

Appendix A

Computeranalyse van kwalitatieve data

Veel procedures voor *kwantitatieve* data-analyse zijn standaard en worden wereldwijd gebruikt door wetenschappers. Aansluitend daarop worden ook bepaalde programmapakketten voor computers over de hele wereld gebruikt, zoals SPSS en SAS.

In het *kwalitatieve onderzoek* worden voor de analyse van de data eveneens veelvuldig computerprogramma's gebruikt, maar is er veel minder standaardisering. Een van de redenen daarvoor is ongetwijfeld dat de analyse van data ingewikkelder is, omdat data, althans in eerste instantie, niet uitgedrukt zijn in getallen, maar in woorden, niet-numerieke symbolen, geluids- en beeldopnamen, enzovoort. Analyseprocedures zijn gevoelig voor persoonlijke vertekeningen, en in ieder geval tijdrovend door het vele codeerwerk.

Verder spreken we wel over 'kwalitatief onderzoek', maar gaat het om een heterogene verzameling van aanpakken of stromingen, zoals symbolisch interactionisme, etnografie, fenomenologie, etnomethodologie, conversatieanalyse, delen van inhoudsanalyse en nog vele meer. Binnen de meeste van deze tradities worden analyseprocedures gebruikt die specifiek aansluiten bij die stroming, en die minder geschikt zijn voor andere.

Ten slotte merken we op dat het gehele terrein in snelle ontwikkeling is en in een soms onvoorspelbare richting gaat, reden waarom het codificeren en standaardiseren van analyseprocedures sterk wordt bemoeilijkt.

Het gebrek aan standaardisering leidt soms tot het advies om meer dan één standaardprogramma te gebruiken voor de analyse van je data, omdat programma A heel goed kan zijn in functie X, maar niet functie Y kan uitvoeren, wat weer wel kan met programma B. Dit advies kan echter ook erg onpraktisch zijn, in die zin dat bij het invoeren van de data er al verschillen optreden, wat het werk verdubbelt.

De meeste computerprogramma's zijn ontwikkeld door Amerikaanse, Australische, Britse of Duitse onderzoekers en worden verkocht via Scolari, een afde-

ling van Sage Publications (www.scolari.co.uk). In de meeste gevallen kun je hier ook demonstratiepakketten downloaden.

Bekende pakketten zijn bijvoorbeeld: QSR Nvivo 8.0, ATLAS.ti 6.1, MAXQ-DA, The Ethnograph 6.0 en C-I-SAID. In Nederland wordt het programma KWALITAN 5.0 (voor informatie zie www.kwalitan.net) veel toegepast. De ATLAS- en KWALITAN-programma's zijn nogal verankerd in de symbolisch interactionistische traditie, in het bijzonder de zogenoemde 'grounded theory approach' (in de kern een inductivistische onderzoekstraditie). Overigens zijn bijna alle programma's *data-driven*, en in deze zin inductief. In het reclamemateriaal treffen we beweringen aan zoals 'relating text fragments leads you to the discovery of the texture of the data' en 'visual theory building with the semantic network editor allows you to make relationships between emerging concepts visible'. Kwalitatief onderzoek is in de eerste plaats explorerend van karakter. Alle programma's gebruiken als primaire database teksten (bijvoorbeeld interviewprotocollen) en waarnemingsprotocollen. Bij sommige kun je ook met grafische, geografische, audio- en videodata als invoer werken. Toepassingen zijn er in vele disciplines. Als illustraties noemen we in de gezondheidswetenschappen de analyse van röntgenbeelden, tomogrammen en microscopische beelden, en in de criminologie het gecombineerde gebruik van brieven, vingerafdrukken en foto's in de analyse.

Als je een grote hoeveelheid kwalitatieve data wilt analyseren, heb je echter bijna een consumentengids nodig om je weg te vinden in de veelheid van het aanbod. Binnen het kader van dit boek is het niet mogelijk om de lezer een overzicht van de mogelijkheden te bieden. Het zou bovendien weer snel verouderd zijn. We zien vaak dat bij een bepaalde universitaire faculteit om onduidelijke redenen een bepaald programmapakket in gebruik is, dat dan ook voor van alles en nog wat wordt gebruikt, terwijl voor dezelfde mix van doeleinden ergens anders een ander pakket wordt gebruikt. De prijsstelling is helaas nog zodanig dat het voor de individuele gebruiker nauwelijks mogelijk is een pakket aan te schaffen, laat staan méér dan één, en licentiecontracten, zoals voor SPSS, zijn er nauwelijks.

Wat zijn de belangrijkste kenmerken die deze programma's gemeen hebben?

1. De primaire documenten worden gescand en ingevoerd. Het gaat om interviewprotocollen, artikelen, boeken, foto's, films, observatieprotocollen, enzovoort.
2. Elk document wordt in secties gesplitst. Als het om teksten gaat, kan een sectie een paragraaf, een zin, een deel van een zin of een woord zijn.
3. Aan elke sectie worden één of meer codes toegekend. Normaliter is dat een term of een centraal kenmerk van de inhoud die in die sectie gevonden wordt.

4. Codes worden verzameld in wat je een codeboek zou kunnen noemen. Dit wordt in het bestand van het project opgenomen.
5. De stappen 1-3 worden voor elk primair document doorlopen.
6. Gedurende het gehele proces schrijft de onderzoeker memo's. Een memo is een notitie over een gevolgde procedure, een plotseling verkregen inzicht, een theoretische interpretatie, een methodologische kanttekening, enzovoort.
7. De primaire documenten vormen samen met de toegekende codes, de memo's en het 'codeboek' het projectbestand.

We kunnen nu een begin maken met een eerste, 'gemakkelijke' analyse. We kunnen documenten vergelijken op de frequentie van allerlei codes, we kunnen kruistabellen maken voor de combinatie van codes, we kunnen dergelijke analyses uitvoeren voor een gedeelte van de secties, voor een gedeelte van de onderzochten, enzovoort.

Dit kan allemaal erg gemakkelijk lijken, maar er blijven twee belangrijke problemen:

1. Hoe moet je je primaire documenten in secties verdelen?
2. Hoe moet je secties coderen?

Ad 1. Hier kan in het algemeen geen antwoord op worden gegeven. Het hangt geheel van de aard van het ruwe materiaal af, en van de probleemstelling. Bij interviewverslagen kan de opsplitsing erg voor de hand liggen als er een paar goed omliggende onderwerpen zijn; dat deze onderwerpen dan de secties vormen, is vanzelfsprekend.

Ad 2. Coderen is nogal ingewikkeld omdat het verschillende doeleinden dient. In de eerste plaats kan een sectie worden geïdentificeerd als een code is toegekend. Hierdoor worden het opslaan en de toegankelijkheid van de data vergemakkelijkt.

In de tweede plaats, en afhankelijk van de probleemstelling, wordt een code toegekend die karakteristiek is voor de inhoud, de aan- of afwezigheid van een bepaalde gedachte of een bepaalde term, enzovoort. Soms is dit gemakkelijk (zoals het al dan niet noemen van een financieel motief in een interview over het maken van de keuze voor een academische studie), soms vereist het dat je in de eerste fasen van het codeerwerk geleidelijk aan een thesaurus of een dictionaire opbouwt. Het is vergelijkbaar met woordenboeken als je een vreemde taal bestudeert, of met het invoeren van synoniemenlijstjes in een wordprocessor. In interviews over politieke houdingen en gedragingen zou de

onderzoeker bijvoorbeeld alle negatieve uitlatingen van de respondent over de politiek op een bepaald terrein willen classificeren als ‘tegen de politiek’. Om dat te kunnen doen moet je wel eerst enkele tientallen interviewverslagen lezen om een min of meer complete lijst van negatieve uitlatingen op te kunnen stellen. Nadat een nieuw woord of een nieuwe uitdrukking is ingevoerd, herkent de computer dat woord of die uitdrukking, en codeert de volgende documenten in deze zin. Op sommige terreinen, bijvoorbeeld bij de communicatiewetenschappen, bestaan dergelijke dictionaires en thesauri. Het coderen vereist het zorgvuldig lezen en vergelijken (bijvoorbeeld of er in bepaalde secties van observatieprotocollen van spelende kinderen sprake is van ‘voor de ander zorgen’). Met betrekking tot deze functie zijn we al halverwege naar meer geavanceerde soorten van analyse.

Er is nog een derde functie waarmee we op het terrein komen van de meer geavanceerde analyse. De inhoud van een sectie wordt vaak in eerste instantie uitgedrukt in een code op een laag niveau van abstractie, maar kan geleidelijk aan worden vervangen door, of aangevuld worden met, een code op een hoger abstractieniveau. Een voorbeeld: regen / neerslag / het weer / het klimaat.

Op deze manier kunnen we verbindingen maken tussen codes die op een verschillend niveau van abstractie liggen, we kunnen indelingen en categorieën maken om classificaties van een hogere orde te bereiken, om uitspraken te formuleren in termen van een theorie die bij de data past, en om zulke uitspraken te toetsen of ze ook op andere delen van het materiaal van toepassing zijn. Een dergelijk abstraheringsproces kan diensten bewijzen als het gaat om het construeren van (‘grounded’) theorieën, dat wil zeggen theorieën die gebaseerd zijn op observaties.

Resultaten van de analyse van kwalitatieve data kunnen worden uitgedrukt in frequentietellingen, het laten zien hoe woorden bij bepaalde contexten horen, en in plaatjes van grafische netwerken tussen de ontwikkelde concepten en hun onderlinge relaties. Die relaties hoeven niet altijd causaal te zijn: ‘hoort bij’ en ‘is een soort van’ zijn ook mogelijk. Het is belangrijk om in dit opzicht precies te zijn, vooral omdat het gebruik van termen als ‘variabelen’, ‘waarden’, ‘causaliteit’ en ‘samenhang’ in bepaalde takken van het kwalitatief onderzoek sterk verschilt van de betekenisgeving in het kwantitatieve (mainstream-) onderzoek.

Literatuur

- Bazeley, P. & Richards, L. (2000). *The Nvivo qualitative project book*. London: Sage.
- Kuckartz, U. (2001). *An introduction to the computer analysis of qualitative data*. London: Sage.
- Lewins, C. & Silver, C. (2007). *Using software in qualitative research: A step by step guide*. London: Sage.
- Richards, L. (2010). *Handling qualitative data. A practical guide* (2nd ed.). London: Sage.

