie over gaat. Een voorbeeld hoeft overigens niet visueel te zijn; denk bijvoorbeeld ook aan verschillende soorten slogans of andere boodschappen die je als voorbeeld kunt citeren. Denk bij het kiezen van je voorbeelden wel aan de mogelijke schadelijke effecten die de voorbeelden op de lezer kunnen hebben (zie ook de discussie in het hoofdstuk over Ethiek (§8.9) over aanstootgevende boodschappen).

Verwachtingen
Vaak worden hypothesen besproken in een apart theoretisch kader, waarbij je vanuit bestaande theorie of eerdere resultaten beargumenteert wat je verwacht te gaan vinden in je onderzoek.[2] Soms wordt het theoretisch kader ook samengevoegd met de inleiding, waar-

2 Een theoretisch kader is ook de uitgesproken plaats om dieper in te gaan op de bestaande literatuur, en om conflicten te identificeren tussen verschillende studies. Die ruimte heb je minder in een inleiding,

door de hypothesen al in de inleiding uitgesproken worden. Als je verwachtingen hebt over de uitkomst: beschrijf je verwachtingen, en geef kort aan waar deze verwachtingen vandaan komen.

10.3 HOE SCHRIJF JE DE METHODE-SECTIE?

Voor een methode-sectie is het belangrijk dat anderen voldoende informatie hebben om te beoordelen of je inderdaad op deze wijze een antwoord kunt geven op de onderzoeksvraag. Dat betekent dat je ook moet motiveren waarom je verschillende keuzes maakt, en moet aangeven welke beperkingen deze keuzes tot gevolg hebben. Het is dus nadrukkelijk geen narratief ("eerst hebben we X gedaan, en toen deden we Y"), maar een systematisch overzicht van de benadering die jullie gekozen hebben om bepaalde doelen (een representatieve steekproef, een betrouwbaar codeerproces) te behalen.

10.3.1 HERHAALBAARHEID

Voor methode-secties is het van belang dat het onderzoek zo gerapporteerd wordt dat de studie in principe herhaalbaar is. Met andere woorden: dat het verslag eigenlijk een stappenplan geeft dat anderen zouden kunnen volgen om op dezelfde manier tot dezelfde resultaten te komen. Een van de grootste uitdagingen is de asymmetrie tussen schrijver en lezer: als schrijver is het lastig om je in de positie van de lezer te verplaatsen, die niets van jouw onderzoek weet. Daardoor kun je belangrijke details over het hoofd zien, omdat ze voor jou zo vanzelfsprekend zijn dat ze niet meer benoemd hoeven te worden. Dit probleem wordt ook wel *the curse of knowledge* genoemd (Camerer et al., 1989). Hoewel je bijna niet aan deze vloek ontkomt, zijn er wel manieren om het probleem te verkleinen:

- Maak aantekeningen tijdens het onderzoek: schrijf bij alle grote en kleine beslissingen op wat je gekozen hebt, en waarom je dat gekozen hebt.
- Laat je werk lezen door anderen, met de expliciete vraag of ze genoeg informatie hebben om het onderzoek zelf uit te voeren. (De overtreffende trap is om anderen dit ook daadwerkelijk te laten doen. Dat wordt tegenwoordig wel eens gedaan in replicatiestudies.)
- Zoek vergelijkbare studies en kijk welke details zij vermelden. Bekijk daarna of diezelfde details ook vermeld staan in jouw verslag.
- Gebruik bestaande checklists om te controleren of je niets vergeten bent. (Je kunt het template voor het onderzoeksvoorstel in de appendix hiervoor gebruiken.)

10.3.2 KEUZES, BEPERKINGEN, EN MOTIVATIE

De meeste onderzoeken beginnen met een algemene onderzoeksvraag, die gaandeweg verder gespecificeerd wordt. Die specificatie kan een theoretische reden hebben, maar vaak is er ook

waar het meer de bedoeling is om een overzicht te geven dan om de diepte in te gaan.

een praktische oorzaak: je kunt met één studie maar beperkt antwoord geven op een algemene vraag. Veel vragen zijn complex, en vereisen een hele batterij aan studies voordat we ze goed kunnen beantwoorden. Een oplossing is dan om algemene vragen op te splitsen in deelvragen, en stapsgewijs antwoord te geven op de verschillende deelvragen. Zo vormt ieder onderzoek een puzzelstuk waarmee we toewerken naar antwoorden op grotere vragen. Het is belangrijk om je bewust te zijn van de beperkingen van je onderzoek, om het onderzoek ook een plaats te kunnen geven binnen de literatuur: met welk puzzelstukje zijn we nu precies bezig? De vraag is dus niet *of* jouw studie beperkt is, maar *welke* beperkingen er zijn voor jouw studie, en wat voor implicaties die beperkingen hebben. Dat vraagt om een duidelijke motivatie.

Een goede motivatie vooraf bespaart een hoop kopzorgen achteraf. Het is jammer als je pas na de data-analyse beseft dat je eigenlijk alles anders had moeten doen. Vaak is dat niet zo, maar voer je een studie uit om een bepaalde reden, nadat je de voors en tegens hebt gewogen. De studie is dan het resultaat van die overwegingen, maar de eventuele nadelen wegen (hopelijk) niet op tegen de voordelen van de gekozen opzet. Nadelen kunnen geadresseerd worden in vervolgonderzoek.

10.3.3 SAMPLE

Zorg voor een heldere beschrijving van je sample, zodat duidelijk is welke data je hebt verzameld, wanneer je de data hebt verzameld, hoe je dat hebt gedaan, en volgens welke criteria je de selectie hebt gemaakt. Vertel daarbij ook waarom je deze keuzes hebt gemaakt. Die keuzes kunnen gemotiveerd worden op basis van de overwegingen in Hoofdstuk 3, maar ook aan de hand van eerdere studies waarin vergelijkbare keuzes worden gemaakt. Veelvoorkomende valkuilen zijn:

- Het weglaten van relevante informatie over de steekproef (zie: *the curse of knowledge* hierboven).
- Een post-hoc motivatie geven voor de gemaakte keuzes. Probeer keuzes niet achteraf goed te praten, maar houd tijdens het onderzoek bij waarom je bepaalde keuzes hebt gemaakt.
- Het foutief omschrijven van de sampling-methode. Bijvoorbeeld: het etiket ‘random sampling’ is alleen van toepassing op situaties waar er een automatische selectie is gemaakt op basis van een volledige lijst van items.

10.3.4 CODEBOEK

De ontwikkeling van het codeboek verdient een prominente plaats in de methode; het is immers de kern van je onderzoek! Tegelijkertijd is het niet eenvoudig om een goed beeld te geven van het codeboek. Enerzijds is de ontwikkeling van een codeboek een ingewikkeld proces waarachter allerlei overwegingen schuilgaan (die het eigenlijk ook wel verdienen om besproken te worden), maar anderzijds moet de omschrijving van het codeboek ook niet te lang worden. Gelukkig kun je het codeboek zelf ook toevoegen als bijlage.

In de hoofdtekst van het verslag schrijf je in ieder geval op welke categorieën je hanteert, wat die categorieën betekenen, en specificeer je aan de hand van referenties waar die categorieën vandaan komen. Geef het ook expliciet aan als je zelf nieuwe categorieën hebt toegevoegd, of als je bepaalde aspecten van de operationalisatie hebt aangepast.[3] Dat helpt lezers om te begrijpen hoe de studie tot stand is gekomen en om eenvoudiger verbanden te kunnen leggen tussen verschillende studies. Net als bij de inleiding kan het helpen om voorbeelden te gebruiken, maar dan wel spaarzaam; we willen de lezer ook niet overweldigen met informatie.

In de bijlage zet je het volledige codeboek, met verdere uitleg van de categorieën, voorbeelden, herkenningspunten (IFIDs) enzovoorts. Daarnaast kun je ook het codeerformulier toevoegen als bijlage.

Je kunt er in de hoofdtekst voor kiezen om alle categorieën te omschrijven in de lopende tekst, maar soms is het beter om het codeboek weer te geven als een tabel, zodat het verhaal niet onderbroken wordt door een ingewikkelde opsomming.

10.3.5 CODEERPROCES

Het codeerproces is het hart van jouw onderzoek, en het stappenplan moet zo omschreven worden dat jouw onderzoek in principe herhaalbaar is. Daarbij geef je ook aan welke maatregelen je hebt genomen om de resultaten zo betrouwbaar mogelijk te maken. Hieronder volgt een vragenlijst die je kunt gebruiken om het codeerproces te omschrijven.[4]

Algemeen

- Wie zijn de codeurs? (Zijn dit de auteurs of zijn er onafhankelijke codeurs? En hebben zij belangrijke demografische eigenschappen die we zouden moeten weten?)
- Hoeveel codeurs hebben er meegewerkt aan deze studie?
- Hoe zijn de items verdeeld tussen de pilotstudie en de hoofdcodeerfase?

De pilotstudie

- Hoeveel items zijn er (dubbel) gecodeerd in de pilotstudie?[5]
- Hoe betrouwbaar waren de resultaten van de pilotstudie?
- Welke wijzigingen heb je doorgevoerd na de pilotstudie?

3 Zoals gezegd is het ook belangrijk om zulke keuzes te motiveren. Bijvoorbeeld door aan te geven dat je in eerste instantie een bestaande definitie hebt gebruikt, maar dat je na de pilotstudie genoodzaakt was om de operationalisatie aan te passen omdat er te veel onenigheid was tussen de codeurs.

4 Let op: geen enkele checklist is volledig! Afhankelijk van het ontwerp van je studie kunnen er nog meer details zijn die genoemd zouden moeten worden.

5 Alle items in de pilotstudie zijn natuurlijk dubbel gecodeerd; daarvoor is het een pilot.

De hoofdcodeerfase

- Hoe zijn de items verdeeld tussen de codeurs?
- Hoeveel items hebben de verschillende codeurs gecodeerd?
- Hoeveel items zijn dubbel gecodeerd?
- Hoe lang duurde het codeerproces?
- Hoe zijn eventuele onenigheden tussen de codeurs opgelost?

De uiteindelijke betrouwbaarheid van de hoofdcodeerfase kun je rapporteren in de resultatensectie (zie hieronder).

10.3.6 ETHISCHE OVERWEGINGEN

In de eerdere hoofdstukken in dit boek heb je gelezen over de verschillende ethische uitdagingen rondom het uitvoeren van een inhoudsanalyse. In veel gevallen zullen die uitdagingen hebben geleid tot het maken van verschillende keuzes in de opzet van je onderzoek. Als je bijvoorbeeld posts van sociale media hebt onderzocht, is het goed om uit te leggen op welke manier je de socialemediagebruikers in jouw sample hebt geprobeerd te beschermen tegen potentieel nadelige effecten van jouw studie. Afhankelijk van de situatie kun je ervoor kiezen om ethische kwesties te bespreken in de relevante paragraaf (bijvoorbeeld over sampling), of om een aparte paragraaf te wijden aan jouw ethische overwegingen. Als je ook nog iets wil zeggen over mogelijke problemen omtrent *dual use* (oneigenlijk gebruik van jouw onderzoeksresultaten), dan kun je die informatie ook kwijt in de discussie aan het eind van je studie.

10.4 HOE RAPPORTEER JE RESULTATEN?

Resultatensecties volgen meestal dezelfde structuur: eerst beschrijf je de descriptieve statistieken (frequenties en betrouwbaarheid), waarna je de toetsende statistieken (chi-kwadraattoets, t-toets, ANOVA, enzovoorts) rapporteert. Mogelijk leiden de resultaten tot vervolgvragen, die je soms kunt bestuderen in een exploratieve (verkennende) analyse.[6]

10.4.1 DESCRIPTIEVE STATISTIEKEN

Descriptieve statistieken kun je vaak rapporteren met behulp van een tabel. In Tabel 10.1 hieronder zie je een fictief voorbeeld, waarbij de frequenties en betrouwbaarheidsscores in één oogopslag te zien zijn:

6 Bij gepreregistreerd onderzoek vallen alle niet-gepreregistreerde analyses onder de noemer "exploratieve analyse."

Categorie	Groep 1	Groep 2	Totaal	Krippendorff's α	Overeenstemming
A	50	60	110	.59	65%
B	20	30	50	.75	80%
C	40	55	95	1.00	100%

TABEL 10.1 *Voorbeeld van een frequentietabel waarin meteen ook betrouwbaarheidsscores worden gerapporteerd.*

Wat de beste manier is om je resultaten te presenteren, hangt af van de exacte onderzoeksvraag en het ontwerp van je studie. Aandachtspunten zijn in ieder geval:

- Het gebruik van de juiste Griekse letters. Voor Krippendorff's alpha[7] gebruik je α, en voor Cohen's kappa gebruik je de κ.
- Het formatteren van tabellen volgens de APA-richtlijnen. Denk daarbij ook aan de uitlijning van de getallen (rechts uitlijnen om getallen eenvoudiger te kunnen vergelijken) en het ontwerp van de tabel (onder andere: zo min mogelijk verticale lijnen).
- Het duiden van de statistieken: wat betekenen de scores die je gevonden hebt eigenlijk? Wanneer je een tabel presenteert in je verslag, is het belangrijk om de lezer aan de hand te nemen en te beschrijven wat er te zien valt. Zo kun je de ogen van de lezer naar de juiste plaats sturen.

In jouw studie heb je waarschijnlijk al eerder gemiddelden, percentages, of absolute aantallen gerapporteerd in een verslag. Maar hoe bespreek je betrouwbaarheidsscores eigenlijk?

Duiding van betrouwbaarheidsscores

De beste manier om te leren hoe je betrouwbaarheidsscores moet rapporteren, is om te kijken naar voorbeelden van eerdere studies. Hieronder zullen we een citaat analyseren van Van Hooijdonk & Liebrecht (2018) die een studie hebben uitgevoerd naar online klantenservice van Nederlandse gemeenten op sociale media. Daarbij hebben de auteurs zich vooral gericht op het taalgebruik van de gemeenten, en dan specifiek op het gebruik van conversationele elementen die in schril contrast staan met een meer formele, *corporate* manier van communiceren.

Van Hooijdonk en Liebrecht presenteren de betrouwbaarheidsscores in een tabel met de volgende kolommen: *Conversationeel linguïstisch element* (bijvoorbeeld: persoonlijk ondertekenen van berichten, of het gebruik van afkortingen), *Voorbeeld* (teksten die iedere categorie

7 Er zijn ook andere statistieken die dezelfde Griekse letters gebruiken, en die worden soms door elkaar gehaald bij het rapporteren van inhoudsanalyses. Een veelvoorkomende fout is om te verwijzen naar Cronbach's alpha (een maat voor de interne consistentie van vragenlijsten) in plaats van Krippendorff's alpha.

illustreren), *Betrouwbaarheid* (een kappa-score), en *Overeenstemming* (*percentage agreement*). De absolute en relatieve frequenties voor iedere categorie worden apart gerapporteerd, zodat de auteurs zich in hun tekst eerst kunnen richten op het bespreken van de betrouwbaarheidsscores. Dat gebeurt onder de tabel, met de volgende tekst.

> "Het blijkt dat het identificatie-instrument grotendeels leidde tot een voldoende ($\kappa > 0{,}60$) tot perfecte ($\kappa = 1{,}00$) overeenstemming tussen de codeurs. De drie linguïstische elementen die personalisatie operationaliseren, konden op betrouwbare wijze geïdentificeerd worden."
>
> *(Van Hooijdonk & Liebrecht, 2018)*

In deze tekst geven ze een algemene interpretatie van de betrouwbaarheidsscores voor de eerste categorie (*Personalisatie*), waarbij ze aangeven wat het bereik van de scores is. De observatie dat de kappa-scores voldoende zijn betekent ook dat het zinvol is om de resultaten verder te interpreteren. Van Hooijdonk en Liebrecht vervolgen:

> "Dit gold eveneens voor het merendeel van de linguïstische elementen behorende bij informeel taalgebruik, hoewel afkortingen en verkortingen tot minder eenduidige coderingen leidden. Bij de identificatie van afkortingen waren de codeurs het in dertien van de 143 gevallen oneens, bij verkortingen gold dit voor zes gevallen. Wanneer deze gevallen bekeken worden, blijken er twee oorzaken aan deze onnauwkeurige identificatie ten grondslag te liggen."
>
> *(Van Hooijdonk & Liebrecht, 2018)*

De volgende categorie (informeel taalgebruik) wordt hier vergeleken met de eerste categorie, waarna de auteurs concluderen dat de betrouwbaarheidsscores voor het identificeren van *afkortingen* en *verkortingen* beduidend lager zijn. Dat roept natuurlijk vragen op bij de lezer: waarom zijn die scores eigenlijk lager? Die vraag wordt dan ook meteen beantwoord; er zijn twee oorzaken, namelijk:

> "Ten eerste kunnen deze linguïstische elementen makkelijk over het hoofd worden gezien, vooral bij vrij gangbare afkortingen, zoals *NS* in plaats van *Nederlandse Spoorwegen* en *gft* voor *groente, fruit en tuinafval*. Ten tweede bleken de categorieën soms verwisseld te worden. Zo werd een aantal keer *ok* en *info* ten onrechte als afkorting gecodeerd in plaats van een verkorting."
>
> *(Van Hooijdonk & Liebrecht, 2018)*

Wat je dus moet doen als auteur, is nagaan waar lagere betrouwbaarheidsscores vandaan komen. Dat doe je door te kijken naar de dubbel gecodeerde data, en onderling te bespreken waarom de verschillende codeurs voor een andere categorie hebben gekozen. Het zou bij jouw onderzoek ook goed kunnen zijn dat codeurs gewoon iets over het hoofd hebben gezien, of twee categorieën door elkaar gehaald hebben. In ieder geval is het belangrijk om te doorgronden wat er precies fout is gegaan, en om regelmatigheden te vinden in het gedrag van de codeurs. Hieronder volgt een laatste citaat van waarin ze nog een oorzaak (**dikgedrukt**) van een lage betrouwbaarheid bespreken. Daarnaast noemen ze ook verschillende taalkundige patronen die ze gevonden hebben in het gedrag van de codeurs:

> "Wat betreft de strategie *uitnodigende retoriek* konden drie van de vijf conversationele linguïstische elementen betrouwbaar geïdentificeerd worden: *bedanken*, *verontschuldigen* en *het gebruik van humor*. De codeurs bleken het echter lastiger te vinden om te herkennen of een organisatie sympathie of empathie toonde in de webcare, ze waren het in veertien van de 143 gevallen oneens. Wellicht is het **ontbreken van meerdere concrete indicatoren van woorden en uitdrukkingen (IFID's)** die sympathie en empathie uitdrukken hier debet aan. Hoewel de codeurs overeenstemden in het identificeren van de uitdrukking *wat vervelend*, bleken ze te verschillen in het aanmerken van bijwoorden van modaliteit (zoals *helaas*, *gelukkig*) en uitdrukkingen zoals *dat is goed om te horen* als expressie van sympathie of empathie."
>
> *(Van Hooijdonk & Liebrecht, 2018)*

Door betrouwbaarheidsscores op deze manier te bespreken, kunnen vervolgstudies rekening houden met de verschillende factoren die het lastig maken om teksten betrouwbaar te coderen.

Test jezelf

Zoek een artikel waarin betrouwbaarheidsscores worden gerapporteerd en analyseer hoe de auteurs deze scores rapporteren. Probeer daarbij ook te bedenken waarom de auteurs dat op deze manier hebben gedaan. Hieronder staan enkele vragen om je te helpen.

1) Worden de betrouwbaarheidsscores in de tekst gerapporteerd, of staan ze in een tabel?
2) Worden de scores apart gerapporteerd, of samen met andere descriptieve statistieken?
3) Wordt er veel aandacht besteed aan de scores, of vindt er verder geen duiding plaats?
4) Als de auteurs een lagere betrouwbaarheidsscore rapporteren, proberen ze die reden ook te achterhalen? Wat is die reden?

Door artikelen op deze manier kritisch te bevragen, ontdek je vaak ook manieren om de rapportage te verbeteren. Wat zou jij anders gedaan hebben?

10.4.2 TOETSENDE STATISTIEK

Toetsende statistiek rapporteer je op dezelfde manier als je geleerd hebt bij eerdere statistiekvakken. De aandachtspunten bij de rapportage zijn dan ook vergelijkbaar:

- Maak duidelijk welke hypothese je toetst met welke toets. Na het lezen van de methodesectie zijn lezers waarschijnlijk vergeten wat "Hypothese 3" ook alweer was, dus het is goed om de hypothese opnieuw te omschrijven.
- Als een statistische toets aannames doet over de data, is het goed om die aannames ook te toetsen. Dat hoeft echter niet allemaal in de hoofdtekst te staan, maar kan ook in een voetnoot of een appendix.
- Als je heel veel statistische toetsen uitvoert, kan het verstandig zijn om een correctie toe te passen, om Type 1-fouten tegen te gaan.
- Als je een chikwadraattoets uitvoert, zorg er dan ook voor dat het duidelijk is met welke getallen je die toets hebt uitgerekend. Dat kun je bijvoorbeeld doen door te verwijzen naar een eerdere frequentietabel, of door nog een aparte kruistabel te maken met de relevante getallen.
- Laat zien dat je de resultaten ook goed kunt interpreteren. Als je een verschil of een bepaalde verdeling vindt, wat betekent dat dan voor jouw hypothesen? Een verdere contextualisering van de resultaten vindt plaats in de discussie/conclusie, maar het is wel goed om alvast iets te zeggen over de uitkomst van de toetsen die je gebruikt.
- Denk na over de volgorde waarin je de verschillende hypothesen toetst, en probeer ook een verhaal te maken van je verslag (in plaats van alleen maar een opsomming van getallen).

10.4.3 VISUALISATIES

Zeker bij inhoudsanalyses is een resultatensectie niet compleet zonder visualisatie van de data. Iedere inhoudsanalyse die we tot nu toe besproken hebben, heeft wel een tabel, grafiek, of een ander soort figuur om een indruk te geven van de resultaten. Maar hoe kies je de juiste visualisatiemethode? Hieronder bespreken we een aantal strategieën om jouw resultaten optimaal te presenteren.

Wat wil je laten zien?

De eerste vraag die je voor jezelf moet beantwoorden voordat je een visualisatie kiest is: wat wil je laten zien? Het antwoord op die vraag is afhankelijk van twee factoren:

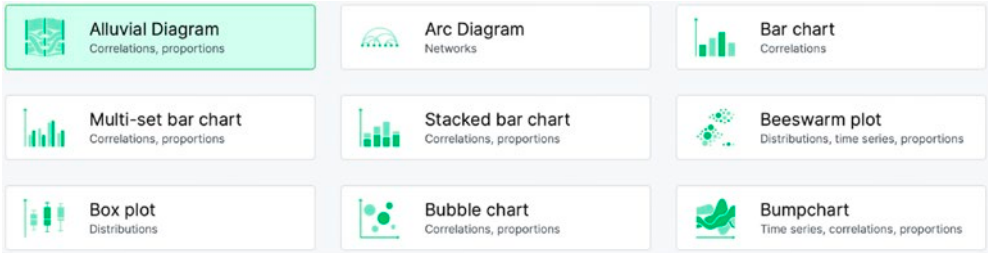
Factor 1: de onderzoeksvraag. In Hoofdstuk 2 hebben we besproken wat voor soort onderzoeksvragen je kunt beantwoorden met een inhoudsanalyse, en met welke doelen die onderzoeksvragen samenhangen. In de resultatensectie geef je antwoord op de onderzoeksvraag, en de manier waarop je de resultaten presenteert moet aansluiten op het doel van jouw

onderzoek. Je hebt een ander soort visualisatie nodig om een algemene verdeling te laten zien, dan om een verandering over tijd in kaart te brengen.[8]

Factor 2: De plaats van de visualisatie in de narratief van je resultatensectie. Aan het begin van de resultatensectie is het vaak goed om een algemeen overzicht te geven van je data, bijvoorbeeld met een tabel of een staafgrafiek met alle categorieën. (We hebben eerder in dit hoofdstuk al zo'n tabel gezien bij de bespreking van descriptieve statistieken.) Maar later in de resultatensectie wil je misschien juist inzoomen op een bepaald aspect van de data, of zelfs op één specifiek voorbeeld.

Wat zijn de mogelijkheden voor jouw soort data?

Data visualisatie is een vaardigheid die je moet leren. Niet alleen moet je leren hoe je verschillende programma's gebruikt om visualisaties te maken,[9] maar je moet ook bekend worden met de verschillende mogelijkheden om data te visualiseren. Daarom is het nuttig om eerst eens te verkennen welke opties er zijn, bijvoorbeeld door verschillende artikelen te lezen waarin een inhoudsanalyse gerapporteerd wordt, maar ook door in verschillende programma's het keuzemenu te bekijken. Figuur 10.1 toont bijvoorbeeld een klein deel van de opties uit het gratis online programma RAWGraphs (Mauri et al., 2017).



FIGUUR 10.1 *Voorbeelden van visualisaties die je met RAWGraphs zou kunnen maken: stroomdiagrammen, arc diagrams, verschillende soorten staafgrafieken, verschillende grafieken om de verdeling van de data te laten zien (box plots, zwermplots), enzovoorts.*

8 Het datavisualisatieteam van de Financial Times heeft een toegankelijke poster gemaakt die verschillende soorten visualisaties categoriseert aan de hand van het doel van de visualisatie. Je vindt de poster op http://ft.com/vocabulary

9 Je kunt visualisaties maken in je favoriete spreadsheetprogramma (Excel, Google Sheets, Calc, Numbers), maar er zijn nog veel meer gratis opties. RAWGraphs (Mauri et al., 2017) is bijvoorbeeld een populaire optie. Als je kunt programmeren in Python, dan is de Seaborn-module een aanrader. In R is ggplot2 een veelgebruikte oplossing.

Als je het leuk vindt om te programmeren, zijn er ook online galerijen met voorbeelden die andere mensen hebben gemaakt (bijvoorbeeld https://r-graph-gallery.com/). Je kunt vaak de code downloaden en aanpassen aan jouw eigen data.[10]

Hoe toegankelijk is de visualisatie?

Visualisaties dienen ter ondersteuning van je verhaal, en zijn er om de lezer te helpen begrijpen wat de resultaten zijn van je onderzoek. Daarvoor moeten die visualisaties echter wel toegankelijk zijn, zodat iedereen de visualisatie kan lezen en begrijpen. Jonathan Schwabish et al. (2022) hebben hiervoor twee verschillende checklists gemaakt.[11]

Checklist 1: Inhoud

De eerste checklist gaat over het taalgebruik en de getallen die je presenteert. Bij een groepsproject kun je de grafieken en tabellen die je maakt laten controleren door je groepsgenoten, maar het kan ook de moeite waard zijn om de tekst en de visualisaties te laten beoordelen door iemand die niet betrokken is bij het project.

- Zorg er bij je visualisaties altijd voor dat de grafiek of tabel een duidelijke titel heeft, die beschrijft waar de visualisatie over gaat. In je verslag zet je deze informatie in het bijschrift.
- Daarnaast is het belangrijk om in de tekst een samenvatting te geven van wat de visualisatie laat zien. Het is dus niet voldoende om iets te zeggen als: "Tabel 1 geeft aan hoe vaak mannen en vrouwen worden onderbroken tijdens het praten." Deze omschrijving is incompleet; je moet ook nog aangeven wat er dan precies te zien valt. Bijvoorbeeld: "De resultaten laten zien dat vrouwen in talkshows twee keer zo vaak (na iedere 20 woorden) worden onderbroken door mannen dan andersom (na iedere 41 woorden)."[12]
- Visualisaties zelf dienen begrijpelijke taal te bevatten, waarbij de X-as en de Y-as van grafieken duidelijk gelabeld worden.
- Getallen moeten zo gepresenteerd worden dat lezers zelfstandig kunnen bepalen hoe vaak iets voorkomt.

Checklist 2: visuele eigenschappen.

De tweede checklist gaat over de vormgeving van de visualisatie. Probeer, naast de controle op je beeldscherm, ook alle visualisaties te printen in zwart-wit. Als je geen goed onderscheid

10 Slechte voorbeelden zijn er ook, zodat je kunt vermijden om dezelfde fouten te maken. Zie bijvoorbeeld *de Grafiekpolitie* (https://nieuwscheckers.nl/tag/grafiekpolitie/) of *Junk Charts* (https://junkcharts.typepad.com/).

11 Het volledige rapport is een aanrader. Zij beschrijven hoe je informatie toegankelijk kunt maken voor mensen met een visuele beperking.

12 Deze resultaten zijn fictief en dienen slechts ter illustratie.

kunt maken tussen de categorieën op basis van de grijswaarden, dan kun je het beste een ander kleurenpalet kiezen.[13]

- Tekstgrootte: visualisaties zijn onbegrijpelijk als de tekst te klein is om te lezen. Zorg ervoor dat alle tekst een grootte heeft van minimaal 9 punten, maar liefst 12 punten.
- Contrast:
  - Zorg ervoor dat er genoeg contrast is tussen tekst en de achtergrond waarop die tekst staat.[14] Voor maximaal contrast kun je zwarte tekst altijd op een witte achtergrond zetten.
  - Zorg ervoor dat er genoeg contrast is tussen de verschillende vormen en iconen en hun achtergrond.[15]
- Redundantie: zorg ervoor dat kleur niet het enige onderscheid is tussen de verschillende categorieën. Denk bijvoorbeeld ook aan verschillende patronen die je kunt gebruiken voor de staven van je staafgrafiek, of verschillende vormen die je kunt gebruiken in een scatterplot. Daarnaast dien je ervoor te zorgen dat alle relevante onderdelen van de visualisatie voorzien zijn van een label.
- Gebruik een schreefloos lettertype in je grafieken, zoals Arial, Verdana, of Helvetica.

10.4.4 EXPLORATIEVE ANALYSES

Als je eenmaal een gecodeerde dataset hebt, kun je niet alleen kijken of de resultaten in lijn zijn met je verwachtingen, maar je kunt de data ook gebruiken om nieuwe vragen te verkennen waar je geen hypothesen voor hebt opgesteld. Zo'n 'extra' data-analyse wordt ook wel een exploratieve analyse genoemd. De status van je bevindingen hangt af van de omvang van je dataset.

- Als je een **census** hebt uitgevoerd, dan kun je met een exploratieve analyse vaststellen wat de eigenschappen zijn van de populatie.
- Als je een **steekproef** hebt genomen, dan kun je met een exploratieve analyse vaststellen wat de eigenschappen zijn van je sample, en de hypothese opstellen dat een nieuwe steekproef dezelfde eigenschappen zou hebben. Die hypothese kan getoetst worden in vervolgonderzoek. Als je dan inderdaad dezelfde eigenschappen waarneemt, kun je stellen dat je observatie generaliseerbaar is naar de gehele populatie.

---

13 Sommige programma's bevatten ook een optie om visualisaties geschikt te maken voor mensen die kleurenblind zijn. In Seaborn (python) wordt gebruikgemaakt van het *Cubehelix* systeem waarbij er automatisch gezorgd wordt voor een palet dat geschikt is voor mensen die kleurenblind zijn.

14 Het rapport noemt een ratio van 4.5:1. Dit zou je kunnen controleren door de kleurcode van tekst en achtergrond te bepalen (als je de kleurcode niet hebt, kun je een *color picker* gebruiken, bijvoorbeeld via de website https://imagecolorpicker.com/), en de waarden in te vullen in een programma dat de contrastwaarde voor je berekent. Bijvoorbeeld via de website: https://webaim.org/resources/contrastchecker/

15 De aanbevolen ratio is hier minimaal 3:1.

10.5 DISCUSSIE

Na de resultatensectie bespreek je de implicaties van je bevindingen, in het licht van de bestaande literatuur. Allereerst vat je de resultaten samen: heb je inderdaad gevonden wat je verwachtte, of is de situatie toch anders dan je dacht?

10.5.1 VERWACHTE BEVINDINGEN

Als je hebt gevonden wat je eerder al had verwacht, kun je hier expliciet aangevingen dat jullie resultaten in lijn zijn met eerder onderzoek. Hierbij kun je ook bespreken wat dat dan betekent voor het vakgebied, bijvoorbeeld: "het patroon dat vrouwen in Nederlandse talkshows twee keer zo vaak onderbroken worden dan mannen, sluit goed aan bij uitkomsten van onderzoek in andere Europese landen: Smith vond … en Lopez vond …. Daarmee vinden we een verdere ondersteuning van het idee dat …."

10.5.2 ONVERWACHTE BEVINDINGEN

Onverwachte bevindingen werpen een ander licht op de literatuur die je hebt gebruikt om je onderzoek te motiveren. Er zijn grofweg twee soorten onverwachte bevindingen:

1) Wanneer je een resultaat hebt gevonden dat niet overeenkomt met je verwachtingen, of
2) Wanneer je iets tegenkomt waar je nog niet over na had gedacht.

In het eerste geval is het goed om te proberen te verklaren waarom de werkelijkheid anders blijkt te zijn dan je verwachtte. Dat doe je niet alleen op basis van je eigen verstand, maar het is ook de moeite waard om de literatuur in te duiken, om nogmaals te kijken of je niet iets gemist hebt of om te kijken of jouw intuïtieve verklaring ook te koppelen is aan eerder onderzoek. In beide gevallen kunnen jouw bevindingen een aanleiding vormen voor vervolgonderzoek.

10.5.3 BEPERKINGEN

Iedere studie heeft zijn eigen beperkingen. Eerder in dit hoofdstuk hebben we al besproken dat het er niet om gaat *of* een studie beperkt is, maar *welke* beperkingen een studie heeft. In de discussie kun je nogmaals reflecteren op de beperkingen van je studie en de mate waarin je die beperkingen ook terugziet in de resultaten. Naast de geanticipeerde beperkingen van je studie zijn er misschien ook onverwachte beperkingen. Misschien heb je bijvoorbeeld bij het coderen gemerkt dat de steekproef minder representatief is dan gedacht. In dat geval is het belangrijk om te bespreken waar het probleem vandaan komt (is het mogelijk te voorkomen?) en wat de implicaties zijn voor de generaliseerbaarheid van jouw studie. Tot slot kun je met de kennis van nu overwegen of het de moeite waard is om een vervolgstudie uit te voeren om de beperkingen van je huidige studie weg te nemen.

10.5.4 SUGGESTIES VOOR VERVOLGONDERZOEK

Ieder onderzoek bevat suggesties voor vervolgonderzoek. Het idee is dat je nadenkt over mogelijke vervolgstappen om voort te bouwen op hetgeen je gevonden hebt. (De suggesties zijn dus niet alleen gerelateerd aan jouw onderzoek, maar jouw onderzoek vormt ook de basis voor die vervolgstappen.) Als je suggesties doet voor vervolgonderzoek, zorg er dan voor dat je meer doet dan het noemen van een mogelijke onderzoeksrichting. Leg concreet uit hoe zo'n vervolgonderzoek eruit zou kunnen zien, waarom het relevant zou zijn om dat te bestuderen, en beschrijf indien mogelijk je verwachtingen (op basis van jullie bevindingen), of wat de implicaties zouden zijn van het voorgestelde onderzoek.

10.6 CONCLUSIE

In de conclusie vat je nogmaals kort samen wat de hoofdbevindingen zijn van je onderzoek, en wat de implicaties daarvan zijn. Daarnaast kun je ook aanbevelingen doen op basis van je bevindingen, waarbij het belangrijk is om bescheiden te blijven, en te zorgen dat de aanbevelingen proportioneel zijn. Met andere woorden: zorg dat je de implicaties van je onderzoek niet overdrijft.

10.7 VERANTWOORDING

Als je samen met anderen een wetenschappelijke tekst schrijft, dan staan de namen van alle auteurs bovenaan de tekst om te laten zien dat jullie de tekst samen hebben geschreven, en dat jullie samen garant staan voor de inhoud.[16] Tegelijkertijd is het ook logisch dat je onderling de taken verdeelt; je gaat niet met zijn allen tegelijkertijd schrijven aan de inleiding, maar je bespreekt vooraf wat er in het verslag moet staan en geeft iedereen een ander onderdeel om aan te werken. Achteraf kun je dan controleren wat de anderen geschreven hebben.

10.7.1 RAPPORTAGE VAN DE TAAKVERDELING: HET LOGBOEK

Bij de groepsopdracht in dit boek hoort ook een logboek. Dat is een grote tabel die beschrijft wie wat wanneer heeft gedaan. De basis voor het logboek is de taakverdeling die je maakt aan het begin van het project. Je kunt het volgende stappenplan gebruiken:

1) Maak duidelijke afspraken over de wijze waarop jullie willen samenwerken, deadlines willen halen, en elkaar van feedback willen voorzien. Spreek ook af wat de consequenties zijn als deze afspraken niet worden nagekomen. Deze afspraken leg je vast in een groepscontract.

16 Dat betekent bij opdrachten aan de universiteit ook dat alle auteurs verantwoordelijk zijn voor eventuele problemen met de tekst, zoals plagiaat of fouten in de rapportage. Iedereen is immers akkoord met de inhoud, en het verslag is ondertekend door alle auteurs. Lees dus altijd de volledige tekst!

2) Maak een lijst van taken die uitgevoerd moeten worden om het onderzoek succesvol af te ronden. (Bijlage D is hiervoor een goed vertrekpunt.) Deze lijst zal waarschijnlijk nog groeien tijdens het project, omdat er altijd taken zijn die je over het hoofd ziet.
3) Spreek af wie wat doet, voor welke deadline, en wie de hoofdverantwoordelijke is voor ieder onderdeel. De hoofdverantwoordelijke houdt in de gaten of er genoeg voortgang geboekt wordt om de deadline inderdaad te halen. Leg deze informatie vast in een document.
4) Laat iedereen bijhouden hoeveel tijd ze kwijt zijn aan het onderzoek, zodat je het latere werk eventueel ook kunt herverdelen als de eerdere verdeling niet eerlijk blijkt te zijn.

Een onderdeel is nooit de verantwoordelijkheid van slechts 1 of 2 personen; wees ervan bewust dat alle onderdelen van dit groepsproces en het rapport de verantwoordelijkheid is van jullie allemaal.

10.7.2 VOOR LATER: AUTEURSCHAP IN DE WETENSCHAP

Wanneer je een groepsopdracht schrijft voor een cursus aan de universiteit, dan is het meestal wel duidelijk van wie de namen op het verslag moeten staan: iedereen die heeft meegewerkt aan de opdracht. Als een groepslid niet meewerkt, dan kun je dat eerst onderling proberen op te lossen, en als dat niet helpt kun je naar de docent stappen om een oplossing te vinden. In de wetenschap is auteurschap genuanceerder. Vind je bijvoorbeeld dat de namen van de codeurs ook op het onderzoeksverslag moeten komen te staan, als ze daar niet aan meegeschreven hebben? Welke bijdragen betekenen dat je je "auteur" mag noemen, en welke bijdragen worden misschien alleen in het dankwoord genoemd? Het internationale comité voor redacteuren van medische tijdschriften (ICMJE) heeft hiervoor de volgende richtlijnen voor opgesteld (Nederlandse vertaling door Leeuw et al. (2014)):[17]

1) men heeft een substantiële bijdrage geleverd aan: het uitdenken en uitwerken van het beschreven onderzoek of aan het verwerven, analyseren of interpreteren van de onderzoeksgegevens; en
2) men heeft meegewerkt aan het schrijven en kritisch reviseren van het manuscript; en
3) men heeft zijn of haar goedkeuring gegeven aan de laatste versie van het manuscript; en
4) men stemt ermee in om volledige verantwoordelijkheid voor het manuscript te dragen en om ervoor te zorgen dat eventuele problemen met betrekking tot de accuraatheid of de integriteit van het beschreven werk of het manuscript naar behoren kunnen worden onderzocht en opgelost.

Daarnaast is het belangrijk om de *volgorde* van de auteursnamen te bespreken. Zo is bijvoorbeeld de eerste auteur het meest prominent wanneer een artikel geciteerd wordt: *Achternaam et al., 2024*. In veel vakgebieden hecht men waarde aan die eerste positie, omdat die persoon

17 Zie https://publicationethics.org voor verdere discussie over auteurschap.

dan ook wordt gezien als de drijvende kracht achter het artikel. Binnen de communicatiewetenschap wordt daarna de tweede auteur vaak gezien als de meest belangrijke auteur, terwijl in de medische wetenschappen de laatste auteur vaak belangrijker is (daar is de laatste auteur vaak de meest ervaren begeleider van het onderzoek). Maar er zijn ook andere opties. In de wiskunde worden auteursnamen bijvoorbeeld vaak in alfabetische volgorde geplaatst (Waltman, 2012), en sommige auteurs laten de volgorde bepalen door een muntje op te gooien (Miller & Ballard, 1992), een basketbalwedstrijd tussen de auteurs (Fauth & Resetarits, 1991), of een wedstrijdje steen-papier-schaar (Kupfer et al., 2004).[18] Hoe dan ook: als je na deze cursus samenwerkt met anderen aan een 'echte' publicatie, is het belangrijk om de volgorde vooraf te bespreken om te voorkomen dat er later vervelende discussies over ontstaan door verschillende verwachtingen van de teamleden over de auteursvolgorde.

Hoewel alle auteurs de verantwoordelijkheid dragen voor de kwaliteit hun werk, is het ook belangrijk om de verschillende soorten bijdragen van de auteurs te erkennen en waarderen. Steeds meer tijdschriften vragen auteurs om hun individuele bijdragen te omschrijven via de CRediT-taxonomie: een internationaal erkende categorisatie van mogelijke bijdragen die auteurs kunnen leveren aan wetenschappelijke publicaties (Brand et al., 2015). Deze ontwikkeling gaat gepaard met de groeiende aandacht voor *team science*: onderzoek doe je vrijwel nooit helemaal alleen, maar altijd in samenwerking met anderen. Aan het eind van wetenschappelijke publicaties, zoals die van Brand et al. (2015) zelf, staat er dan een *author statement* met een overzicht van de verschillende bijdragen en wie daar verantwoordelijk voor is.

10.8 ALGEMENE TIPS

Tot slot volgen hieronder nog wat algemene tips voor het schrijven van een verslag. De meeste tips zullen je niet verbazen, of misschien zelfs overbodig lijken. De ervaring leert echter dat veel studenten hier punten laten liggen.

1) Academisch schrijven is niet hetzelfde als "moeilijk taalgebruik." Het allerbelangrijkste is om duidelijk te zijn over wat je precies bedoelt. Het is niet de bedoeling om lezers te imponeren met ingewikkelde woorden of constructies, maar juist om complexe materie toegankelijk te maken voor een zo breed mogelijk publiek. Je weet dat je een goede tekst geschreven hebt als lezers instemmend meeknikken, en denken: *Natuurlijk! Dat kan ook niet anders.*
2) Let op je formulering, spelling, en grammatica. Als Word of een andere automatische schrijfhulp aangeeft dat iets fout is, is daar vaak een goede reden voor. Tegelijkertijd is het ook belangrijk om zelf na te blijven denken, en niet zomaar mee te gaan met de verbetersuggesties die deze programma's geven. Vraag groepsgenoten om te controleren wat

18 Met dank aan Meghan Duffy voor deze voorbeelden.

je geschreven hebt (en maak dit onderdeel van je taakverdeling – met duidelijke afspraken en interne deadlines kun je veel onenigheid voorkomen).

3) Wees voorzichtig met verwijswoorden. Wanneer je woorden zoals *het*, *die*, *deze* en andere verwijzende voornaamwoorden gebruikt, moet het voor de lezer duidelijk zijn waar die woorden naar verwijzen. Aan het begin van een alinea is het meestal beter om verwijswoorden te vermijden. In plaats van een frase als "dit idee" kun je beter iets zeggen als "het voorstel om iedereen gratis geld te geven" (of waar "dit idee" dan ook naar verwijst). Dan is het meteen duidelijk wat je bedoelt.

NUTTIGE BRONNEN

Om je verslag nog verder te verbeteren, kun je onder andere de volgende bronnen raadplegen:

- Clark, H. H. (2000). Everyone can write better (and you are no exception). In K. A. Keough & J. Garcia (Red.), *Social psychology of gender, race, and ethnicity.* McGraw-Hill New York, NY.
- Linger, N., & Stam, T. (2020). Academische uitdrukkingen [Gepubliceerd op de website van Faculteit der Geesteswetenschappen van de Vrije Universiteit, onder redactie van Nel de Jong en Margreet Onrust. Verkregen op 15 Oktober 2024]. https://www2.fgw.vu.nl/taalmateriaal/acad- uitdr/index.php
- The Writing Center. (g.d.). *Paragraphs* [Verkregen van de website van de *University of North Carolina at Chapel Hill*, geraadpleegd op 15 oktober 2024]. https://writingcenter.unc.edu/tips-and- tools/paragraphs/

BIBLIOGRAFIE

Albers, C., & Lakens, D. (2018). When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias. *Journal of Experimental Social Psychology*, *74*, 187–195. https://doi.org/https://doi.org/10.1016/j.jesp.2017.09.004

Antheunis, M. L., Schouten, A. P., Valkenburg, P. M., & Peter, J. (2011). Interactive uncertainty reduction strategies and verbal affection in computer-mediated communication. *Communication Research*, *39*(6), 757–780. https://doi.org/10.1177/0093650211410420

Arcas, B. A. y, Mitchell, M., & Todorov, A. (2023). Physiognomy in the age of AI. In J. Browne, S. Cave, E. Drage, & K. McInerney (Eds.), *Feminist AI: Critical perspectives on algorithms, data, and intelligent machines* (pp. 208–236). Oxford University Press.

Armstrong, C. L. (2004). The influence of reporter gender on source selection in newspaper stories. *Journalism & Mass Communication Quarterly*, *81*(1), 139–154. https://doi.org/10.1177/107769900408100110

Arts, A., Maes, A., Noordman, L. G. M., & Jansen, C. (2011). Overspecification in written instruction. *Linguistics*, *49*(3), 555–574. https://doi.org/doi:10.1515/ling.2011.017

Austin, J. L. (1962). *How to do things with words: The william james lectures delivered at harvard university in 1955*. Oxford University Press. https://archive.org/details/HowToDoThingsWithWordsAUSTIN

Autoriteit Persoonsgegevens. (n.d.). *De AVG in het kort*. https://www.autoriteitpersoonsgegevens.nl/themas/basis-avg/avg-algemeen/de-avg-in-het-kort

Ayers, J. W., Caputi, T. L., Nebeker, C., & Dredze, M. (2018). Don't quote me: Reverse identification of research participants in social media studies. *Npj Digital Medicine*, *1*(1). https://doi.org/10.1038/s41746-018-0036-2

Bagchi, C., Menczer, F., Tarafdar, M., Paik, A., & Grabowicz, P. A. (2024). Social media algorithms can curb misinformation, but do they? In *Science*. https://www.science.org/doi/10.1126/science.abp9364

Bandura, A. (1977). *Social learning theory*. Prentice-Hall.

Banga, A., Poelmans, P., Sweep, J., & Verhagen, V. (Eds.). (2022). *Handboek taalkunde* (1st ed.). Bussum: Uitgeverij Coutinho.

Bayerl, P. S., & Paul, K. I. (2011). What determines inter-coder agreement in manual annotations? A meta-analytic investigation. *Computational Linguistics*, *37*(4), 699–725. https://doi.org/10.1162/COLI_a_00074

Bell, A. (1984). Language style as audience design. *Language in Society*, *13*(2), 145–204. https://doi.org/10.1017/s004740450001037x

Berghel, H. (2018). Malice domestic: The cambridge analytica dystopia. *Computer*, *51*(05), 84–89. https://doi.org/10.1109/MC.2018.2381135

Bertolero, M. A., Dworkin, J. D., David, S. U., Lloreda, C. L., Srivastava, P., Stiso, J., Zhou, D., Dzirasa, K., Fair, D. A., Kaczkurkin, A. N., Marlin, B. J., Shohamy, D., Uddin, L. Q., Zurn, P., & Bassett, D. S. (2020). Racial and ethnic imbalance in neuroscience reference lists and intersections with gender. *bioRxiv*. https://doi.org/10.1101/2020.10.12.336230

Bird, S. (2020). Decolonising speech and language technology. In D. Scott, N. Bel, & C. Zong (Eds.), *Proceedings of the 28th international conference on computational linguistics* (pp. 3504–3519). International Committee on Computational Linguistics. https://doi.org/10.18653/v1/2020.coling-main.313

Bird, S., & Yibarbuk, D. (2024). Centering the speech community. In Y. Graham & M. Purver (Eds.), *Proceedings of the 18th conference of the european chapter of the association for computational linguistics (volume 1: Long papers)* (pp. 826–839). Association for Computational Linguistics. https://aclanthology.org/2024.eacl-long.50

Blassnig, S. (2023). Content analysis in the research field of political communication: The self-presentation of political actors. In F. Oehmer-Pedrazzi, S. H. Kessler, E. Humprecht, K. Sommer, & L. Castro (Eds.), *Standardisierte inhaltsanalyse in der kommunikationswissenschaft – standardized content analysis in communication research: Ein handbuch – a handbook* (pp. 301–312). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36179-2_26

Brand, A., Allen, L., Altman, M., Hlava, M., & Scott, J. (2015). Beyond authorship: Attribution, contribution, collaboration, and credit. *Learned Publishing*, *28*(2), 151–155. https://doi.org/https://doi.org/10.1087/20150211

Braun, N., Goudbeek, M., & Krahmer, E. (2019). Language and emotion – a foosball study: The influence of affective state on language production in a competitive setting. *PLOS ONE*, *14*(5), e0217419. https://doi.org/10.1371/journal.pone.0217419

Brunswik, E. (1955a). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, *62*(3), 193–217. https://search.ebscohost.com/login.aspx?direct=true&db=pdh&AN=1956-00035-001&site=ehost-live

Brunswik, E. (1955b). The conceptual framework of psychology. In O. Neurath, R. Carnap, & C. Morris (Eds.), *International encyclopedia of unified science. Volume i, part 2*. University of Chicago Press. https://archive.org/details/B-001-015-450/page/n327/mode/2up

Bryan, J. (2015). *Naming things*. https://speakerdeck.com/jennybc/how-to-name-files

Burgers, C., Konijn, E. A., & Steen, G. J. (2016). Figurative framing: Shaping public discourse through metaphor, hyperbole, and irony. *Communication Theory*, *26*(4), 410–430. https://doi.org/10.1111/comt.12096

Burgers, C., Mulken, M. van, & Schellens, P. J. (2011). Finding irony: An introduction of the verbal irony procedure (VIP). *Metaphor and Symbol*, *26*(3), 186–205. https://doi.org/10.1080/10926488.2011.583194

Camerer, C., Loewenstein, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy*, *97*(5), 1232–1254. https://doi.org/10.1086/261651

Cardoso, B., & Cohn, N. (2022). The multimodal annotation software tool (MAST). In N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the thirteenth language resources and evaluation conference* (pp. 6822–6828). European Language Resources Association. https://aclanthology.org/2022.lrec-1.736

Chavez, C. (2008). Conceptualizing from the inside: Advantages, complications, and demands on insider positionality. *The Qualitative Report*, *13*(3), 474–494. https://doi.org/10.46743/2160-3715/2008.1589

Clark, H. H. (2000). Everyone can write better (and you are no exception). In K. A. Keough & J. Garcia (Eds.), *Social psychology of gender, race, and ethnicity* (p. 379). McGraw-Hill New York, NY.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Commissies Levelt, Noort, en Drenth. (2012). *Falende wetenschap: De frauduleuze onderzoekspraktijken van sociaal-psycholoog diederik stapel*. Koninklijke Nederlandse Academie van Wetenschappen.

Courtney, A. E., & Whipple, T. W. (1983). *Sex stereotyping in advertising*. Lexington Books. https://books.google.nl/books?id=ZAqSAAAAIAAJ

Devaraj, S., Quigley, N. R., & Patel, P. C. (2018). The effects of skin tone, height, and gender on earnings. *PLOS ONE*, *13*(1), e0190640. https://doi.org/10.1371/journal.pone.0190640

Don, J., Meyer, C., & Rispens, J. (Eds.). (2023). *Taal en taalwetenschap* (3rd ed.). Wiley.

Doreleijers, K. (2024). *Styling the local: Hyperdialectisms and the enregisterment of the gender suffix in the "new" dialect of north brabant*. Landelijke Onderzoekschool Taalwetenschap.

Eisner, J. (2010). *Write the paper first*. https://www.cs.jhu.edu/~jason/advice/write-the-paper-first.html

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.

Esau, K. (2023). Content analysis in the research field of incivility and hate speech in online communication. In F. Oehmer-Pedrazzi, S. H. Kessler, E. Humprecht, K. Sommer, & L.

Castro (Eds.), *Standardisierte inhaltsanalyse in der kommunikationswissenschaft – standardized content analysis in communication research: Ein handbuch – a handbook* (pp. 451–461). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36179-2_38

Fauth, J. E., & Resetarits, W. J. (1991). Interactions between the salamander siren intermedia and the keystone predator notophthalmus viridescens. *Ecology*, *72*(3), 827–838. http://www.jstor.org/stable/1940585

Fiesler, C., Beard, N., & Keegan, B. C. (2020). No robots, spiders, or scrapers: Legal and ethical regulation of data collection methods in social media terms of service. *Proceedings of the International AAAI Conference on Web and Social Media*, *14*, 187–196. https://doi.org/10.1609/icwsm.v14i1.7290

Fiesler, C., & Proferes, N. (2018). "Participant" perceptions of twitter research ethics. *Social Media + Society*, *4*(1), 2056305118763366. https://doi.org/10.1177/2056305118763366

Fleisig, E., Blodgett, S. L., Klein, D., & Talat, Z. (2024). The perspectivist paradigm shift: Assumptions and challenges of capturing human labels. In K. Duh, H. Gomez, & S. Bethard (Eds.), *Proceedings of the 2024 conference of the north american chapter of the association for computational linguistics: Human language technologies (volume 1: Long papers)* (pp. 2279–2292). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.naacl-long.126

Fouts, G., & Burggraf, K. (2000). Television situation comedies: Female weight, male negative comments, and audience reactions. *Sex Roles*, *42*(9), 925–932. https://doi.org/10.1023/A:1007054618340

Freitas Netto, S. V. de, Sobral, M. F. F., Ribeiro, A. R. B., & Soares, G. R. da L. (2020). Concepts and forms of greenwashing: A systematic review. *Environmental Sciences Europe*, *32*(1), 19. https://doi.org/10.1186/s12302-020-0300-3

Fulvio, J. M., Akinnola, I., & Postle, B. R. (2021). Gender (im)balance in citation practices in cognitive neuroscience. *Journal of Cognitive Neuroscience*, *33*(1), 3–7. https://doi.org/10.1162/jocn_a_01643

Giles, H., Taylor, D. M., & Bourhis, R. (1973). Towards a theory of interpersonal accommodation through language: Some canadian data. *Language in Society*, *2*(2), 177–192. https://doi.org/10.1017/s0047404500000701

Gillespie, T. (2013). The relevance of algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media technologies: Essays on communication, materiality, and society* (pp. 167–193). The MIT Press.

Gilly, M. C. (1988). Sex roles in advertising: A comparison of television advertisements in australia, mexico, and the united states. *Journal of Marketing*, *52*(2), 75–85. https://doi.org/10.1177/002224298805200206

Gottfried, J. (2024). *Americans' social media use*. Pew Research Center. https://www.pewresearch.org/internet/2024/01/31/americans-social-media-use/

Gray, M. L., & Suri, S. (2019). *Ghost work: How to stop silicon valley from building a new global underclass*. Houghton Mifflin Harcourt.

Gruzd, A., Mai, P., & Vahedi, Z. (2022). Studying anti-social behaviour on reddit with communalytic. In A. Quan-Haase & L. Sloan (Eds.), *The SAGE handbook of social media research methods* (pp. 503–520). SAGE Publications Ltd. https://doi.org/10.4135/9781529782943.n36

Guess, A. M., Malhotra, N., Pan, J., Barberá, P., Allcott, H., Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gentzkow, M., González-Bailón, S., Kennedy, E., Kim, Y. M., Lazer, D., Moehler, D., Nyhan, B., Rivera, C. V., Settle, J., Thomas, D. R., … Tucker, J. A. (2023). How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science*, *381*(6656), 398–404. https://doi.org/10.1126/science.abp9364

Guidry, J. P. D., Carlyle, K., Messner, M., & Jin, Y. (2015). On pins and needles: How vaccines are portrayed on pinterest. *Vaccine*, *33*(39), 5051–5056. https://doi.org/https://doi.org/10.1016/j.vaccine.2015.08.064

Gwet, K. L. (2014). *Handbook of inter-rater reliability, 4th edition: The definitive guide to measuring the extent of agreement among raters*. Advanced Analytics, LLC.

Hall, J. A., Horgan, T. G., & Murphy, N. A. (2019). Nonverbal communication [Journal Article]. *Annual Review of Psychology*, *70*(Volume 70, 2019), 271–294. https://doi.org/https://doi.org/10.1146/annurev-psych-010418-103145

Hall, J. A., & Knapp, M. L. (Eds.). (2013). *Nonverbal communication*. De Gruyter Mouton. https://doi.org/doi:10.1515/9783110238150

Harper, C. A. (2020). Ideological measurement in social and political psychology. *PsyArXiv*. https://doi.org/https://doi.org/10.31234/osf.io/wpsje

Haspelmath, M. (2023). Defining the word. *WORD*, *69*(3), 283–297. https://doi.org/10.1080/00437956.2023.2237272

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77–89. https://doi.org/10.1080/19312450709336664

Hinds, J., Williams, E. J., & Joinson, A. N. (2020). "It wouldn't happen to me": Privacy concerns and perspectives following the cambridge analytica scandal. *International Journal of Human-Computer Studies*, *143*, 102498. https://doi.org/https://doi.org/10.1016/j.ijhcs.2020.102498

Huibers, J., & Verhoeven, J. (2014). Webcare als online reputatiemanagement [Journal Article]. *Tijdschrift Voor Communicatiewetenschap*, *42*(2). https://doi.org/https://doi.org/10.5117/2014.042.002.165

Hunter, M. (2007). The persistent problem of colorism: Skin tone, status, and inequality. *Sociology Compass*, *1*(1), 237–254. https://doi.org/https://doi.org/10.1111/j.1751-9020.2007.00006.x

Hussey, I. (2023). Data is not available upon request. *PsyArXiv*. https://doi.org/https://doi.org/10.31234/osf.io/jbu9r

Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

Jamieson, M. K., Govaart, G. H., & Pownall, M. (2023). Reflexivity in quantitative research: A rationale and beginner's guide. *Social and Personality Psychology Compass*, *17*(4), e12735. https://doi.org/https://doi.org/10.1111/spc3.12735

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus; Giroux.

Kaur, T. (2024). Nicki minaj is beefing with a naughty dog dev, and i'm on the dev's side. *The Gamer*. https://www.thegamer.com/nicki-minaj-beefing-naughty-dog-dev-ai-generated-images/

Kincaid, J. P., Jr., R. P. F., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count, and flesch reading ease formula) for navy enlisted personnel* (56). Institute for Simulation; Training. https://stars.library.ucf.edu/istlibrary/56

Kirk, H., Birhane, A., Vidgen, B., & Derczynski, L. (2022). Handling and presenting harmful text in NLP research. In Y. Goldberg, Z. Kozareva, & Y. Zhang (Eds.), *Findings of the association for computational linguistics: EMNLP 2022* (pp. 497–510). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.findings-emnlp.35

Kleinnijenhuis, J., & Atteveldt, W. van. (2006). Geautomatiseerde inhoudsanalyse, met de berichtgeving over het EU-referendum als voorbeeld. In F. Wester (Ed.), *Inhoudsanalyse: Theorie en praktijk*. Kluwer.

KNAW, NFU, NWO, TO2-Federatie, Vereniging Hogescholen, & VSNU. (2018). *Nederlandse gedragscode wetenschappelijke integriteit*. Data Archiving; Networked Services (DANS). https://doi.org/10.17026/DANS-2CJ-NVWU

Kraf, R., Lentz, L., & Pander Maat, H. (2011). Drie nederlandse instrumenten voor het automatisch voorspellen van begrijpelijkheid – een klein consumentenonderzoek [Journal Article]. *Tijdschrift Voor Taalbeheersing*, *33*(3), 249–265. https://doi.org/https://doi.org/10.5117/TVT2011.3.DRIE411

Krawczyk, M., & Reuben, E. (2012). (Un)available upon request: Field experiment on researchers' willingness to share supplementary materials. *Accountability in Research*, *19*(3), 175–186. https://doi.org/10.1080/08989621.2012.678688

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. SAGE Publications.

Kupfer, J. A., Webbeking, A. L., & Franklin, S. B. (2004). Forest fragmentation affects early successional patterns on shifting cultivation fields near indian church, belize. *Agriculture, Ecosystems & Environment*, *103*(3), 509–518. https://doi.org/https://doi.org/10.1016/j.agee.2003.11.011

L. Haven, T., & Van Grootel, Dr. L. (2019). Preregistering qualitative research. *Accountability in Research*, *26*(3), 229–244. https://doi.org/10.1080/08989621.2019.1580147

Lacy, S., & Riffe, D. (1996). Sampling error and selecting intercoder reliability samples for nominal content categories. *Journalism & Mass Communication Quarterly*, *73*(4), 963–973. https://doi.org/10.1177/107769909607300414

Lakens, D. (2022). Sample Size Justification. *Collabra: Psychology*, *8*(1), 33267. https://doi.org/10.1525/collabra.33267

Lakens, D. (2023). Is my study useless? Why researchers need methodological review boards. *Nature*, *613*(7942), 9. https://doi.org/10.1038/d41586-022-04504-8

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. http://www.jstor.org/stable/2529310

Larson, B. (2017). Gender as a variable in natural-language processing: Ethical considerations. In D. Hovy, S. Spruit, M. Mitchell, E. M. Bender, M. Strube, & H. Wallach (Eds.), *Proceedings of the first ACL workshop on ethics in natural language processing* (pp. 1–11). Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-1601

Lazard, L., & McAvoy, J. (2017). Doing reflexivity in psychological research: What's the point? What's the practice? *Qualitative Research in Psychology*, *17*(2), 159–177. https://doi.org/10.1080/14780887.2017.1400144

Leeuw, P. W. de, Rosenberg, J., & Reyes, H. (2014). Nieuwe richtlijnen vanuit de ICMJE. *Nederlands Tijdschrift Voor Geneeskunde*, *158*(A7602), 1–2.

Lehner, P. E., Adelman, L., Cheikes, B. A., & J. Brown, M. (2008). Confirmation bias in complex analyses. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, *38*(3), 584–592. https://doi.org/10.1109/TSMCA.2008.918634

Liebrecht, C. (2015). *Intens krachtig. Stilistische intensiveerders in evaluatieve teksten* [PhD thesis]. Radboud Universiteit.

Liesenfeld, A., Lopez, A., & Dingemanse, M. (2023). Opening up ChatGPT: Tracking openness, transparency, and accountability in instruction-tuned text generators. *Proceedings of the 5th International Conference on Conversational User Interfaces*. https://doi.org/10.1145/3571884.3604316

Lin, C. A. (1997). Beefcake versus cheesecake in the 1990s: Sexist portrayals of both genders in television commercials. *Howard Journal of Communications*, *8*(3), 237–249. https://doi.org/10.1080/10646179709361757

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer vision – ECCV 2014* (pp. 740–755). Springer International Publishing.

Linger, N., & Stam, T. (2020). *Academische uitdrukkingen*. https://www2.fgw.vu.nl/taalmateriaal/acad-uitdr/index.php

Lischka, J. A. (2023). Content analysis in the research field of corporate communication. In F. Oehmer-Pedrazzi, S. H. Kessler, E. Humprecht, K. Sommer, & L. Castro (Eds.), *Standardisierte inhaltsanalyse in der kommunikationswissenschaft – standardized content analysis in com-*

munication research: Ein handbuch – a handbook (pp. 349–361). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36179-2_30

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, *28*(4), 587–604. https://doi.org/https://doi.org/10.1111/j.1468-2958.2002.tb00826.x

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2010). *Practical resources for assessing and reporting intercoder reliability in content analysis research projects*.

Looman, B. (2024). In toga op de barricade. *Onderzoek*, *15*, 2–4. https://www.nwo.nl/sites/nwo/files/media-files/onderzoek_15_voorjaar_2024_versie_voor_de_website.pdf

Maat, H. P., Kleijn, S., & Frissen, S. (2023). LiNT: Een leesbaarheidsformule en een leesbaarheidsinstrument [Journal Article]. *Tijdschrift Voor Taalbeheersing*, *45*(1), 2–39. https://doi.org/https://doi.org/10.5117/TVT2023.3.002.MAAT

Marwick, A. E., & boyd, danah. (2014). Networked privacy: How teenagers negotiate context in social media. *New Media & Society*, *16*(7), 1051–1067. https://doi.org/10.1177/1461444814543995

Mauri, M., Elli, T., Caviglia, G., Uboldi, G., & Azzi, M. (2017, September). RAWGraphs: A visualisation platform to create open outputs. *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter*. https://doi.org/10.1145/3125571.3125585

McCambridge, J., Witton, J., & Elbourne, D. R. (2014). Systematic review of the hawthorne effect: New concepts are needed to study research participation effects. *Journal of Clinical Epidemiology*, *67*(3), 267–277. https://doi.org/https://doi.org/10.1016/j.jclinepi.2013.08.015

McCook, A. (2016). Publicly available data on thousands of okcupid users pulled over copyright claim. In *Retraction Watch*. https://retractionwatch.com/2016/05/16/publicly-available-data-on-thousands-of-okcupid-users-pulled-over-copyright-claim/

McGrady, R., McGrady, R., Zheng, K., Curran, R., Baumgartner, J., & Zuckerman, E. (2023). Dialing for videos: A random sample of YouTube. *Journal of Quantitative Description: Digital Media*, *3*. https://doi.org/10.51685/jqd.2023.022

Merton, R. K. (1968). The matthew effect in science. *Science*, *159*(3810), 56–63. https://doi.org/10.1126/science.159.3810.56

Michiels, L., Leysen, J., Smets, A., & Goethals, B. (2022, July). What are filter bubbles really? A review of the conceptual and empirical work. *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*. https://doi.org/10.1145/3511047.3538028

Miller, S. D., & Ballard, W. B. (1992). In my experience: Analysis of an effort to increase moose calf survivorship by increased hunting of brown bears in south-central alaska. *Wildlife Society Bulletin (1973-2006)*, *20*(4), 445–454. http://www.jstor.org/stable/3783068

Mizrahi, M., Kaplan, G., Malkin, D., Dror, R., Shahaf, D., & Stanovsky, G. (2024). State of What Art? A Call for Multi-Prompt LLM Evaluation. *Transactions of the Association for Computational Linguistics*, *12*, 933–949. https://doi.org/10.1162/tacl_a_00681

Moeller, J., Helberger, N., et al. (2018). *Beyond the filter bubble: Concepts, myths, evidence and issues for future debates*. University of Amsterdam. https://www.ivir.nl/publicaties/download/Beyond_the_filter_bubble__concepts_myths_evidence _and_issues_for_future_debates.pdf

Mohammad, S. M. (2020). Practical and ethical considerations in the effective use of emotion and sentiment lexicons. *CoRR*, *abs/2011.03492*. https://arxiv.org/abs/2011.03492

Mui, P. H. C., Goudbeek, M. B., Swerts, M. G. J., & Hovasapian, A. (2017). Children's nonverbal displays of winning and losing: Effects of social and cultural contexts on smiles. *Journal of Nonverbal Behavior*, *41*(1), 67–82. https://doi.org/10.1007/s10919-016-0241-0

Naab, T. K., & Küchler, C. (2023). Content analysis in the research field of online user comments. In F. Oehmer-Pedrazzi, S. H. Kessler, E. Humprecht, K. Sommer, & L. Castro (Eds.), *Standardisierte inhaltsanalyse in der kommunikationswissenschaft – standardized content analysis in communication research: Ein handbuch – a handbook* (pp. 441–450). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36179-2_37

Nanne, A. J., Antheunis, M. L., Lee, C. G. V. D., Postma, E. O., Wubben, S., & Noort, G. V. (2020). The use of computer vision to analyze brand-related user generated image content. *Journal of Interactive Marketing*, *50*(1), 156–167. https://doi.org/10.1016/j.intmar.2019.09.003

Navarro, D. (2022). *Project structure* (Version v1.0.0). Zenodo. https://doi.org/10.5281/zenodo.6976003

Neuendorf, K. A. (2017). *The content analysis guidebook*. SAGE Publications, Inc. https://doi.org/10.4135/9781071802878

Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press. https://books.google.nl/books?id=_NN1uGn1Jd8C

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Oehmer-Pedrazzi, F., Kessler, S. H., Humprecht, E., Sommer, K., & Castro, L. (Eds.). (2023). *Standardisierte inhaltsanalyse in der kommunikationswissenschaft – standardized content analysis in communication research: Ein handbuch – a handbook* (1st ed.). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36179-2

Osterholz, S., Breil, S. M., Nestler, S., & Back, M. D. (2021). Lens and dual lens models. In *The Oxford Handbook of Accurate Personality Judgment*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190912529.013.4

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., …

Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*.

Pander Maat, H., & Dekker, N. (2016). Tekstgenres analyseren op lexicale complexiteit met t-scan [Journal Article]. *Tijdschrift Voor Taalbeheersing*, *38*(3), 263–304. https://doi.org/https://doi.org/10.5117/TVT2016.3.PAND

Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. Penguin Books Limited. https://books.google.nl/books?id=-FWOopuw3nYC

Pauwels, A. (1953). *De plaats van hulpwerkwoord verleden deelwoord en infinitief in de nederlandse bijzin*. M. & L. Symons, Leuven.

Payne, B. H., Taylor, J., Spiel, K., & Fiesler, C. (2023, October). How to ethically engage fat people in HCI research. *Computer Supported Cooperative Work and Social Computing*. https://doi.org/10.1145/3584931.3606987

Peeters, S., & Hagen, S. (2022). The 4CAT capture and analysis toolkit: A modular tool for transparent and traceable social media research [Journal Article]. *Computational Communication Research*, *4*(2), 571–589. https://doi.org/https://doi.org/10.5117/CCR2022.2.007.HAGE

Probst, A. K. (2023). *Scientific practice regarding visual accessibility of research papers: Pitfalls and recommendations* [PhD thesis, Tilburg University. Communication; Cognition]. http://arno.uvt.nl/show.cgi?fid=162118

Proferes, N., Jones, N., Gilbert, S., Fiesler, C., & Zimmer, M. (2021). Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics. *Social Media + Society*, *7*(2), 205630512110190. https://doi.org/10.1177/20563051211019004

Qi, P., Zhang, Y., Zhang, Y., Bolton, J., & Manning, C. D. (2020). Stanza: A python natural language processing toolkit for many human languages. In A. Celikyilmaz & T.-H. Wen (Eds.), *Proceedings of the 58th annual meeting of the association for computational linguistics: System demonstrations* (pp. 101–108). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-demos.14

Qutteina, Y., Hallez, L., Mennes, N., De Backer, C., & Smits, T. (2019). What do adolescents see on social media? A diary study of food marketing images on social media. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02637

Riffe, D., Lacy, S., Watson, B. R., & Lovejoy, J. (2023). *Analyzing media messages: Using quantitative content analysis in research*. Routledge. https://doi.org/10.4324/9781003288428

Rogers, A., Kovaleva, O., & Rumshisky, A. (2019). Calls to action on social media: Detection, social impact, and censorship potential. In A. Feldman, G. Da San Martino, A. Barrón-Cedeño, C. Brew, C. Leberknight, & P. Nakov (Eds.), *Proceedings of the second workshop on natural language processing for internet freedom: Censorship, disinformation, and propaganda* (pp. 36–44). Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-5005

Rudy, R. M., Popova, L., & Linz, D. G. (2010). The context of current content analysis of gender roles: An introduction to a special issue. *Sex Roles*, *62*(11), 705–720. https://doi.org/10.1007/s11199-010-9807-1

Rudy, R. M., Popova, L., & Linz, D. G. (2011). Contributions to the content analysis of gender roles: An introduction to a special issue. *Sex Roles*, *64*(3), 151–159. https://doi.org/10.1007/s11199-011-9937-0

Saenger, G. (1955). Male and Female Relations in the American Comic Strip. *Public Opinion Quarterly*, *19*(2), 195–205. https://doi.org/10.1086/266561

Samson-Himmelstjerna, C. von. (2023). Content analysis in the research field of strategic health communication. In F. Oehmer-Pedrazzi, S. H. Kessler, E. Humprecht, K. Sommer, & L. Castro (Eds.), *Standardisierte inhaltsanalyse in der kommunikationswissenschaft – standardized content analysis in communication research: Ein handbuch – a handbook* (pp. 399–410). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36179-2_34

Santos, C., Coelho, A., & Marques, A. (2024). A systematic literature review on greenwashing and its relationship to stakeholders: State of art and future research agenda. *Management Review Quarterly*, *74*(3), 1397–1421. https://doi.org/10.1007/s11301-023-00337-5

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity*, *52*(4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8

Scholman, M. C. J., Evers-Vermeul, J., & Sanders, T. J. M. (2016). Categories of coherence relations in discourse annotation. *Dialogue & Discourse*, *7*(2), 1–28. https://doi.org/10.5087/dad.2016.201

Schwabish, J., Popkin, S. J., & Feng, A. (2022). *Do no harm guide: Centering accessibility in data visualization*. Urban Institute.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press. https://books.google.nl/books?id=4UKbAAAAQBAJ

Sen, M., & Wasow, O. (2016). Race as a bundle of sticks: Designs that estimate effects of seemingly immutable characteristics [Journal Article]. *Annual Review of Political Science*, *19*(Volume 19, 2016), 499–522. https://doi.org/https://doi.org/10.1146/annurev-polisci-032015-010015

Signorielli, N. (2010). Research ethics in content analysis. In A. B. Jordan, D. Kunkel, J. Manganello, & M. Fishbein (Eds.), *Media messages and public health* (pp. 106–114). Routledge. https://doi.org/10.4324/9780203887349-12

Sinclair, J. (2004). Corpus and text: Basic priniciples. In M. Wynne (Ed.), *Developing linguistic corpora: A guide to good practice*. AHDS – Literature, Languages; Linguistics. https://users.ox.ac.uk/~martinw/dlc/chapter1.htm

Spiegelman, M., Terwilliger, C., & Fearing, F. (1953). The content of comics: Goals and means to goals of comic strip characters. *The Journal of Social Psychology*, *37*(2), 189–203.

Spooren, W., & Degand, L. (2010). Coding coherence relations: Reliability and validity. *Corpus Linguistics and Linguistic Theory*, *6*(2), 241–266. https://doi.org/doi:10.1515/cllt.2010.009

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspect Psychol Sci*, *11*(5), 702–712. https://doi.org/10.1177/1745691616658637

Stortenbeker, I., Salm, L., Hartman, T. olde, Stommel, W., Das, E., & Dulmen, S. van. (2022). Coding linguistic elements in clinical interactions: A step-by-step guide for analyzing communication form. *BMC Medical Research Methodology*, *22*(1). https://doi.org/10.1186/s12874-022-01647-0

Swerts, M., Van Heteren, A., Nieuwdorp, C., Von Oerthel, E., & Kloots, H. (2021). Asymmetric forms of linguistic adaptation in interactions between flemish and dutch speakers. *Frontiers in Communication*, *6*. https://doi.org/10.3389/fcomm.2021.716444

Tatman, R. (2018). *What you can, can't and shouldn't do with social media data*. Gepubliceerd via: https://makingnoiseandhearingthings.com/2018/08/31/what-you-can-cant-and-shouldnt-do-with-social-media-data/.

Tausczik, Y. R., & Pennebaker, J. W. (2009). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*(1), 24–54. https://doi.org/10.1177/0261927x09351676

Tedersoo, L., Kãngas, R., Oras, E., Kãster, K., Eenmaa, H., Leijen, Ã., Pedaste, M., Raju, M., Astapova, A., Lukner, H., Kogermann, K., & Sepp, T. (2021). Data sharing practices and data availability upon request differ across scientific disciplines. *Scientific Data*, *8*(1). https://doi.org/10.1038/s41597-021-00981-0

Templeton, A. R. (2013). Biological races in humans. *Stud Hist Philos Biol Biomed Sci*, *44*(3), 262–271. https://doi.org/10.1016/j.shpsc.2013.04.010

The Writing Center. (n.d.). *Paragraphs*. https://writingcenter.unc.edu/tips-and-tools/paragraphs/

Thorp, H. H., & Vinson, V. (2024). Context matters in social media. *Science*, *385*(6716), 1393–1393. https://doi.org/10.1126/science.adt2983

Tiziana Schopper, A. B., & Vogelgsang, L. (2024). Pride or rainbow-washing? Exploring LGBTQ+ advertising from the vested stakeholder perspective. *Journal of Advertising*, *0*(0), 1–18. https://doi.org/10.1080/00913367.2024.2317147

Tolentino, D. (2022). This tiktoker is "consensually doxxing" people to teach them about social media privacy. *NBCnews.com*. https://www.nbcnews.com/pop-culture/tiktoker-consensually-doxxing-people-teach-social-media-privacy-rcna55037

Treadwell, D. (2023). *Introducing communication research: Paths of inquiry* (5th ed.). SAGE Publications, Inc.

Trepte, S. (2020). The Social Media Privacy Model: Privacy and Communication in the Light of Social Media Affordances. *Communication Theory*, *31*(4), 549–570. https://doi.org/10.1093/ct/qtz035

Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 505–514. https://doi.org/10.1609/icwsm.v8i1.14517

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Van Atteveldt, W., Trilling, D., & Calderon, C. A. (2022). *Computational analysis of communication*. Wiley-Blackwell. https://cssbook.net

Van der Velden, M. A. C. G., & Loecherbach, F. (2023). Content analysis in the research field of political coverage. In F. Oehmer-Pedrazzi, S. H. Kessler, E. Humprecht, K. Sommer, & L. Castro (Eds.), *Standardisierte inhaltsanalyse in der kommunikationswissenschaft – standardized content analysis in communication research: Ein handbuch – a handbook* (pp. 85–97). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36179-2_8

Van der Zanden, T., Mos, M., & Schouten, A. (2018). Taalaccommodatie in online datingprofielen [Journal Article]. *Tijdschrift Voor Taalbeheersing*, *40*(1), 83–106. https://doi.org/https://doi.org/10.5117/TVT2018.1.zand

Van Enschot, R., Spooren, W., Bosch, A. van den, Burgers, C., Degand, L., Evers-Vermeul, J., Kunneman, F., Liebrecht, C., Linders, Y., & Maes, A. (2024). Taming our wild data: On intercoder reliability in discourse research. *Dutch Journal of Applied Linguistics*, *13*. https://doi.org/10.51751/dujal16248

Van Hintum, M. (2024). Banden verbroken. *Onderzoek*, *15*(8–10). https://www.nwo.nl/sites/nwo/files/media-files/onderzoek_15_voorjaar_2024_versie_voor_de_website.pdf

Van Hooijdonk, C., & Liebrecht, C. (2018). "Wat vervelend dat de fiets niet is opgeruimd! Heb je een zaaknummer voor mij? ^EK" [Journal Article]. *Tijdschrift Voor Taalbeheersing*, *40*(1), 45–81. https://doi.org/https://doi.org/10.5117/TVT2018.1.hooi

Van Hooijdonk, C., & Liebrecht, C. (2021). Sorry but no sorry: The use and effects of apologies in airline webcare responses to NeWOM messages of flight passengers. *Discourse, Context & Media*, *40*, 100442. https://doi.org/https://doi.org/10.1016/j.dcm.2020.100442

Van Maanen, G. (2020). Ethics washing: Een introductie [Journal Article]. *Algemeen Nederlands Tijdschrift Voor Wijsbegeerte*, *112*(4), 462–467. https://doi.org/https://doi.org/10.5117/ANTW2020.4.020.VANM

Voigt, P., & Bussche, A. von dem. (2017). *The EU general data protection regulation (GDPR)*. Springer International Publishing. https://doi.org/10.1007/978-3-319-57959-7

Vromans, R. D., van Eenbergen, M. C., Pauws, S. C., Geleijnse, G., van der Poel, H. G., van de Poll-Franse, L. V., & Krahmer, E. J. (2019). Communicative aspects of decision aids for localized prostate cancer treatment – a systematic review. *Urologic Oncology: Seminars and Original Investigations*, *37*(7), 409–429. https://doi.org/https://doi.org/10.1016/j.urolonc.2019.04.005

Waltman, L. (2012). An empirical analysis of the use of alphabetical authorship in scientific publishing. *CoRR*, *abs/1206.4863*. http://arxiv.org/abs/1206.4863

Ware, O. R., Dawson, J. E., Shinohara, M. M., & Taylor, S. C. (2020). Racial limitations of fitzpatrick skin type. *Cutis*, *105*(2), 77–80.

Wester, F. P. J. (Ed.). (2006). *Inhoudsanalyse: Theorie en praktijk*. Kluwer.

Wicke, N. (2023). Content analysis in the research field of science communication. In F. Oehmer-Pedrazzi, S. H. Kessler, E. Humprecht, K. Sommer, & L. Castro (Eds.), *Standardisierte inhaltsanalyse in der kommunikationswissenschaft – standardized content analysis in communication research: Ein handbuch – a handbook* (pp. 411–425). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36179-2_35

Williams, D. N., & Williams, K. A. (2020). Sample size considerations: Basics for preparing clinical or basic research. *Annals of Nuclear Cardiology*, *6*(1), 81–85. https://doi.org/10.17996/anc.20-00122

Wilson, C., Janes, G., & Williams, J. (2022). Identity, positionality and reflexivity: Relevance and application to research paramedics. *British Paramedic Journal*, *7*(2), 43–49. https://doi.org/10.29045/14784726.2022.09.7.2.43

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. *CoRR*, *abs/1308.5499*. http://arxiv.org/abs/1308.5499

Winter, B. (2019). *Statistics for linguists: An introduction using r*. Taylor & Francis. https://books.google.nl/books?id=8cbADwAAQBAJ

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In N. Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odijk, & D. Tapias (Eds.), *Proceedings of the fifth international conference on language resources and evaluation (LREC'06)*. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2006/pdf/153_pdf.pdf

Zhang, W., Deng, Y., Liu, B., Pan, S., & Bing, L. (2024). Sentiment analysis in the era of large language models: A reality check. In K. Duh, H. Gomez, & S. Bethard (Eds.), *Findings of the association for computational linguistics: NAACL 2024* (pp. 3881–3906). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.findings-naacl.246

Zhang, Y., Zhang, Y., Qi, P., Manning, C. D., & Langlotz, C. P. (2021). Biomedical and clinical English model packages for the Stanza Python NLP library. *Journal of the American Medical Informatics Association*, *28*(9), 1892–1899. https://doi.org/10.1093/jamia/ocab090

Zom, D.-J., & Van der Deijl, E. (2024). Tussen onderzoek en opinie. *Onderzoek*, *15*, 11–13. https://www.nwo.nl/sites/nwo/files/media-files/onderzoek_15_voorjaar_2024_versie_voor_de_website.pdf

Zook, M., Barocas, S., boyd, danah, Crawford, K., Keller, E., Gangadharan, S. P., Goodman, A., Hollander, R., Koenig, B. A., Metcalf, J., Narayanan, A., Nelson, A., & Pasquale, F. (2017). Ten simple rules for responsible big data research. *PLOS Computational Biology*, *13*(3), e1005399. https://doi.org/10.1371/journal.pcbi.1005399

BIJLAGEN

A

Terminologie

Hieronder vind je een begrippenlijst met veelgebruikte termen.

Agreement
: Overeenstemming tussen twee of meer codeurs. Om dit te kwantificeren wordt vaak de procentuele overeenstemming (Engels *percentage agreement*) gerapporteerd: als twee codeurs allebei dezelfde items coderen, kijk je hoevaak ze dat hetzelfde hebben gedaan. De formule voor procentuele overeenstemming is dan: (Aantal hetzelfde gecodeerd/Totaal aantal items * 100). Andere maten die iets zeggen over de overeenstemming tussen codeurs zijn **Cohen's Kappa** en **Krippendorff's alpha**.

Anonimisatie
: Anonimiseren betekent dat je de identificerende gegevens van iemand *verwijdert*, zodat de data niet meer aan die persoon gekoppeld kunnen worden.

Automatisch coderen
: Wanneer de verschillende units in een corpus door de computer worden gecategoriseerd, zonder tussenkomst van menselijke **codeurs**, spreken we van automatisch **coderen**.

Beeldherkenning
: Het automatisch herkennen van entiteiten en eigenschappen binnen een afbeelding wordt beeldherkenning genoemd.

Categorie
: Iedere variabele die bestudeerd wordt in een inhoudsanalyse omvat twee of meer **categorieën** die toegekend moeten worden aan de verschillende **units** in het **corpus**.

Codeboek
: Een handleiding voor **codeurs** met instructies die aangeven hoe zij te werk moeten gaan. Een codeboek bevat in ieder geval een overzicht van **variabelen** die gecodeerd moeten worden, definities en voorbeelden van die variabelen en de verschillende **categorieën** die ze omvatten, eventuele indicatoren die codeurs kunnen gebruiken om te beslissen hoe ze een bepaalde variabele moeten coderen, en een bronvermelding.

Codeerformulier
: Het formulier (vaak een spreadsheet) waarin de codeurs de data coderen.

Codeur
: Een codeur is iemand die voor alle **variabelen** in het **codeboek** kijkt welke **categorie** van toepassing is op de verschillende **units** in het **corpus** dat bestudeerd wordt.

Coderen
: Het toekennen van **categorieën** aan de **units** in een corpus.

Cohen's Kappa
: Een statistiek die iets zegt over de overeenstemming tussen twee **codeurs**, waarbij gebruik gemaakt wordt van **kanscorrectie**. Cohen's Kappa is alleen geschikt voor nominale data, die gecodeerd wordt door twee codeurs. Voor andersoortige data en voor een ander aantal codeurs dienen andere maten voor intercodeursbetrouwbaarheid gebruikt te worden.

Corpus
: Een corpus is een verzameling van teksten of andere boodschappen, die samengesteld is om een representatief beeld te geven van teksten of andere boodschappen in een bepaald domein.

Dagboekstudie
: Een dagboekstudie is een studieopzet waarbij alle deelnemers een dagboek/logboek bijhouden met hun eigen observaties. Deze dagboeken vormen samen het corpus dat bestudeerd wordt.

Dual use
: Wanneer onderzoeksresultaten voor andere (vaak negatieve) doeleinden gebruikt (kunnen) worden dan de oorspronkelijke onderzoeksdoeleinden, spreken we van *dual use*.

Dubbel coderen
: Bij het uitvoeren van een **inhoudsanalyse** wordt het **corpus** vaak (deels) **gecodeerd** door meerdere **codeurs**, om zo te kunnen bekijken hoe betrouwbaar het codeerinstrument is (via een **intercodeursbetrouwbaarheidsscore**).

Experiment
: Een experiment is een studie waarbij participanten worden toegewezen verschillende experimentele condities, waarbij ze worden blootgesteld aan een bepaalde stimulus, en waarna vervolgens wordt gemeten of de condities verschillen teweeg hebben gebracht tussen de verschillende groepen participanten.

Ethische commissie
: Een ethische commissie is een groep van mensen die wetenschappelijke voorstellen beoordelen, waarbij gekeken wordt of de voorstellen voldoen aan bestaande ethische richtlijnen..

Fonetiek
: Vakgebied binnen de taalwetenschap dat zich richt op de perceptie en productie van spraakgeluiden.

Formule
: Een functie in spreadsheetprogramma's waarmee berekeningen kunnen worden uitgevoerd op basis van de inhoud van andere cellen in de spreadsheet.

Frequentietabel
: Een frequentietabel geeft een overzicht van de aantallen waarin de verschillende **categorieën** in een **inhoudsanalyse** voorkomen binnen de verschillende groepen die je bestudeert.

IFID
: Illocutionary Force Indicating Devices; (combinaties van) woorden die vaak samengaan met berichten van een bepaalde aard (bijvoorbeeld verontschuldigingen, doorverwijzingen, enzovoorts).

Inhoudsanalyse
: Inhoudsanalyse is een onderzoeksmethode waarbij boodschappen systematisch gecodeerd worden om conclusies te kunnen trekken over thema's of patronen in de data. *Kwantitatieve* inhoudsanalyse is vaak deductief: op basis van de literatuur wordt er een codeboek op gesteld, waarna codeurs de data categoriseren, en we conclusies trekken op basis van de frequenties van verschillende categorieën. *Kwalitatieve* inhoudsanalyse is een meer inductief proces, waarbij een codeerschema wordt ontwikkeld op basis van de data. Vervolgens categoriseren de codeurs de data, maar worden er vooral conclusies getrokken op basis van de verschillende thema's die gevonden zijn in de data.

Intercodeursbetrouwbaarheid
: De mate waarin verschillende **codeurs** de data op dezelfde manier **coderen**.

Intercodeursbetrouwbaarheidsscore
: Een statistische score (zoals **Krippendorff's alpha** of **Cohen's kappa**) die aangeeft in hoeverre er overeenstemming is tussen twee of meer **codeurs**. Vaak wordt er in deze score ook **kanscorrectie** toegepast.

Hawthorne effect
: De term *Hawthorne effect* verwijst naar het idee dat mensen zich anders gedragen wanneer ze geobserveerd worden. Om dit effect te vermijden, kun je bijvoorbeeld gebruik maken van bestaande data om te zien hoe mensen zich 'in het echt' gedragen.

Hoofdcodeerfase
: De fase waarin het gehele corpus gecodeerd wordt aan de hand van het **codeboek**. Deze fase volgt na de **pilotstudie**, waarin aangetoond dient te worden dat de **inter-**

codeursbetrouwbaarheid goed genoeg is om aan het eind van de studie generaliseerbare conclusies te kunnen trekken.

Kanscorrectie
: Kanscorrectie zorgt ervoor dat er bij de berekening van intercodeursbetrouwbaarheidsscores rekening gehouden wordt met de kans dat twee codeurs het toevallig (puur op basis van frequentie) eens zijn met elkaar.

Krippendorff's Alpha
: Een statistiek die iets zegt over de overeenstemming tussen twee **codeurs**, waarbij gebruik gemaakt wordt van **kanscorrectie**. Krippendorff's alpha is niet alleen geschikt voor nominale data, maar werkt bijvoorbeeld ook voor intervallen. Een ander voordeel van Krippendorff's Alpha ten opzichte van Cohen's Kappa is dat Krippendorff's alpha ook werkt voor meer dan twee codeurs, en ook kan werken met ontbrekende datapunten.

Latent
: Een eigenschap is latent als deze niet direct herkenbaar is aan de fysieke eigenschappen van het item dat je codeert. (Tegenovergestelde van Manifest.)

LIWC
: *Linguistic Inquiry and Word Count* is een bekende methode om op basis van woordfrequenties iets te zeggen over de mentale toestand of persoonskenmerken van de auteur van een tekst.

Machine learning
: Een vakgebied dat zich richt op de ontwikkeling van algoritmes die computers in staat stellen om bepaalde taken steeds beter te kunnen uitvoeren naarmate ze blootgesteld zijn aan steeds meer data. (Dit blootstellen aan data om de prestaties van de computer te verbeteren wordt ook wel **trainen** genoemd.)

Manifest
: Een eigenschap is manifest als deze herkenbaar is aan de fysieke eigenschappen van het item dat je codeert. (Tegenovergestelde van Latent.)

Morfologie
: Morfologie is een vakgebied binnen de taalwetenschap dat bestudeert hoe woorden zijn opgebouwd uit verschillende betekenisvolle componenten (die *morfemen* worden genoemd).

Non-verbale communicatie
: Vormen van communicatie die naast geschreven en gesproken taal bestaan, zoals lichaamstaal, houding, en gezichtsuitdrukkingen.

PRISMA
: Een onderzoeksmethode voor literatuurstudies, die ook gebruikt kan worden om de dataverzameling bij een inhoudsanalyse helder in kaart te brengen.

Pilotstudie
: Uitprobeerfase waarin gekeken wordt naar de praktische haalbaarheid van de inhoudsanalyse, en waarin aangetoond dient te worden dat de intercodeursbetrouwbaarheid goed genoeg is om aan het eind van de studie generaliseerbare conclusies te kunnen trekken. Als dat zo is, kan er begonnen worden met de hoofdcodeerfase.

Positionaliteit
: Betreft het expliciet maken van je identiteit en je houding en aannames ten opzichte van het onderwerp dat je onderzoekt. Dit past binnen een **reflexieve** onderzoekshouding.

Preregistratie
: Voordat je een studie uitvoert, kun je het volledige studieprotocol, de hypothesen, en het analyseplan vastleggen en indienen bij een centraal archief (zoals OSF of AsPredicted.org) waar deze gegevens worden opgeslagen. Dit wordt preregistratie genoemd. Bij het publiceren van je resultaten kun je vervolgens laten zien in hoeverre de uiteindelijke studie overeen komt met of afwijkt van het oorspronkelijke plan. Daarmee kun je als onderzoeker transparantie bieden over je werkwijze.

Pseudonimiseren
: Pseudonimiseren betekent dat je de identificerende gegevens van iemand *versleutelt* (bijvoorbeeld iedere naam vervangen door een uniek nummer), waardoor de identiteit van de betrokken personen verborgen is, maar je verschillende soorten data van dezelfde persoon (bijvoorbeeld verschillende posts op sociale media) wel aan elkaar kunt koppelen.

Reflexiviteit
: Betreft het kritisch bevragen van jouw **positionaliteit** en de effecten die jouw positionaliteit kan hebben op je onderzoek.

Sampling
: Het nemen van een steekproef. Hierbij zijn verschillende samplingstrategieën mogelijk, die elk verschillende voor- en nadelen hebben in de representativiteit van de steekproef en de praktische uitvoerbaarheid van de methode.

Semantiek
: Vakgebied binnen de taalwetenschap dat zich richt op het bestuderen van de letterlijke betekenis van talige uitingen. Binnen de semantiek zijn er twee richtingen: lexicale semantiek (waarbij gekeken wordt naar de betekenis van *individuele* woorden), en 'algemene' semantiek (waarbinnen gekeken wordt naar de betekenis van *combinaties van woorden*).

Semi-automatisch coderen
: Codeerstrategie waarbij mensen en computers samenwerken om efficiënter tek-

sten te **coderen**. Hierbij zijn er twee volgordes mogelijk: 1. Een menselijke **codeur** controleert de resultaten van een computersysteem dat automatisch boodschappen heeft gecodeerd, of 2. een computersysteem controleert een menselijke codeur.

Sentiment-analyse
: Parapluterm voor automatische analyses die een indicatie geven van het positieve danwel negatieve sentiment van een tekst.

Sociolinguïstiek
: Vakgebied binnen de taalwetenschap dat bestudeert hoe sociale verhoudingen ons taalgebruik beïnvloeden.

Syntax
: Vakgebied binnen de taalwetenschap dat zich richt op het begrijpen van de zinsbouw van natuurlijke talen, zoals het Nederlands, het Swahili of het Zweeds.

Spreadsheet
: Een spreadsheet (Nederlands: *rekenblad*) is een grote elektronische tabel waarbij je gegevens kunt invoeren. In een spreadsheetprogramma zoals Excel, Numbers, Google Sheets, of LibreOffice Calc zitten verschillende functies om allerlei berekeningen uit te voeren. Daarbij wordt vaak gebruik gemaakt van voorgedefinieerde formules, die je ook weer verder kunt combineren om nog meer berekeningen uit te voeren.

Taalhandeling
: De term *taalhandeling* verwijst naar het idee dat je met een talige uiting niet alleen informatie overbrengt, maar ook veranderingen teweeg kan brengen in de wereld. Je kunt met een uitspraak iets *doen*, zoals beloven, verontschuldigen, uitnodigen, vragen, enzovoorts.

Taalmodel
: Een taalmodel is een computationeel model dat op basis van grote verzamelingen tekst in staat is om van een tekstfragment te voorspellen hoe dat fragment verder zou kunnen gaan. Grote taalmodellen (Engels: *large language models*) zijn doorontwikkeld om grotere tekstbestanden te kunnen analyseren en vragen van gebruikers te beantwoorden. (Daarbij genereert het model *plausibele* en *overtuigend klinkende* antwoorden die niet noodzakelijkerwijs correct zijn.)

Trainen
: Binnen inhoudsanalyse spreken we over twee verschillende soorten training:

1) Het trainen van codeurs: hiermee wordt bedoeld dat **codeurs** opgeleid worden om de data allemaal op dezelfde manier te **coderen**, om zo de betrouwbaarheid van de codeertaak te waarborgen.

2) Het trainen van programma's om automatisch te coderen: hiermee wordt bedoeld dat computers via **machine learning** in staat worden gesteld om menselijke codeurs te imiteren.

Units

Er zijn drie verschillende soorten eenheden die we onderscheiden voor inhoudsanalyse:

1) Sampling unit: eenheden die je verzamelt. Bijvoorbeeld: gesprekken op Twitter, televisie-afleveringen.
2) Data collection unit: eenheden die je codeert. (Soms ook *Coding unit* genoemd.) Bijvoorbeeld: iedere reactie binnen het gesprek wordt apart gecodeerd. Of misschien codeer je juist het gesprek als geheel.
3) Unit of analysis: eenheden die je analyseert bij dataverwerking. Op welk niveau aggregeer en analyseer je de gecodeerde data? Afhankelijk van je onderzoeksvraag.

Variabele

De term 'variabele' verwijst naar een bepaalde eigenschap die je wil meten. Een variabele omvat bij een inhoudsanalyse vaak meerdere **categorieën**. Het **coderen** van een variabele gebeurt in een **codeerformulier**, waarin de juiste waardes worden ingevuld door de **codeur**.

Versiebeheer

Versiebeheer gaat over de manier waarop je met onderzoeksdata, codeboeken, programmeercode, en ander onderzoeksmateriaal omgaat. Het idee is dat je ervoor moet zorgen dat alle verschillende versies netjes opgeslagen worden, en dat iedereen altijd met de juiste versie van alle relevante bestanden aan het werken is.

Vulgreep

Functionaliteit in spreadsheetprogramma's om de inhoud van een cel (of reeks cellen) te kopiëren naar een andere cel (of reeks cellen). Wanneer het spreadsheetprogramma detecteert dat de originele cel een formule bevat, wordt de formule toegepast op de cellen waar je de vulgreep overheen sleept. Wanneer het spreadsheetprogramma detecteert dat de originele reeks cellen een serie van waardes bevat (bijvoorbeeld 1,2,3,4), zal het programma proberen de reeks te vervolgen (5,6,7,8).

B

Leesvragen

Onderstaande leesvragen zijn van toepassing op ieder artikel dat een inhoudsanalyse bevat. Probeer deze vragen voor jezelf te beantwoorden wanneer je de artikelen voor deze cursus aan het lezen bent.

1) Algemeen: wat is de onderzoeksvraag?

2) Sampling: samenstelling van het corpus.
   a) Wat voor samplingstrategie hebben de auteurs gebruikt?
   b) Hoeveel items hebben de auteurs verzameld?
   c) Stel je zou zelf een vergelijkbaar corpus willen samenstellen; bevat het artikel dan alle relevante details om de datacollectiemethode te repliceren, of denk je dat er iets ontbreekt?

3) Coderen: labels toewijzen aan de data.
   a) Op welke categorieën zijn de data geanalyseerd?
   b) Waar komen die categorieën vandaan? Waar zijn ze op gebaseerd?
   c) Wat zijn de *data collection units?* Met andere woorden: welke discrete units worden er geannoteerd?
   d) Hoe zorgen de auteurs ervoor dat de annotaties betrouwbaar zijn?
   e) Welke categorie was het meest betrouwbaar te coderen? En welke het minste?

4) Statistiek en rapportage
   a) Hoe is de resultatensectie gestructureerd? (Bijvoorbeeld: in welke volgorde worden verschillende soorten statistiek gepresenteerd?)
   b) Welke statistieken worden er gerapporteerd over de geannoteerde data?
   c) Hoe worden de resultaten gevisualiseerd? (Bijvoorbeeld: tabellen, grafieken.)
   d) Als er geen of beperkte visualisaties zijn: hoe zou je de resultaten anders kunnen weergeven?

5) Ethiek
   a) Zijn er specifieke ethische aandachtspunten voor deze studie?

b) Is het duidelijk hoe de auteurs hiermee om zijn gegaan?
c) Zie je hier mogelijkheden tot verbetering?

6) Algemeen oordeel
   a) Wat zijn de sterke punten van deze studie?
   b) Wat zijn de minder sterke punten van deze studie. (Denk aan het design, of aan de (on)volledigheid van de rapportage.
   c) Als er minder sterke punten zijn, hoe zou je die dan kunnen oplossen?

C

Opdracht: een inhoudsanalyse uitvoeren

C.1 INLEIDING

Inmiddels heb je al meerdere onderzoeken gelezen die gebruik maken van inhoudsanalyse, en al voorzichtig geoefend met meerdere onderdelen van het onderzoeksproces. Met deze opdracht gaan we een stap verder: in een groep van maximaal vier personen ga je zelf een inhoudsanalyse opzetten en uitvoeren, naar aanleiding van eerder onderzoek. Daarvoor krijg je uitgebreid de tijd: het hele semester, te starten met het vormen van het team en het kiezen van een onderwerp tot het stap voor stap uitwerken van een onderzoeksvoorstel en uiteindelijk de uitvoering en rapportage van het onderzoek. Deze bijlage geeft een algemene beschrijving van de groepsopdracht. De methode wordt gedurende het boek verder uitgewerkt.

C.2 DOEL VAN DE OPDRACHT

In deze opdracht komen alle verschillende elementen van de cursus aan bod. Door de geleerde theorie toe te passen in een onderzoek, heb je aan het eind van deze cursus niet alleen de *kennis* over de verschillende onderdelen, maar ook de *vaardigheden*.

- Je leert een corpus van bestaande data samen te stellen, of via een experiment ongestructureerde data te verwerven die vervolgens geschikt zijn voor inhoudsanalyse (bijvoorbeeld tekst of gesproken uitingen).
- Je leert een codeboek op te stellen waarin je variabelen selecteert, definieert en operationaliseert die passen bij de onderzoeksvraag.
- Je leert data handmatig en/of automatisch te annoteren aan de hand van het codeboek.
- Je leert deze data te verwerken en analyseren om tot het eindverslag te komen.
- Je leert de inhoudsanalyse in een onderzoeksverslag te rapporteren en te reflecteren op de gekozen methodiek.

De opdracht bereidt je ook voor op het tentamen, dat zal bestaan uit een combinatie van open en gesloten vragen. Bij de open vragen word je bijvoorbeeld gevraagd om te reflecteren op de voor- en nadelen van verschillende methodologische keuzes.

NB. **Lees de opdracht goed door. Wanneer er niet voldaan wordt aan de eisen van de opdracht, behouden wij het recht voor om de opdracht niet ontvankelijk te verklaren, en dus niet te beoordelen.**

C.3 WERKWIJZE

In deze opdracht werken we in twee blokken:

Blok 1: Het onderzoeksvoorstel
In het eerste blok werken jullie eerst toe naar een voorstel om een inhoudsanalyse uit te voeren. In het hoofdstuk 'Template Onderzoeksvoorstel' kun je zien uit welke onderdelen het voorstel moet bestaan, zodat je hier vanaf het begin van de cursus aan kunt werken.

Blok 2: Uitvoeren en rapporteren
In het tweede blok gaan jullie, na goedkeuring van het onderzoeksvoorstel, de inhoudsanalyse uitvoeren en rapporteren. Alleen het eindresultaat van de opdracht (het groepsrapport) krijgt een cijfer. Op Canvas zijn het schema en de deadlines te vinden. Bij het eindrapport worden alle relevante databestanden ingeleverd, zoals jullie codeboek, corpus, de bestanden met de coderingen, en gebruikte files voor data-analyse (in R, Python, Excel, of SPSS).

C.4 ONDERWERP

De eerste stap van de groepsopdracht is het bedenken van **een onderwerp dat aansluit op de vakken die je gevolgd hebt, of de vakken die je momenteel aan het volgen bent**. De reden hiervoor is dat we in deze methodologische cursus geen tijd hebben voor een uitgebreide literatuurstudie, maar het is wel de bedoeling dat je kunt verantwoorden waarom jullie onderwerp relevant is. Daarnaast helpt de kennis uit je eerdere vakken bij het ontwikkelen van een valide codeboek. In het hoofdstuk over Onderwerpen en onderzoeksvragen kijken we naar verschillende manieren om een onderwerp te bepalen en te motiveren.

C.5 STAPPENPLAN

Voor de groepsopdracht werken we volgens het stappenplan dat hieronder wordt beschreven.

1) **Teams samenstellen**
   Voor deze opdracht werk je in teams; elk team bestaat uit 3-4 studenten. Deze teams worden door de docenten samengesteld.

2) **Groepscontract opstellen**
   Maak afspraken over de onderlinge communicatie en samenwerking, en over de consequenties van het niet nakomen van deze afspraken.

3) **Logboek aanmaken**
   Voor deze opdracht maak je meteen een logboek aan, waarin je bijhoudt wie wat doet. Leg daarbij ook de onderlinge afspraken en deadlines vast. Die afspraken en deadlines hoef je niet in te leveren, maar het is wel belangrijk om elkaar aan te kunnen spreken op de verantwoordelijkheid die jullie samen dragen voor dit project.

4) **Onderwerp bepalen**
   Jullie mogen zelf een onderwerp kiezen. Met het onderwerp als uitgangspunt, kunnen jullie nadenken over de aanleiding en motivatie voor jullie studie. Dit vormt de basis voor de inleiding van het onderzoeksvoorstel – en uiteindelijk het onderzoeksverslag (zie ook de appendix 'Template Onderzoeksvoorstel').

5) **Onderzoeksvraag bepalen**
   Als jullie een onderwerp hebben gekozen, kan de onderzoeksvraag worden bepaald. Deze vraag staat centraal in de hele opdracht en bepaalt de volgende stappen in het proces. De onderzoeksvraag bevat een *vergelijkende component*, die sturing geeft aan de data-analyse. Bijvoorbeeld: *tijdstip X versus tijdstip Y, gesprekken met versus zonder beeld, de ene versus de andere aanleiding voor een webcare-gesprek, mannen versus vrouwen, TikTok versus Instagram, etc.* Zorg voor een goede motivatie van de onderzoeksvraag (waarom wil je deze twee groepen vergelijken?), en formuleer verwachtingen bij de onderzoeksvraag.

6) **Sample samenstellen**
   Voor deze opdracht stel je zelf een corpus samen. Dat kan bijvoorbeeld een verzameling socialemediaberichten, YouTube-video's of televisiefragmenten zijn die je zelf aanlegt, maar ook literatuur of een sample genomen uit een bestaande dataset (zoals LexisNexis of OpenSONAR). Belangrijk is dat de dataset groot genoeg moet zijn om

je codeboek te testen, en om gegronde conclusies te kunnen trekken. Wat "groot genoeg" precies inhoudt, ligt aan het soort data dat je verzamelt (zie ook 'Template Onderzoeksvoorstel'). Zorg dat de data netjes opgeslagen worden en op een gemakkelijke manier gecodeerd kunnen worden (bijvoorbeeld: een Excelbestand met één regel per observatie en één kolom per eigenschap).

7) **Codeboek ontwikkelen**
Jullie gaan zelf een codeboek ontwikkelen dat past bij de onderzoeksvraag. Je kunt je hierbij laten inspireren door de literatuur en variabelen uit bestaande codeboeken toespitsen op jullie eigen onderzoek. Het is niet de bedoeling dat je klakkeloos een bestaand codeboek volledig overneemt; het is belangrijk dat je laat zien de vaardigheid te hebben om een valide codeboek te ontwikkelen. Maak ook een planning waarin je laat zien hoe lang jullie ongeveer bezig denken te zijn met het coderen (zie ook 'Template Onderzoeksvoorstel').

8) **Peer feedback**
rste onderdelen van het onderzoeksvoorstel (algemeen, sampling, codeboek) worden tijdens het werkcollege in een peer feedback-setting besproken. Door het onderzoeksvoorstel te bespreken met een ander team, ontwikkel je je eigen kennis en vaardigheden verder op het gebied van inhoudsanalyses, kun je een ander team helpen, en zelf ook profiteren van de adviezen van anderen.

9) **Data-analyseplan en werkplan**
Het laatste onderdeel dat nog aan het onderzoeksvoorstel toegevoegd moet worden voordat het ingeleverd wordt bij de docenten, is het maken van een data-analyseplan en werkplan. Immers, voordat je het onderzoek gaat uitvoeren, is het verstandig om na te denken hoe je de verzamelde data gaat analyseren. Om bovendien niet in tijdnood te komen bij de uitvoering van het onderzoek, is het verstandig om met het onderzoeksteam een plan te maken met werkafspraken per week (zie 'Template Onderzoeksvoorstel').

10) **Inleveren onderzoeksvoorstel**
Deadline is week 7. Tijdens de midterms wordt het onderzoeksvoorstel door de docenten bekeken en van feedback voorzien, zodat jullie (na goedkeuring) in de eerste week van het tweede blok kunnen starten met de inhoudsanalyse.

11) **Start inhoudsanalyse: pilot**
De eerste stap bij de uitvoering is het samenstellen van het sample. Daarna volgt het

testen van het codeboek voordat het daadwerkelijke onderzoek begint. Dit wordt een 'pilot' genoemd. De pilot voer je als volgt uit:

a) Jullie codeboek test je op een subsample dat je in de beschikbare tijd (bijvoorbeeld 2 uur) kunt coderen (afhankelijk van het type units dat je verzamelt voor het sample).
b) Onafhankelijk van de andere teamleden analyseert *elk teamlid* deze (zelfde) units met het codeboek, waarna jullie de intercoder reliability berekenen.
c) Bespreek tijdens een onderlinge meeting de problematische variabelen: welke verbeteringen kun je aanbrengen in het codeboek om bij het hoofdonderzoek een betere overeenstemming te bereiken?
d) Dit alles leidt tot een verbeterd codeboek. Hiermee voeren jullie de daadwerkelijke inhoudsanalyse uit op het gehele sample.

12) **Analyse van het sample**

Nu je codeboek gefinetuned is, kan de analyse van het werkelijke corpus beginnen. Verdeel het hoofdcorpus onder de teamleden zodat elk lid evenveel units codeert. Het is noodzakelijk dat *elk teamlid* een deel van het corpus codeert. Daarnaast wordt een deel van het corpus dubbelgecodeerd door *alle* teamleden om opnieuw de intercoder reliability te checken.[1]

13) **Data-analyse**

Het doel van deze opdracht is tweeledig:

Enerzijds: intercodeursbetrouwbaarheid berekenen om te achterhalen hoe valide je onderzoeksinstrument (codeboek) was. Die bereken je op basis van de dubbelcoderingen.

Anderzijds: data analyseren om het antwoord te kunnen geven op je onderzoeksvraag. Die bereken je op basis van het gehele corpus van je hoofdonderzoek. (Dus alle berichten die door jullie zijn gecodeerd.)

1 Een vuistregel die vaak wordt genoemd is dat minstens 10% van je sample dubbel gecodeerd moet zijn (Neuendorf, 2017; Stortenbeker et al., 2022). Maar het belangrijkste is dat je genoeg items dubbel codeert om met enige zekerheid uitspraken te kunnen doen over de betrouwbaarheid van de codeurs. Bij kleine steekproeven is 10% niet genoeg, en zul je meer items dubbel moeten coderen.

14) **Rapportage**
Het eindproduct van deze opdracht is een rapport waarin je, na een korte inleiding, de methode, resultaten, conclusie en discussie rapporteert. Met het rapport toon je de werkwijze van je onderzoek, beschrijf je de bevindingen, en geef je methodologische leer- en discussiepunten. Het is vooral methodologisch van aard, dus het is niet nodig om een uitgebreid theoretisch kader toe te voegen. De onderzoeksvraag en verwachtingen moeten uiteraard wel beargumenteerd worden. Als leidraad voor de conclusie/discussie: bespreek 3 leermomenten uit de opdracht, dus geef aan welke problemen er optraden en hoe je die in de toekomst zou kunnen voorkomen. Geef ook 3 voorbeelden die goed zijn gegaan en die je als tips wilt meegeven aan andere onderzoekers. Als service voorafgaand aan het eindrapport organiseren we een Q&A waarin je de laatste, prangende vragen aan ons kunt stellen.

15) **Inleveren groepsrapport**
Deadline is week 15. Het rapport, en alle andere relevante bestanden (codeboek, gecodeerd sample, SPSS-bestanden, etc.) lever je in via Canvas.

C.6 PLANNING

Deze opdracht kost veel tijd, maar het is allemaal goed te doen als je iedere week wat schrijft. Bijvoorbeeld:

- Wanneer je het onderwerp bepaald hebt, kun je gelijk opschrijven waarom jullie voor dat onderwerp gekozen hebben, wat de aanleiding is, welke onderzoeksvraag jullie hebben geformuleerd, en wat jullie verwachtingen zijn.
- Bij iedere methodologische keuze schrijf je meteen op (1) dat je die keuze hebt gemaakt, (2) waarvoor je hebt gekozen, en (3) waarom je daarvoor hebt gekozen.

"Write the paper first"

Voor eigenlijk alle onderdelen van een onderzoek geldt dat je al iets kunt schrijven voordat je ook maar begonnen bent. Eisner (2010) schreef daar een erg leesbaar essay over, getiteld "Write the paper first." Zijn idee is dat het verhaal achter het onderzoek leidend moet zijn in de keuzes die je maakt, en dat schrijven helpt bij het concreet maken van je plannen. Door alles meteen op te schrijven, ben je je eerder bewust van gaten of zwakke plekken in het onderzoek. Zelfs de resultatensectie kun je meteen al schrijven nadat je de onderzoeksvraag hebt bepaald. Weliswaar heb je nog geen exacte getallen, maar je kunt al wel bedenken hoe die sectie van je verslag eruit moet zien, inclusief tabellen en visualisaties ter ondersteuning van je verhaal. Met die informatie kun je een betere planning maken voor het onderzoek. Je kunt meer lezen over rapportage in Hoofdstuk 10.

D

Template onderzoeksvoorstel

Het onderzoeksvoorstel is de basis voor je uiteindelijke verslag. Hierin schrijf je wat je wil gaan doen, waarom je dat wil gaan doen, en hoe je dat precies wil gaan doen. Afhankelijk van de fase van je onderzoek kun je eerst een kort antwoord proberen te formuleren op iedere vraag, waarna je de tekst kunt herschrijven tot een lopend geheel. Deze stappen hebben ieder een eigen functie: de vragen helpen om het voorstel compleet te maken, terwijl het herschrijven helpt om dieper na te denken over het onderzoek en de manier waarop alle onderdelen op elkaar aansluiten.

Zoals je misschien al hebt geconcludeerd, is het verstandig om het onderzoeksvoorstel te zien als een eerste versie van je verslag. De inleiding en methodesectie kun je nu al schrijven zoals je dat ook in het eindverslag zou doen, inclusief referenties naar de wetenschappelijke literatuur die je tot nu toe geraadpleegd hebt.

D.1 INLEIDING

Schrijf een inleiding voor jullie studie. Deze inleiding vormt de motivatie voor jullie onderzoek. In de inleiding moeten de volgende punten aan bod komen:

- Wat is de aanleiding voor jullie onderzoek?
- Wat is jullie onderzoeksvraag? (Zorg hierbij voor een vergelijkende component)
- Waarom is die onderzoeksvraag relevant?
- Wat verwachten jullie te gaan vinden? Waarom?
- Wat voor data gaan jullie verzamelen?
- Hoe gaan jullie die data analyseren? (globaal, in één of twee zinnen)
- Wat draagt jullie onderzoek daarmee bij aan het probleem dat de aanleiding vormt van jullie studie?

D.2 METHODE

Sampling

Jullie hebben ongeveer twee weken om het corpus (het sample dat je gaat coderen) samen te stellen.

- Wat gaan jullie precies verzamelen?
- Hoeveel data gaan jullie verzamelen?
- Hoe gaan jullie die data verzamelen?
- Hoeveel tijd kost het (per item of in totaal) om de data te verzamelen?
- Wat zijn de beoogde eigenschappen van de data die jullie willen verzamelen? Bijvoorbeeld, maar niet uitsluitend:
  - Welke groepen gaan jullie vergelijken?
  - Aantal items per groep.
  - Aantal verschillende auteurs.
  - Diversiteit (denk bijvoorbeeld aan: onderwerp en lengte van de tekst, gender en sociaal-economische status van de auteur, ...).
- Hoe organiseren jullie de data zodanig dat de oorsprong en andere metadata bewaard blijven?

Codeboek

Beschrijf jullie codeboek voor de groepsopdracht. Zorg ervoor dat alle onderstaande vragen beantwoord worden in jullie onderzoeksvoorstel.

- Hoe gaan jullie die data coderen?
  - Hoeveel variabelen gaan jullie coderen per item?
  - Uit hoeveel dimensies bestaat iedere variabele? (Hoe fijnmazig zijn die dimensies? Binair ja/nee, of meerdere niveaus?)
  - Hoeveel daarvan doen jullie handmatig?
- Op welk niveau gaan jullie coderen? (zin, woord, bericht, gesprek, ...)
- Coderen jullie in één keer, of zijn er meerdere stappen nodig? (Bijvoorbeeld: tekst opsplitsen in zinnen, en dan zinnen coderen.) Als dit zo is, moet je goed nadenken over de evaluatie van de betrouwbaarheid.
- Hoe ziet jullie codeboek eruit?
- Hoe worden alle variabelen gedefinieerd?
- Waar komen de eigenschappen uit het codeboek vandaan? (Geef bronnen indien van toepassing.)

Codeerproces

Jullie hebben straks ongeveer drie weken om te coderen. Dat proces omvat de volgende stappen:

- 1-2 weken: pilotstudie
  - Neem een kleine steekproef van items om te coderen.
  - Coderen met initieel codeboek.
  - Betrouwbaarheid berekenen per categorie.
  - Herzien codeboek op basis van de pilot.
- 2-3 weken: coderen
  - Coderen met definitief codeboek.
  - Betrouwbaarheid berekenen per categorie.

Voor het coderen hebben jullie ongeveer 6-8 uur per persoon per week (dit vak kost je een dag per week, maar tijdens college kun je natuurlijk niet coderen). Wij willen dat jullie binnen die tijd een realistische hoeveelheid data coderen. Dus niet te weinig (klaar in een uurtje) en niet te veel (je moet ook geen maand fulltime bezig zijn natuurlijk). Daarom moeten jullie een plan maken, met een berekening om te laten zien dat je ongeveer twee weken bezig bent met de hoofdcodeerfase. (Dus ongeveer 2 keer 6-8 uur per week, keer het aantal groepsleden.) Reserveer ook tijd voor het berekenen van de betrouwbaarheid en het herzien van het codeerschema.

- Hoeveel tijd kost het om één item te coderen?
- Hoeveel items gaat iedereen coderen?
- Hoeveel items gaan jullie daarnaast nog dubbel coderen?
- Hoeveel tijd kost het dan om alles te coderen?

Daarnaast moet je ook alvast nadenken over het berekenen van de betrouwbaarheid.

- Hoe gaan jullie de betrouwbaarheid berekenen voor de verschillende categorieën?

D.3 ANALYSE

Dit is de voorloper van de resultatensectie. Als alle data zijn gecodeerd, en de uiteindelijke betrouwbaarheid is berekend, kun je de gecodeerde data analyseren. Het is verstandig om vooraf al te bedenken hoe je dat gaat doen, en wat je wil gaan rapporteren. Dit schrijf je daarom ook op in het onderzoeksvoorstel.

- Welke descriptieve statistieken gaan jullie rapporteren?
- Wat voor figuren en tabellen gaan jullie rapporteren?
- Welke statistische toets gaan jullie gebruiken?
- Wanneer geven jullie resultaten wel/geen ondersteuning voor jullie hypothesen?

D.4 BIJLAGE: CODEBOEK

Zorg voor een eerste versie van je codeboek en voeg deze toe aan het onderzoeksvoorstel. Zoals je weet uit Hoofdstuk 5 bevat een codeboek:

- De namen van de categorieën die je gaat coderen.
- Een definitie van de categorie.
- Een beschrijving van hoe de categorie gecodeerd moet worden. (Liefst met duidelijke indicatoren/IFIDs.)
- Een voorbeeld om codeurs te helpen. (Liefst met uitleg.)
- Een verwijzing naar de wetenschappelijke literatuur waar de operationalisatie op gebaseerd is.

Daarnaast is het goed om nagedacht te hebben over het codeerformulier. Hoe ziet de spreadsheet eruit waarin je gaat coderen?

D.5 BIJLAGE: WERKPLAN

Maak een plan voor wat jullie iedere week van blok 4 gaan doen. Zorg ervoor dat iedereen elke week iets doet, en dat jullie allemaal genoeg tijd besteden aan het project. Daarnaast moet iedereen betrokken zijn bij het coderen, en iedereen moet betrokken zijn bij het schrijven van het verslag. Als startpunt staan hier de weken van blok 4:

- Week 1: Sampling.
- Week 2: Sampling, begin van de pilot.
- Week 3: Pilot.
- Week 4: Hoofdcodeerfase.
- Week 5: Hoofdcodeerfase.
- Week 6: Afronden coderen, betrouwbaarheid.
- Week 7: Data-analyse.
- Week 8: Verslag afmaken.

Advies: Zorg ervoor dat je het schrijven niet tot het eind uitstelt. Maak notities tijdens het onderzoek, zodat alle belangrijke beslissingen (en de redenen daarvoor) al op papier staan. Deze notities vormen samen met dit onderzoeksvoorstel de basis van je verslag.

E

Richtlijnen groepsrapport

In het groepsrapport doe je verslag van de werkwijze van je onderzoek, beschrijf je de bevindingen, en geef je methodologische leer- en discussiepunten. Het is dus vooral methodologisch van aard, het is niet nodig om een theoretisch kader toe te voegen. Het rapport draagt voor 40% bij aan je eindcijfer voor deze cursus.

E.1 ALGEMENE RICHTLIJNEN

Het rapport bestaat uit ongeveer 4000 woorden, exclusief de titelpagina, tabellen, referenties en bijlagen. Het is geschreven in academisch Nederlands, volgens de APA-richtlijnen, Times New Roman 12 pts, regelafstand 1,5. Eén teamlid uploadt het rapport en de additionele bestanden in Canvas. Hanteer in het rapport onderstaande tussenkoppen.

1. Introductie

Hier wordt het onderwerp geïntroduceerd, uitmondend in de onderzoeksvraag. Je verwijst hierbij op bondige wijze naar een aantal relevante bronnen, maar een uitgebreid literatuuronderzoek is niet nodig. Daarnaast beschrijf je in een paar zinnen hoe jullie onderzoek er globaal uitziet, en beschrijf je ook jullie verwachtingen: welke verschillen denken jullie te gaan vinden? Dit onderdeel hoeft maar 1-2 pagina's lang te zijn.

2. Methode

Hierin beschrijf je in ieder geval de vier onderstaande onderdelen, met een duidelijke motivatie voor de gemaakte keuzes en een reflectie op de ethische vragen die naar voren zijn gekomen tijdens het opzetten van het onderzoek.

- Corpus: beschrijf duidelijk op basis van welke criteria het corpus voor jouw onderzoek is samengesteld. Wat was de grootte van het corpus, waar heb je op gelet bij het samenstellen, wat waren inclusiecriteria, in welke tijdsperiode is het verzameld?
- Pretest: beschrijf de uitkomsten van de pretest en tot welke aanpassingen in het codeboek die nog (eventueel) heeft geleid voor het hoofdonderzoek.
- Codeboek: beschrijf de variabelen die in het codeboek zijn gebruikt. Geef per variabele de gehanteerde definities en omschrijf de wijze waarop ze in jullie onderzoek

zijn geoperationaliseerd. Verwijs ook naar de bijlage, waar het gehele codeboek te vinden is. Dit is het codeboek dat jullie hebben gebruikt voor het hoofdonderzoek.
- Procedure: beschrijf hoe de codeurs te werk gingen in het hoofdonderzoek en welke instructies zijn meegegeven.

3. Resultaten

Deze sectie bestaat uit twee onderdelen. (Zie bijvoorbeeld hoe Van Hooijdonk en Liebrecht (2018) hebben gerapporteerd in hun artikel over webcaregesprekken van Nederlandse gemeenten.)
- Betrouwbaarheid van het analyse-instrument: rapporteer de intercodeursbetrouwbaarheid **per variabele** en bespreek op welke punten de codeurs het met elkaar eens/oneens waren, en waarom.
- Resultaten van de analyse: bereken de statistieken zodat je antwoord kunt geven op je onderzoeksvraag. De resultaten worden, waar relevant, ondersteund door visualisaties van de data (dat wil zeggen: tabellen of grafieken).

4. Conclusie en Discussie

Deze sectie bestaat eveneens uit twee onderdelen:
- Het antwoord op je onderzoeksvraag
- Betrouwbaarheid van het analyse-instrument
  - 3 lessen of verbeterpunten uit het huidige onderzoek
  - 3 voorbeelden van dingen die wel goed gingen

5. Appendix

Naast het rapport dienen ook andere documenten ingeleverd te worden. Zie Canvas.

E.2 BEOORDELINGSFORMULIER GROEPSOPDRACHT

Hieronder lees je alle criteria die gebruikt worden binnen het beoordelingsformulier. Zorg ervoor dat je verslag aan de criteria voldoet.

E.2.1 Inleiding

- De introductie begint aantrekkelijk en motiveert om door te lezen.
- De inleiding maakt het onderwerp van het onderzoek duidelijk.
- De relevantie van het onderzoek is helder.
- Het is duidelijk waarop deze studie is gebaseerd.
- Er is/zijn RQ's, onderbouwd met argumentatie.
- Waar relevant worden verwachtingen over de uitkomst uitgesproken.
- Het onderwerp is origineel.

E.2.2 Methode

- De steekproef is duidelijk beschreven (wat voor content, waar verzameld, grootte van het sample, tijdsperiode, etc.)
- Het project is ambitieus
- Codeboek is helder (welke variabelen, hoe gedefinieerd en geoperationaliseerd)
- De methode geeft duidelijk aan waar de categorieën uit het codeboek op gebaseerd zijn, en welke categorieën nieuw zijn ten opzichte van de bestaande literatuur
- Pretestmethode en resultaten zijn gerapporteerd, plus tot welke aanpassingen het heeft geleid in het finale codeboek
- Procedure is nauwkeurig omschreven
- Er is helder beschreven hoe getracht is om tot betrouwbare coderingen te komen
- Relevante ethische vragen omtrent de opzet van het onderzoek worden besproken

E.2.3 Resultaten

Betrouwbaarheid van het analyseinstrument

- Intercodeursbetrouwbaarheid per gecodeerde variabele is gerapporteerd
- In tekst wordt gereflecteerd op de hoogte van de betrouwbaarheidsscores

Beantwoording onderzoeksvraag

- De data-analyse sluit aan bij de RQ's
- De resultaten worden ondersteund door heldere figuren of tabellen
- De tabellen/figuren zijn beschreven en geïnterpreteerd in de tekst
- Statistiek is correct uitgevoerd, gerapporteerd en toegelicht
- De dataset is voldoende benut om relevante resultaten eruit te genereren

E.2.4 Conclusie & Discussie

Beantwoording onderzoeksvraag

- De sectie begint bondig met de RQ's en aanleiding, doel en methode van de studie
- De antwoorden op de RQ's zijn bondig en dekken de bevindingen van de corpusanalyse
- De bevindingen worden verklaard en zijn kort gerelateerd aan relevante literatuur
- De tekst geeft een inhoudelijke bespreking van de generaliseerbaarheid van de studie

Betrouwbaarheid van het analyseinstrument

- Er worden gedegen conclusies getrokken over de betrouwbaarheid van het analyseinstrument
- Kanttekeningen / tekortkomingen van het onderzoek worden helder besproken
- Implicaties van het onderzoek worden helder besproken.
- Er zijn concrete suggesties voor vervolgonderzoek

E.2.5 Structuur & lay-out

- Goede logische structuur en opbouw
- Gedegen redeneringen en argumentatiekwaliteit
- Correct wetenschappelijk Nederlands, geen spelfouten
- Juiste tijdsvormen (verleden/tegenwoordige tijd) in de secties
- APA is correct toegepast, bronnen zijn correct gebruikt
- Lengte: maximaal 4000 woorden (+10%) zonder titelblad, bijlagen en referentielijst
- Appendix is compleet
- Plagiaatcheck wijst uit dat het rapport eigen werk is

F

Principes uit de Nederlandse gedragscode

Hieronder staan de vijf principes uit de Nederlandse gedragscode wetenschappelijke integriteit in het kort (KNAW et al., 2018; eigen formattering)".

Eerlijkheid
"houdt onder andere in dat men:
- geen ongefundeerde claims doet,
- over het onderzoeksproces correct rapporteert,
- data of bronnen niet verzint of vervalst,
- alternatieve visies en tegenargumenten serieus neemt,
- open is over onzekerheidsmarges, en
- resultaten niet gunstiger dan wel ongunstiger voorstelt dan ze zijn."

Zorgvuldigheid
"houdt onder andere in dat men wetenschappelijke methoden gebruikt en optimale precisie betracht bij het ontwerp, de uitvoering, verslaglegging en disseminatie van het onderzoek."

Transparantie
"houdt onder andere in dat het voor anderen helder is:
- op welke data men zich heeft gebaseerd,
- hoe deze zijn verkregen,
- welke resultaten men heeft bereikt en langs welke weg, en
- wat de rol van externe belanghebbenden is geweest.

Als delen van het onderzoek of van de data niet toegankelijk worden gemaakt, dient de onderzoeker goed gemotiveerd aan te geven waarom dat niet mogelijk is.

De wijze van uitvoering en fasering van het onderzoeksproces moet tenminste voor vakgenoten te volgen zijn. Dit betekent in ieder geval dat de argumentatie helder moet zijn en dat de stappen in het onderzoeksproces controleerbaar moeten zijn."

Onafhankelijkheid
"houdt onder andere in dat men zich in de keuze van de methode, bij de beoordeling van de data en in de weging van alternatieve verklaringen, maar ook bij het beoordelen van onderzoek of onderzoeksvoorstellen van anderen, niet laat leiden door buiten-wetenschappelijke overwegingen (bijvoorbeeld overwegingen van commerciële of politieke aard). Aldus geformuleerd omvat onafhankelijkheid ook onpartijdigheid. Onafhankelijkheid is in elk geval vereist bij de opzet en uitvoering van en rapportage over het onderzoek; bij de keuze van het onderzoeksobject en van de onderzoeksvraag is onafhankelijkheid niet altijd nodig"

Verantwoordelijkheid
"houdt onder andere in dat men zich rekenschap geeft van het feit dat men als onderzoeker niet in isolement opereert, en daarom binnen de grenzen van het redelijke rekening houdt met de legitieme belangen van bij het onderzoek betrokken personen en dieren, van eventuele opdrachtgevers en financiers, en van de omgeving. Verantwoordelijkheid houdt ook in dat men onderzoek doet dat wetenschappelijk en/of maatschappelijk relevant is."

G

Programmeren in Excel

G.1 INLEIDING

Dit hoofdstuk geeft een kort overzicht van de functies en mogelijkheden van Microsoft Excel en andere spreadsheetprogramma's.[1] Hoewel het in eerste instantie wel wat tijd kost om alles onder de knie te krijgen, betaalt deze kennis zich dubbel en dwars uit: niet alleen kun je veel sneller data coderen en voorbereiden voor de statistische analyse, maar ook in het bedrijfsleven wordt er veel gebruik gemaakt van Excel.[2] Voor meer informatie staan er ontzettend veel handleidingen en tips op internet, bijvoorbeeld op de website: https://exceljet.net/[3]

G.1.1 CELLEN IN EXCEL

Eigenlijk is een Excelbestand gewoon een oneindig grote tabel waarin je gegevens kunt invullen. Alle cellen in Excel zijn te identificeren aan de hand van de kolom en de rij waarin de cel zich bevindt. Alle kolommen in Excel worden aangeduid met een letter (A-Z, gevolgd door AA-ZZ, enzovoorts), en alle cellen worden aangeduid met een cijfer. Figuur G.1 geeft hiervan een voorbeeld. De cel C2 bevindt zich in de derde kolom, op de tweede rij.

	A	B	C	D	E	F
1						
2			Hallo			
3						
4						

FIGUUR G.1 *Voorbeeld van een deel van een spreadsheet zoals gebruikt wordt in Excel.*

1 Voorbeelden zijn *Google Sheets*, *LibreOffice Calc*, of Apple's *Numbers*. In dit hoofdstuk wordt de naam "Excel" gebruikt omdat dit programma in de praktijk het meest gebruikt wordt.

2 Wist je trouwens dat er ook wereldkampioenschappen Excel worden gehouden?

3 Voor dit hoofdstuk wil ik Jacqueline Dake bedanken voor het beschikbaar stellen van haar (veel uitgebreidere) cursusmateriaal over Excel.

G.1.2 FORMULES IN EXCEL

Data-analyse in Excel gebeurt aan de hand van *formules*. Een formule begint altijd met het is-teken (=), gevolgd door een combinatie van verwijzingen naar cellen, functies (voorgedefinieerde handelingen), en operatoren (tekens voor optellen, aftrekken, delen, en vermenigvuldigen: +, -, /,*). Een voorbeeld van een formule is `=ROUND(A3,2)`. Hiermee zeg je eigenlijk: de waarde van deze cel is gelijk aan de waarde uit cel A3, afgerond (met de `ROUND`-functie) tot twee decimalen achter de komma.

Voorbeeld: cijfers afronden

Zoals gezegd kun je verschillende functies in Excel ook met elkaar combineren. Laten we daarvoor kijken naar een specifiek voorbeeld: het afronden van tentamencijfers. In Tilburg geldt een speciale regel voor cijfers die verplicht voldoende moeten zijn:

- Cijfers tussen de 5 en 6 worden afgerond op hele punten (een 5 of een 6).
- Andere cijfers worden afgerond op halve punten (het dichtstbijzijnde veelvoud van 0.5)).

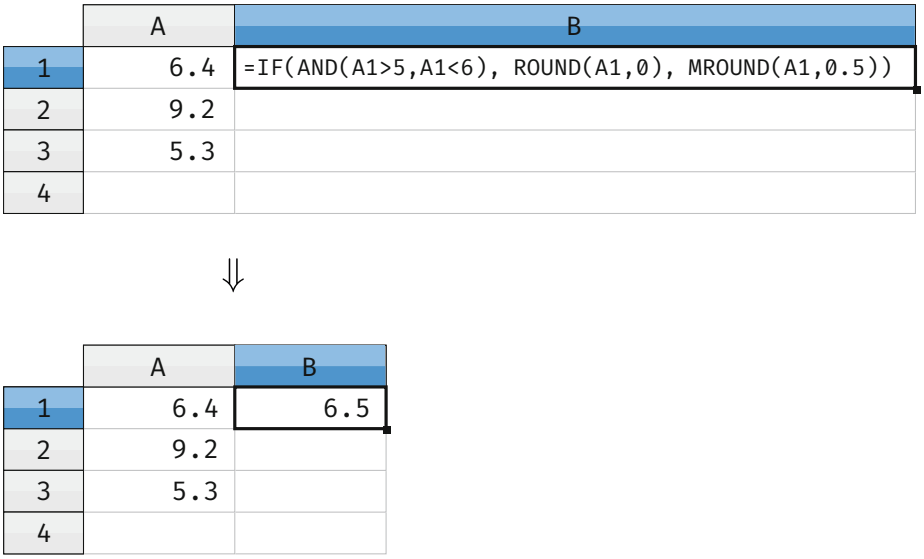
Om deze regel om te zetten in een formule hebben we de volgende ingrediënten nodig:

- `ROUND(X,Y)`: rondt getallen af op N decimalen achter de komma. Bij afronden op hele getallen kun je voor `Y` het getal 0 invullen.
- `MROUND(X,Y)`: rondt getallen af naar veelvouden van `Y`. In ons voorbeeld is dat 0.5.
- `IF(X,Y,Z)`: controleert of aan een bepaalde voorwaarde voldaan wordt. Als `X`, dan wordt Y uitgevoerd, en anders Z. In ons voorbeeld willen we controleren of het cijfer tussen de 5 en de 6 ligt.
- `AND(X,Y)`: controleert of X en Y allebei waar zijn of niet. In ons voorbeeld kijken we of het getal zowel groter is dan 5 als kleiner dan 6. Dan weten we dat het getal tussen de 5 en de 6 in zit.

Als het cijfer dat je wil afronden in de cel A1 staat, dan kunnen we dat cijfer afronden met de volgende formule: `=IF(AND(A1>5,A1<6), ROUND(A1,0), MROUND(A1,0.5))`. Dat ziet er in Excel uit als in Figuur G.2.

Hoe maak je zelf een formule?

Het bovenstaande voorbeeld is misschien wat intimiderend. Hoe weet je nu eigenlijk welke functies je allemaal moet gebruiken? Het antwoord is dat je maar een handvol functies hoeft te kennen. De meeste problemen die je tegenkomt zijn al eerder door andere mensen opgelost, en dan kun je hun oplossingen gewoon aanpassen naar jouw situatie (bijvoorbeeld door een verwijzing te veranderen van D3 naar A1). Met de basale, logische functies `AND, OR`, en `IF` kun je de resultaten vervolgens automatisch combi-



FIGUUR G.2

neren. Binnen de meeste spreadsheetprogramma's zit ook een formule-editor (met zoekfunctie!) die je ondersteunt bij het schrijven van formules.

G.1.3 DE VULGREEP

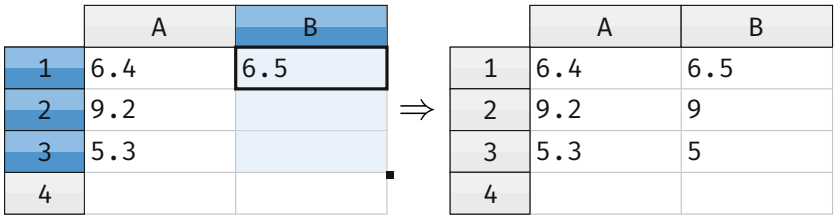
Je kunt een formule toepassen op een hele kolom door gebruik te maken van de vulgreep. Dat is het vierkante blokje dat rechtsonder een cel staat wanneer je hem selecteert. Als je met je muis over de vulgreep gaat, verandert de muis van een pijltje naar een plusje. Je kunt de vulgreep dan dubbelklikken, of naar beneden slepen om de formule op de rest van de kolom toe te passen. De verwijzingen in de formule (zoals A1) worden dan automatisch aangepast aan de relevante rij. Figuur G.3 geeft aan hoe dit eruit ziet als we de formule uit Figuur G.2 (om cijfers af te ronden) toepassen met de vulgreep.

Kolommen en reeksen

Een andere praktische functie in Excel is om de vulgreep te gebruiken om kolommen te vullen of om reeksen automatisch uit te breiden.

Kolommen vullen

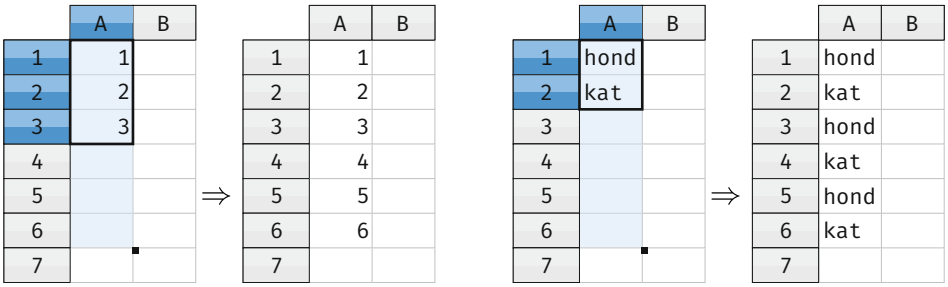
Je kunt een kolom vullen met dezelfde waarde door in de bovenste cel een bepaalde waarde in te vullen (bijvoorbeeld de tekst *bla*) en vervolgens de vulgreep op dezelfde manier te gebruiken als hierboven: dubbelklik het kleine vierkantje, of sleep het vierkantje net zover als dat je de kolom wil vullen met diezelfde waarde (*bla, bla, bla, bla, bla*).



FIGUUR G.3 *Spreadsheet waarbij de vulgreep wordt toegepast door het vierkantje vanaf cel B1 naar beneden te slepen (links), en het resultaat in Excel (rechts). De (nu onzichtbare) formule in A1 wordt automatisch toegepast op de cellen A2 en A3, en het resultaat wordt geplaatst in B2 en B3.*

Reeksen aanvullen

Een reeks kan zowel bestaan uit getallen of tekst. Excel probeert automatisch een patroon te vinden in de opeenvolgende waarden die je hebt ingevuld, en kan die reeks dan automatisch verder aanvullen. Figuur G.4 laat twee voorbeelden zien. In het linker voorbeeld zie je dat de reeks met getallen (1,2,3) automatisch uitgebreid wordt met de volgende getallen (…3,4,5). In het rechter voorbeeld zie je dat de twee tekstuele waarden (*hond* en *kat*) aangevuld worden door de twee waarden steeds af te wisselen.



FIGUUR G.4 *Spreadsheet waarbij de vulgreep wordt toegepast door eerst cellen A1–A3 te selecteren, en vervolgens het vierkantje vanaf cel A3 naar beneden te slepen, en het resultaat in Excel. De reeks waarden in A1–A3 wordt automatisch vervolgd in de cellen eronder.*

Wanneer gebruik je de vulgreep?

Je kunt de vulgreep gebruiken om heel snel een codeerformulier te maken. Stel je voor dat je wil vergelijken hoe bedrijven zichzelf presenteren op verschillende sociale media, en dat je berichten verzamelt voor twee bedrijven (*A* en *B*) op twee platforms (*Tiktok* en *Instagram*). Dat betekent dat je een spreadsheet moet vullen met op iedere regel een

nieuw bericht. Voor de data-analyse (en überhaupt om alles bij te kunnen houden) is het nuttig om:

- Een kolom te hebben met een uniek nummer voor iedere boodschap.
- Kolommen te hebben met de bedrijfsnaam en het platform.
- Als de data hierarchisch is, bijvoorbeeld verschillende gesprekken met meerdere berichten per gesprek: losse kolommen aan te maken met gespreksnummer en berichtnummer.

Als je zelf kunt programmeren en de data automatisch verzamelt met R of Python, kun je die kolommen het beste automatisch aanmaken binnen R of Python. Werken in Excel blijft relatief foutgevoelig, omdat het uiteindelijk toch handwerk blijft. (Je moet bijvoorbeeld wel zorgen dat je precies de juiste cellen vult.)

G.1.4 VERSCHILLENDE TALEN EN DIALECTEN

Verschillende spreadsheetprogramma's hanteren soms verschillende functies en symbolen om dezelfde operaties uit te voeren. Afhankelijk van de variant die je gebruikt, kun je de volgende verschillen verwachten:

- Het decimaalteken is in het Engels een punt, en in het Nederlands een komma. Dus Amerikanen schrijven 3.75, maar Nederlanders schrijven 3,75.
- Argumenten van functies worden soms gescheiden door een puntkomma in plaats van door een komma. In de voorbeelden hierboven is de standaard Engelse variant gebruikt: `=MROUND(A1,0.5)`. Maar soms werkt dat niet, en moet de functie er zo uitzien: `=MROUND(A1;0.5)`.
- Namen van functies worden niet overal geschreven in het Engels. Bij de Nederlandse Excel zijn er ook Nederlandse commando's, waardoor je bijvoorbeeld niet `=IF(…)` schrijft maar =ALS(…).
- Verschillende programma's hebben ook verschillende functionaliteiten. Zo kun je met Google Sheets bijvoorbeeld automatisch de inhoud van een cel vanuit het Duits naar het Nederlands vertalen met het volgende commando: `=GOOGLETRANSLATE(A1, "DE","NL")`.

Om je te helpen heeft Excel een ingebouwde vertaalfunctie die Engelse formules naar het Nederlands kan omzetten. Daarnaast zijn er gespecialiseerde online vertaalmachines om formules van het internet te vertalen naar het Nederlands.[4]

4 Probeer wel eerst de formule te begrijpen voordat je hem uitvoert; onbekende formules moet je nooit zomaar uitvoeren.

G.2 INHOUDSANALYSE AUTOMATISEREN MET EXCEL

Nu hebben we genoeg informatie om voorbeelden te geven waarmee je verzamelingen tekst automatisch kunt coderen. Bij de formules hieronder wordt aangenomen dat de kolom A alle berichten bevat die geanalyseerd moeten worden, en dat het eerste bericht zich bevindt in cel A1. Ziet jouw data er anders uit? Dan moet je de formule even aanpassen zodat hij naar de goede cel verwijst.

G.2.1 LENGTE VAN EEN BERICHT METEN

Je kunt de lengte van een bericht eenvoudig meten door middel van de `LEN`-functie. Deze functie geeft je het aantal karakters van een tekst (inclusief spaties). Als je wil weten hoeveel *woorden* een boodschap bevat, dan wordt de formule wat moeilijker. In de meeste versies van Excel (en soortgelijke programma's) kun je het aantal woorden in een tekst meten als volgt:

```
=IF(LEN(TRIM(A1))=0, 0, LEN(TRIM(A1))-LEN(SUBSTITUTE(A1," ","")) + 1)
```

Deze formule maakt gebruik van het idee dat alle woorden in een tekst gescheiden worden door spaties. Door alle spaties te tellen, krijg je een aardig idee van het aantal woorden in een tekst: dat is gelijk aan het aantal spaties +1. (Een tekst met één woord heeft geen spaties nodig, en een tekst met twee woorden heeft maar één spatie nodig, enzovoorts.) De formule vergelijkt dus de lengte van de boodschap mét en zonder spaties, en telt 1 op bij het verschil tussen die twee lengtes.

Is de bovenstaande formule omslachtig? Jazeker! Excel is eigenlijk niet ontworpen voor dit soort bewerkingen, maar het werkt wel. Modernere versies van Excel bevatten inmiddels extra functies waarmee we eenvoudiger het aantal woorden in een tekst kunnen berekenen:[5]

```
=COUNTA(TEXTSPLIT(TRIM(A1)," "))
```

Deze formule zegt eigenlijk: verander de inhoud van cel A1 in een lijst met woorden, door spaties aan het begin en eind van de tekst te verwijderen, en daarna de tekst op te splitsen bij iedere spatie. Daarna tellen we het aantal items in de lijst. (Dat kunnen woorden of getallen zijn, maar in dit voorbeeld behandelen we getallen ook als woorden.)

5 `TEXTSPLIT` is in 2022 geïntroduceerd.

G.2.2 CONTROLEREN OF EEN BOODSCHAP EEN WOORD BEVAT

Veel inhoudsanalyses zijn te automatiseren door te kijken of een boodschap bepaalde woorden bevat. Eerder in dit boek hebben we al gekeken naar het artikel van Van Hooijdonk & Liebrecht (2021), waarin verontschuldigingen werden gecodeerd. Dit proces zou mogelijk versneld kunnen worden door alle voorkomens van het woord "sorry" te zoeken.

Een eerste oplossing: zoeken naar woordvormen

We kunnen het zoeken naar woorden eenvoudig automatiseren door middel van de volgende formule:[6]

```
=ISNUMBER(SEARCH("tekst", A1))
```

Deze formule kun je lezen door te kijken naar de haakjes, en te beginnen bij de functie die het verst naar binnen staat: SEARCH("tekst", A1). Deze functie zoekt naar het woord "tekst" in cel A3, en geeft als resultaat een getal dat verwijst naar het eerste karakter in de cel A1 waar het woord "tekst" begint. Als het woord "tekst" niet in de cel staat, dan heeft SEARCH geen numeriek resultaat. Nu begrijpen we ook waarom de functie ISNUMBER gebruikt wordt: deze functie geeft als resultaat TRUE als het argument (hetgeen er tussen de haakjes staat) een getal is, en anders geeft de functie het resultaat FALSE. Met andere woorden: als het woord "tekst" voorkomt in A1, dan staat er TRUE, en anders staat er FALSE.

Strings versus woorden

Een probleem met bovenstaande formule is dat de formule ook TRUE oplevert als een zoekterm overeenkomt met een deel van het woord. Bijvoorbeeld: als je zoekt op "aap" dan krijg je ook een positief resultaat als de tekst het woord "schaap" bevat, want de laatste drie letters van dat woord komen overeen met je zoekterm. Excel zoekt dus naar *strings* (opeenvolgingen van karakters) en niet naar woorden.

Een tweede oplossing: zoeken naar exacte termen

Sinds de introductie van de TEXTSPLIT-functie in Excel, kunnen we ook proberen om alle boodschappen zelf te *tokeniseren*, om vervolgens te kijken of onze zoekterm(en) overeen komen met de losse woorden in de boodschappen. Laten we dat stap voor stap

6 De Nederlandse vertaling van onderstaande formule is: =ISGETAL(VIND.SPEC("tekst", A1)).

doen. Met de volgende formule kun je een tekst omzetten in een lijst met woorden, door de tekst op te splitsen bij iedere spatie, punt, of komma:[7]

```
=TEXTSPLIT(A1,{".","," " "},,1)
```

En op deze manier kunnen we kijken of het woord "sorry" voorkomt in die lijst met woorden:

```
=COUNT(XMATCH("sorry",TEXTSPLIT(A1,{".","," " "},,1)))>0
```

In deze formule wordt eerst de tekst in A1 omgezet in een lijst met woorden, vervolgens wordt er gekeken of het woord "sorry" in die lijst staat, dan wordt het aantal voorkomens van dat woord geteld, waarna Excel kijkt of dat aantal groter is dan 0. (Bonus: als je >0 weghaalt aan het eind van de formule krijg je het aantal keer dat "sorry" voorkomt in de tekst.)

Deze tweede oplossing heeft als voordeel dat je de exacte termen kunt vinden waarnaar je op zoek bent, maar tegelijkertijd kan dat ook een nadeel zijn: meervouden of vervoegingen van werkwoorden worden niet herkend, terwijl dat met de eerste oplossing wel gebeurt. De eerste oplossing is misschien iets minder accuraat (denk aan het aap-schaapprobleem), maar je vindt wel meer potentieel relevante boodschappen. Denk dus goed na of je op zoek bent naar woordvormen of exacte termen.

G.2.3 CONTROLEREN OF EEN BOODSCHAP WOORDEN UIT EEN LIJST BEVAT

Op een vergelijkbare manier als hierboven kunnen we controleren of een boodschap woorden uit een lijst bevat. Een lijst met woorden kun je op drie manieren maken in Excel:

1) Met een speciale lijst-notatie: {"aap";"noot";"mies"}. Daarbij gebruik je dus accolades (ook wel krulhaakjes genoemd) aan het begin en eind van de lijst. Woorden staan tussen dubbele aanhalingstekens en zijn gescheiden door puntkomma's.
2) Met een absolute verwijzing: woorden staan in verschillende cellen onder elkaar, bijvoorbeeld A1: *aap*, A2: *noot*, A3: *mies*. Je kunt dan naar de lijst verwijzen op de volgende manier: $A$1:$A$3. (Als je die dollartekens weglaat uit een formule, en je past de vulgreep toe, dan begrijpt Excel niet dat die verwijzing *niet* aangepast moet worden aan de volgende regel.)

7 Let op: dit is een erg basale manier om tekst te tokeniseren. Hierbij wordt bijvoorbeeld geen rekening gehouden met andere leestekens of emoji, en ook afkortingen zijn problematisch.

3) Met een zogenaamde *named range*: een ingebouwde manier om een groep geselecteerde cellen eein eigen naam te geven. Zoek in de handleiding van jouw favoriete spreadsheetprogramma op hoe dat moet.

Wanneer je een lijst hebt gemaakt in Excel, kun je die lijst gebruiken in dezelfde functie als voorheen. Bijvoorbeeld:

```
=COUNT(XMATCH({“aap”;“noot”;“mies”},TEXTSPLIT(A1,{“.”,”,”,” “},,1)))>0
```

Maar in plaats van de lijstnotatie kun je dus ook de absolute verwijzing gebruiken. Als je lijst met woorden in B1–B3 staat, ziet de formule er zo uit:

```
=COUNT(XMATCH($B$1:$B$3,TEXTSPLIT(A1,{“.”,”,”,” “},,1)))>0
```

Het probleem dat je vervolgens bij deze benadering tegenkomt is dat de woorden toch ergens in het Excelbestand moeten staan, liefst niet tussen de data die je wil coderen. Je kunt daarvoor een andere *Sheet* gebruiken: open een nieuw tabblad, en zet daar de woordenlijst in één van de kolommen (bijvoorbeeld in kolom D). Je kunt dan verwijzen naar de woorden met de volgende notatie: `Sheet2!$D$1:$D$3`. Het laatste voorbeeld maakt gebruik van een *named range* die we in dit geval “woordenlijst” hebben genoemd, zodat de formule meteen duidelijk is:

```
=COUNT(XMATCH(woordenlijst,TEXTSPLIT(A1,{“.”,”,”,” “},,1)))>0
```

Je hoeft de woordenlijst natuurlijk niet “woordenlijst” te noemen. Probeer een descriptievere naam te vinden, bijvoorbeeld “verontschuldigingen” of “scheldwoorden.”

G.2.4 VARIABELEN HERCODEREN

Stel je voor dat je uitspraken hebt verzameld van politici van vier verschillende politieke partijen: GroenLinks, PvdA, SGP, en VVD. Iedere uitspraak staat op een nieuwe regel in je spreadsheet, en er zijn verschillende kolommen met metadata, waaronder een kolom met de politieke partij. Nu wil je de partijen indelen in twee groepen: “Links” en “Rechts,”

waarbij de eerste twee partijen als Links worden gezien en de laatste twee partijen als Rechts.[8] Je zou de variabelen kunnen hercoderen met de volgende formule:

```
=IF(ISNUMBER(MATCH(A1,{“PvdA”;“GroenLinks”},0)), “Links”, “Rechts”)
```

Deze formule kijkt of de waarde in cel A1 in de lijst met linkse politieke partijen voorkomt. Als dat zo is, dan wordt de partij “Links” genoemd, en anders “Rechts.” Het is hierbij wel belangrijk dat de partijnamen correct zijn ingevoerd. Als iemand per ongeluk “Partij voor de Arbeid” heeft ingevuld in plaats van “PvdA” dan wordt de variabele gecodeerd als “Rechts” omdat de langere partijnaam niet in de lijst van linkse partijen voorkomt. Je kunt dit probleem op twee manieren oplossen:

1) Defensief programmeren: expliciet aangeven welke partijen als links en rechts gecategoriseerd moeten worden, en overige waarden categoriseren als “Overig.” Bijvoorbeeld door de formule uit te breiden en de waarde “Rechts” te vervangen met een tweede controle:

```
=IF(ISNUMBER(MATCH(A1,{“PvdA”;“GroenLinks”},0)),
“Links”,
IF(ISNUMBER(MATCH(A1,{“VVD”;“SGP”},0)), “Rechts”, “Overig”))
```

2) De oorspronkelijke data valideren: vooraf of achteraf controleren wat er ingevoerd is. In de meeste spreadsheetprogramma’s kun je aangeven wat de toegestane waarden zijn binnen een kolom, waardoor je ook niets anders kunt invullen.

G.3 EXTRA FUNCTIONALITEITEN

Verschillende organisaties bieden ook extra functionaliteiten aan die je kunt toevoegen aan je favoriete spreadsheetprogramma. Zo is er bijvoorbeeld software om automatisch een sentiment-analyse uit te voeren binnen Excel of Google Sheets, waardoor je automatisch kunt bepalen of een tekst *positief* of *negatief* is. Voor deze software worden verschillende termen gebruikt; online kun je zoeken naar *plugins*, *add-ins* of *add-ons*.

Hoewel het natuurlijk fijn is dat er zoveel analyses mogelijk worden gemaakt via externe software, is het belangrijk om kritisch te blijven. Probeer te begrijpen hoe de software werkt, en vraag jezelf af wat er met de data gebeurt: blijft de data lokaal (en

8 Dit is een versimpeling van de werkelijkheid. Het is nog niet zo eenvoudig om partijen in te delen op het Links-Rechts spectrum, zie bijvoorbeeld https://www.parlement.com/id/vh8lnhrp8wsy/links_en_rechts voor een bespreking.

dus veilig volgens de AVG), of worden alle gegevens naar een onbekende (en dus onveilige) server gestuurd? Daarnaast is het ook belangrijk om alle software te testen voordat je de resultaten gebruikt in je onderzoek. Pas wanneer we weten dat de resultaten betrouwbaar en valide zijn, kunnen we ook degelijke conclusies trekken.

G.4 RESULTATEN BEREKENEN

Wanneer je alle data gecodeerd hebt, kun je de resultaten gaan analyseren. Toetsende statistieken kun je het beste berekenen met programma's zoals *R*, *JASP*, of *JAMOVI*, maar Excel bevat ontzettend veel nuttige functies om descriptieve statistieken te berekenen.

G.4.1 RESULTATEN CONTROLEREN

Voordat je de resultaten gaat analyseren of exporteren naar jouw favoriete statistiekprogramma is het belangrijk om te controleren of je alles goed hebt ingevoerd. Hiervoor kun je drie verschillende functies gebruiken:

1) Gegevensvalidatie (Engels: *data validation*). Je kunt vooraf aangeven welke waarden zijn toegestaan binnen iedere kolom. Je krijgt dan een waarschuwing als je een onverwachte waarde invoert.
2) Voorwaardelijke opmaak (Engels: *conditional formatting*). Als je alle data hebt ingevoerd, kun je voorwaardelijke opmaak gebruiken om verschillende kleuren toe te kennen aan alle cellen in je tabel, afhankelijk van de waardes die in die cellen staan. Dit kan je helpen om te zien of er bepaalde waardes uitspringen.
3) De data visualiseren in een grafiek. Welke grafiek je hiervoor moet gebruiken hangt af van de context. Denk bijvoorbeeld aan een spreidingsdiagram (Engels: *scatter plot*).

G.4.2 RELEVANTE FUNCTIES

Tabel G.1 geeft een overzicht van functies om descriptieve statistieken te berekenen voor een reeks getallen. Deze functies zijn vooral nuttig als je kolommen hebt met verschillende aantallen, of als je verschillende eigenschappen binair hebt gecodeerd met een 1 wanneer die eigenschap aanwezig is, en een 0 wanneer de eigenschap afwezig is.

Functie	Wat doet de functie?	Voorbeeld	Uitleg
SUM	Telt een reeks getallen bij elkaar op.	=SUM(A:A)	Som van alle getallen in kolom A.
AVERAGE	Berekent het gemiddelde cijfer.	=AVERAGE(B:B)	Gemiddelde van kolom B.
MODE	Berekent het meestvoorkomende getal.	=MODE(C:C)	Modus van kolom C.
MEDIAN	Berekent de mediaan.	=MEDIAN(D:D)	Mediaan van kolom D.
MIN	Berekent het kleinste getal in een reeks.	=MIN(E:E)	Kleinste getal van kolom E.
MAX	Berekent het grootste getal in een reeks.	=MAX(F:F)	Grootste getal van kolom F.

TABEL G.1 *Functies om descriptieve statistieken te berekenen voor een reeks getallen.*

Wanneer je kolommen hebt met tekstuele labels, dan heb je andere functies nodig. Als je wil tellen hoe vaak een bepaalde categorie voorkomt, dan kun je de functie COUNTIF gebruiken, bijvoorbeeld:

```
=COUNTIF(A:A, “appel”)
```

Deze formule telt hoe vaak de waarde “appel” voorkomt in kolom A. Als je een volledige frequentietabel zou willen maken voor een bepaalde kolom, dan kun je een aparte formule maken voor iedere waarde, maar de meeste spreadsheetprogramma’s hebben hiervoor ook een ingebouwde functie: *draaitabellen* (Engels: *pivot tables*). Lees in de handleiding van je favoriete programma hoe dat moet.

Werkboek Inhoudsanalyse is een praktische handleiding om te leren hoe je op een betrouwbare en verantwoordelijke manier kunt analyseren hoe mensen met elkaar communiceren. Dit boek richt zich op het ontwerpen en uitvoeren van studies die gebruik maken van kwantitatieve inhoudsanalyse als onderzoeksmethode. Deze eerste editie is ontwikkeld in het kader van de opleiding Communicatie- en Informatiewetenschappen (CIW) van Tilburg University. Binnen één semester voeren schrijven studenten een onderzoeksvoorstel en voeren ze het voorstel vervolgens ook uit. Deze groepsopdracht is volledig uitgewerkt in de bijlagen van het boek.

Emiel van Miltenburg is universitair docent bij de opleiding Communicatie- en Informatiewetenschappen aan Tilburg University. Hij is opgeleid als taalkundige en doet onderzoek naar de ontwikkeling en verantwoorde inzet van taaltechnologie. Sinds 2020 geeft hij jaarlijks college over inhoudsanalyse.

OPEN PRESS TiU