



Centraal Bureau voor de Statistiek

Programma Impact ICT

Onderzoeksrapport nr. 6

Medegefinancierd door het Ministerie van Economische Zaken en Prima

**Eindrapportage
Onderzoek inzet internetwaarneming bij de
vacaturestatistieken
2011 en 2012**

N. Heerschap

op basis van notities van J. Daalmans, D. ter Steege en M. Smeets

Kennisgeving:

De in dit rapport weergegeven opvattingen zijn die van de auteurs en komen niet noodzakelijk overeen met het beleid van het Centraal Bureau voor de Statistiek

Inhoud

1	Doel van notitie	4
2	Achtergrond en aanleiding	5
3	De vacaturestatistieken en internet	7
	3.1 De vacaturestatistieken	7
	3.2 Internet en vacatures	8
4	Belangrijkste resultaten onderzoek 2011 en 2012	12
	4.1 Eerste fase: onderzoek 2011	12
	4.2 Tweede fase: onderzoek 2012	13
	4.3 Eindconclusies	17
5	Business case nul-scenario	21
6	Aanbevelingen voor een vervolg	23
7	Referenties	25

Samenvatting

In 2011 en 2012 is verkennend onderzoek gedaan naar de mogelijkheden om internetwaarneming in te zetten bij de productie van statistieken over vacatures bij het CBS. Daarvoor zijn analyses uitgevoerd op bestanden met internetvacatures (Jobfeed) van het bedrijf Textkernel. Dit bedrijf is gespecialiseerd in het waarnemen, structureren en classificeren van op internet waargenomen vacatures.

Doelstelling van de pilot was om te onderzoeken of het mogelijk is om internetwaarneming in te zetten bij het maken van de vacaturestatistieken van het CBS en dan met name 1) ter vervanging van of als aanvulling op de huidige waarneming van de Structuurenquête Vacatures en 2) het beperken van de huidige uitvraag van de Kwartaalstatistiek Vacatures.

Deze nota bevat een evaluatie van de tot nu toe behaalde resultaten, een business case voor de inzet van één van de onderzochte scenario's en mogelijkheden voor vervolgonderzoek.

Belangrijkste conclusies:

De belangrijkste conclusie van het onderzoek is dat de mogelijkheden om internetwaarneming bij de statistieken over vacatures bij het CBS in te zetten vooralsnog beperkter zijn gebleken dan bij het begin van het onderzoek was gedacht. Meer onderzoek is nodig om tot een goede inschatting te komen en een vergelijking te kunnen maken met de bestaande vacaturestatistieken bij het CBS. Wel lijken er kansen te liggen om internetwaarneming met bestaande statistieken te combineren, bijvoorbeeld voor enquêtedrukvermindering, voor verdere detaillering (o.a. regionale cijfers), als aanvullende informatie in het verwerkingsproces en als snelle bèta-indicatoren.

Een aantal factoren speelt een rol bij de beperktere mogelijkheden om op dit moment internetwaarneming in te zetten bij de vacaturestatistieken van het CBS. Hier worden de belangrijkste genoemd. Hoewel de dekkinggraad de laatste jaren sterk lijkt te zijn toegenomen, worden, ten eerste, niet alle vacatures op internet geplaatst. Dit kan per bedrijfstak verschillen. Een tweede factor zijn de vacatures die op internet bij uitzendbureaus worden waargenomen. Door de vaak beperkte informatie over deze vacatures kunnen deze niet altijd aan de juiste bedrijfstak of regio worden toebedeeld, maar ook het ontdebelen van deze groep vacatures is minder gemakkelijk. Het gaat hier om forse aantallen vacatures. Hoewel waarschijnlijk minder groot van omvang, speelt, ten derde, het fenomeen bulkvacatures een rol, waarbij meer dan één vacature wordt gepresenteerd in dezelfde internetvacature. En, ten slotte, spelen bij internetwaarneming allerlei invloeden een rol, die geheel losstaan van de daadwerkelijke ontwikkeling van het aantal vacatures. Te noemen zijn het toenemend gebruik van internet als wervingskanaal, opkomst van nieuwe wervingskanalen zoals sociale media en gratis job boards en het verbeteren van het spideren en de wijze van ontdebelen (methodologie). Deze hiaten en ruis in de waarneming en verwerking maken het op dit moment moeilijk in te schatten hoe representatief de aantallen internetvacatures en de ontwikkeling daarvan zijn voor de daadwerkelijke aantallen en ontwikkelingen van vacatures. Er zijn (nog) onvoldoende gegevens beschikbaar om de omvang en richting van deze factoren goed te kwantificeren. Dat is echter wel nodig om internetwaarneming op een verantwoorde wijze bij de vacaturestatistieken van het CBS te kunnen inzetten.

Wel moet worden opgemerkt dat de meting van en ontwikkelingen in het vacatureaanbod in 2011 via data van Jobfeed beter aansluiten bij de cijfers uit de Kwartaalstatistiek Vacatures dan in 2010 het geval was.

In 2012 zijn er ook acht scenario's onderzocht om met behulp van internetinformatie de uitvraag voor de Kwartaalenquête Vacatures te beperken. De beste mogelijkheden voor een reductie van de uitvraag lijken te liggen bij het zogenaamde nul-scenario¹, waarbij geen vragenlijsten worden uitgezonden naar die bedrijven die in de Jobfeed data geen vacatures hebben. Er is namelijk een relatie geconstateerd tussen bedrijven die bij de Kwartaalenquête Vacatures opgeven geen vacatures te hebben en waar Jobfeed ook geen vacatures op internet waarneemt, en vice versa. Als naar die bedrijven bijvoorbeeld in drie kwartalen geen vragenlijsten zouden worden gestuurd, zou dat al een aanzienlijke beperking van de enquêtedruk opleveren (14 – 45%). Voor dit scenario wordt in deze notitie een business case gepresenteerd. De geschatte eenmalige kosten zijn € 10.000 en de structurele kosten op jaar basis € 22.000 voor met name het verkrijgen van de data van Jobfeed. Daartegenover staan als baten: een beperking van de uitvraag, mogelijk een betere kwaliteit van de statistiek en vooral kansen om het imago van het CBS te verbeteren. Er wordt vanuit gegaan dat er één jaar wordt schaduwgedraaid. Een bottleneck om de voordelen van het nul-scenario ten volle te benutten is het feit dat de uitvraag van de Kwartaalenquête Vacatures gekoppeld is aan de uitvraag van de statistiek Ziekteverzuim.

Juridische aspecten:

Bij het waarnemen van vacatures op internet spelen juridische aspecten een ondergeschikte rol. Het betreffen geen persoonsgegevens². Wel is het zo dat aan een eventueel uitwisseling van gegevens tussen CBS en Textkernel strenge randvoorwaarden moeten worden gesteld.

Suggesties voor vervolg:

Met deze eindrapportage zijn de werkzaamheden voor het Programma Impact ICT afgerond. Verder onderzoek en eventuele implementatie zullen geschieden op basis van eigen middelen van het CBS. Voor 2013 is daarvoor in de beleidsreserves geld gereserveerd. Daarmee heeft het Programma Impact ICT zijn aanjagersfunctie vervuld.

De volgende vervolgstappen worden gesuggereerd:

- De aanschaf van Jobfeed data voor minstens nog één jaar (2012 en mogelijk ook 2013). Hiermee kan worden beoordeeld of de betere fit tussen de ontwikkeling van de internetvacatures (Jobfeed) en de Kwartaalenquête Vacatures van 2011 zich handhaaft. Zoals aangegeven lijkt de dekkingsgraad van de internetvacatures de laatste jaren sterk te zijn toegenomen. Bovendien is de werkwijze van het waarnemen en verwerken van internetvacatures bij Jobfeed verbeterd. Hiermee kan het zijn dat de invloed van de hiaten en ruis in de internetwaarneming geringer is geworden dan uit dit onderzoek naar voren is gekomen.
- Onderzoek om beter grip te krijgen op de omvang en richting van de genoemde hiaten in en de ruis op de internetwaarneming. Lukt het niet om hierop beter grip te krijgen dan blijft het moeilijk internetwaarneming, buiten het nul-scenario om, op een verantwoorde wijze in te zetten voor het maken of ondersteunen van CBS-waardige vacaturestatistieken.
- Eén jaar schaduwdraaien met het nul-scenario, en dan met name variant 4 (grootteklasse 4-7). Het lijkt verstandig de beslissing over deze stap pas te nemen als de stap bij het eerste pijltje, dat wil zeggen: nog een jaar Jobfeed data, is afgerond.

¹ Dit betreft dus een scenario met bedrijven met nul vacatures en niet een scenario waarbij niets wordt gedaan (terminologie business analyse).

² In zeer beperkte mate zou het kunnen gaan om eenmanszaken. In dat geval is er sprake van persoonsgegevens.

- *Verder onderzoek naar andere mogelijkheden om internetwaarneming in te zetten bij de vacaturestatistieken van het CBS, bijvoorbeeld als hulpinformatie bij het verwerkingsproces, voor detaillering van de cijfers (o.a. voor de Structuurenquête Vacatures) en als snelle bèta-indicatoren als aanvulling op de huidige output.*
- *Bij de vacaturestatistieken is men gestart met onderzoek naar de behoefte aan cijfers over de vacaturemarkt. Op basis van die behoefte kan vervolgens gekeken worden naar de bronnen, die nodig zijn om die cijfers te maken. Daarbij kan internetwaarneming een rol spelen. Onderdeel daarvan zou ook moeten zijn om na te gaan in hoeverre gebruikers van de vacaturestatistieken geïnteresseerd zijn in snellere en meer gedetailleerde cijfers, zogenaamde bèta-indicatoren, over de vacaturemarkt, maar dan puur gebaseerd op internetwaarneming (vergelijk de huidige maandelijks vacature indicator van NR)? Ook al weet men dat de representativiteit te wensen overlaat.*

Hierbij behoort ook de vraag in hoeverre het CBS het als haar taak ziet om dit soort bèta-indicatoren te produceren? Is het antwoord op deze vragen ja en gaat men verder met Jobfeed data, dan dient men zich te realiseren dat aan zo'n variant een aanzienlijk hoger kostenplaatje hangt. Dat zal voor de Jobfeed data liggen rond de 100.000 euro per jaar (exclusief BTW). Mogelijke alternatieven zijn: het hierin gezamenlijk optreden met het UWV en om te kijken naar andere aanbieders van internetdata over vacatures. Bij het laatste alternatief speelt ook de kwaliteit en continuïteit van die aanbieders een rol.

Trefwoorden: Jobfeed, internetrobots, waarneming op het internet, vacatures, lastendrukverlichting

1 Doel van notitie

De ontwikkeling van internet verschaft de mogelijkheid om deze als nieuwe databron te gebruiken voor het maken van statistieken. Dat geldt ook voor het terrein van vacatures. In 2011 en 2012 is door het CBS dan ook onderzocht in hoeverre internetdata de waarneming bij statistieken over de vacaturemarkt kunnen ondersteunen danwel vervangen.

In dit kader heeft deze notitie een aantal doelstellingen:

1. Om achtergrond te verschaffen over de tot nu toe uitgevoerde analyses en de behaalde resultaten (hoofdstukken 2 - 4).
2. Een business case te presenteren van één van de onderzochte scenario's die kan leiden tot enquêtedrukvermindering, het zogenaamde nul-scenario (hoofdstuk 5).
3. En te kijken naar mogelijk vervolgonderzoek (hoofdstuk 6).

Deze notitie is, samen met de andere rapportages (m.n. Smeets en Daalmans, 2012, en Daalmans en Smeets, 2012), tevens de eindverantwoording van het project richting het Programma Impact ICT en daarmee het ministerie van Economische Zaken (Prima), dat het onderzoek in 2011 en 2012 voor 50% heeft gefinancierd.

2 Achtergrond en aanleiding

Internet is een fundamentele infrastructuur geworden van de samenleving. Steeds meer activiteiten uit de “fysieke” wereld hebben zich verplaatst naar internet. Daarmee is internet niet alleen een fenomeen geworden dat zelf vraagt om een statistische beschrijving, maar ook een nieuwe bron van informatie, die voor het maken van statistieken kan worden benut. Dit is mogelijk omdat activiteiten op internet digitale sporen, *digital footprints*, achterlaten. Deze digitale sporen kunnen 24 uur per dag en 7 dagen per week worden waargenomen. De waargenomen internetdata vormen vervolgens de basis voor het maken van statistieken. Potentiële voordelen van deze internetstatistieken t.o.v. traditionele statistieken zijn onder meer sneller en gedetailleerder, met minder of geen enquêtedruk. Daarbij gaat het niet alleen om de totale vervanging van de traditionele waarneming, maar vooral om een combinatie van internetwaarneming en bestaande waarneming.

Ook als het gaat om het werven van personeel speelt internet een belangrijke rol (Intermediair, 2011). De meeste bedrijven hebben tegenwoordig wel een website, waarop zij informatie verstrekken over de bij hen beschikbare vacatures. Daarnaast zijn er intermediairs, zogenaamde job boards, die internetservices aanbieden waarop bedrijven hun vacatures kunnen plaatsen. En ten slotte, neemt ook het gebruik van sociale media voor het werven van personeel toe.

Het verzamelen van data van internet kan op verschillende manieren (Dialogic, 2008). Eén van de mogelijkheden is het inzetten van zogenaamde internetrobots, –crawlers of –spiders. Het betreft software (browsers), waarmee internet kan worden afgezocht naar de gewenste informatie. Daarbij kan het gaan om het benaderen van één of meer specifieke websites, bijvoorbeeld de website van een bedrijf of een job board. Maar het kan ook gaan om een scan van een deel van internet, bijvoorbeeld door het Nederlandstalige deel van internet af te zoeken naar begrippen als “vacature” en “baan”. In dat geval wordt gesproken over “web scraping”. De met internetrobots verzamelde informatie wordt centraal opgeslagen in een database en verwerkt. De verwerkte internetinformatie vormt vervolgens de basis voor het maken van statistieken.

Onderzoek naar de toepassingsmogelijkheden van internetrobots voor het maken van statistieken loopt al sinds 2009 bij het CBS. In eerste instantie ging het daarbij om prijzen van producten en diensten, zoals vliegtickets en benzine. Vanaf eind 2010 zijn daar onder de paraplu van het programma Impact ICT (Heerschap, 2011) de domeinen woningmarkt en vacatures aan toegevoegd. Het domein van vacatures is vooral in beeld gekomen, doordat steeds meer private partijen statistieken over het aantal en de ontwikkeling van vacatures publiceren die geheel gebaseerd zijn op internetwaarneming. Hoewel men vaak vraagtekens kan zetten bij de representativiteit van dit soort cijfers, trekt deze nieuwe manier van statistiekmaken steeds meer de aandacht van beleidsmakers en in het geval van vacatures van instanties zoals het UWV, het CPB en het ministerie van Werkgelegenheid en Sociale Zaken (Binnenlands Bestuur, 2012). Daarbij kon het CBS niet aan de zijlijn blijven staan en is eind 2010 gestart met onderzoek naar de mogelijkheden om internetwaarneming in te zetten bij het maken van de vacaturestatistieken. In eerste instantie als vervanging van de waarneming van de tweejaarlijkse Structuurenquête Vacatures en later ook als mogelijkheid om de waarneming bij de Kwartaalstatistiek Vacatures te beperken met als doel lastendrukvermindering.

De eerste gedachte was om zelf robots te bouwen en deze in te zetten bij de waarneming van de top 200 bedrijven. Hiermee zou een aanzienlijk deel van de vacaturemarkt worden gedekt en het zou direct leiden tot enquêtedrukvermindering. Daarvan is afgezien omdat dit zou betekenen het bouwen en onderhouden van 200 robots, maar vooral ook omdat, in eerste instantie via het UWV, de beschikking kon worden verkregen over internetdata uit de vacature-database Jobfeed

van het bedrijf Textkernel. Textkernel neemt al voor een groot aantal jaren “alle” Nederlandse vacatures op internet waar, maar belangrijker: zij zijn vooral gespecialiseerd in de cruciale stappen nadat de webpagina's met vacatureteksten zijn gespiderd. Te noemen zijn het extraheren van de gewenste vacature-informatie uit de gespiderde webpagina's, het proces van het ontdebelen van vacatures die op meer dan één plek op internet zijn geplaatst, het standaardiseren en classificeren van de geëxtraheerde gegevens en, ten slotte, de koppeling van de vacatures aan Kamer van Koophandelnummers. Uiteindelijk leidend tot een database, Jobfeed, met alle unieke vacatures, die over de jaren heen op het Nederlandse deel van internet zijn geplaatst, met variabelen zoals datum voor het eerst verschenen, datum van het internet verdwenen, Kamer van Koophandelnummer, SBI, postcode, naam van het bedrijf, beroep en functie.

In 2011 en 2012 is op basis van de Jobfeed data (cijfers van 2010 en 2011) door het CBS onderzoek gedaan naar de mogelijkheden om internetwaarneming in te zetten bij de vacaturestatistieken. Voordat in hoofdstuk 4 de belangrijkste resultaten worden beschreven van dit onderzoek, wordt in hoofdstuk 3 achtergrondinformatie verschaft over de vacaturestatistieken en internet.

3 De vacaturestatistieken en internet

3.1 De vacaturestatistieken

Bij de vacaturestatistieken gaat het om:

1. *De Kwartaalstatistiek vacatures*. Bij deze statistiek wordt gevraagd naar de openstaande vacatures aan het eind van een kwartaal (standcijfer) en gedurende het kwartaal ontstane, vervulde en vervallen vacatures (stroomvariabelen).

Het CBS houdt deze enquête bij circa 15-20 duizend bedrijven en instellingen per kwartaal in zowel de particuliere- als in de overheidssector, inclusief het onderwijs. Doel is om uiterlijk 45 dagen na afsluiting van een kwartaal cijfers te publiceren over de ontwikkeling van de vacaturemarkt, uitgesplitst naar sector (SBI-groep) en grootteklasse. De statistiek is gebaseerd op een Europese verordening.

De waarneming bestaat uit een combinatie van enquêteformulieren, bellen en elektronisch waarnemen. Hierbij wordt, naast het aantal vacatures, ook gevraagd naar het ziekteverzuim. Dit veroorzaakt enquêtedruk.

Bij het gaafmaken wordt de volgende rekenregel gehanteerd:

$$\text{openstaande vacatures } t = \text{openstaande vacatures } t-1 + \text{nieuwe vacatures} - \text{vervulde en vervallen vacatures}$$

Indien de ruwe gegevens niet aan deze regel voldoen is dat reden voor aanpassing, waarbij de variabele “openstaande vacatures” het minst wordt aangepast. De aanpassingen zijn soms aanzienlijk.

2. *De Structuurenquête vacatures*. De vragen uit de Kwartaalstatistiek zeggen niets over wat voor soort vacatures er zijn. Om hierachter te komen, wordt de vraag naar vacatures eens in de twee jaar uitgebreid met een extra vragenlijst (3^{de} kwartaal). Deze enquête wordt de Structuurenquête genoemd. Daarbij gaat het om de volgende kenmerken:

- Gemeente waar de vacatures openstaan;
- Functie;
- Aantal medewerkers aan wie leiding moet worden gegeven;
- Vereiste opleiding en vakrichting;
- Aantal vacatures waarvoor geen werkervaring is vereist;
- Aantal vacatures waarvoor korter dan één maand wordt gezocht;
- Aantal vacatures waarvoor tussen de één en drie maanden wordt gezocht;
- Aantal vacatures waarvoor een personeelsadvertentie is geplaatst;
- Aantal vacatures dat is aangemeld bij het Centrum voor Werk en Inkomen;
- Aantal vacatures dat u als moeilijk vervulbaar beschouwt;
- Aantal vacatures dat alleen is bestemd voor interne kandidaten.

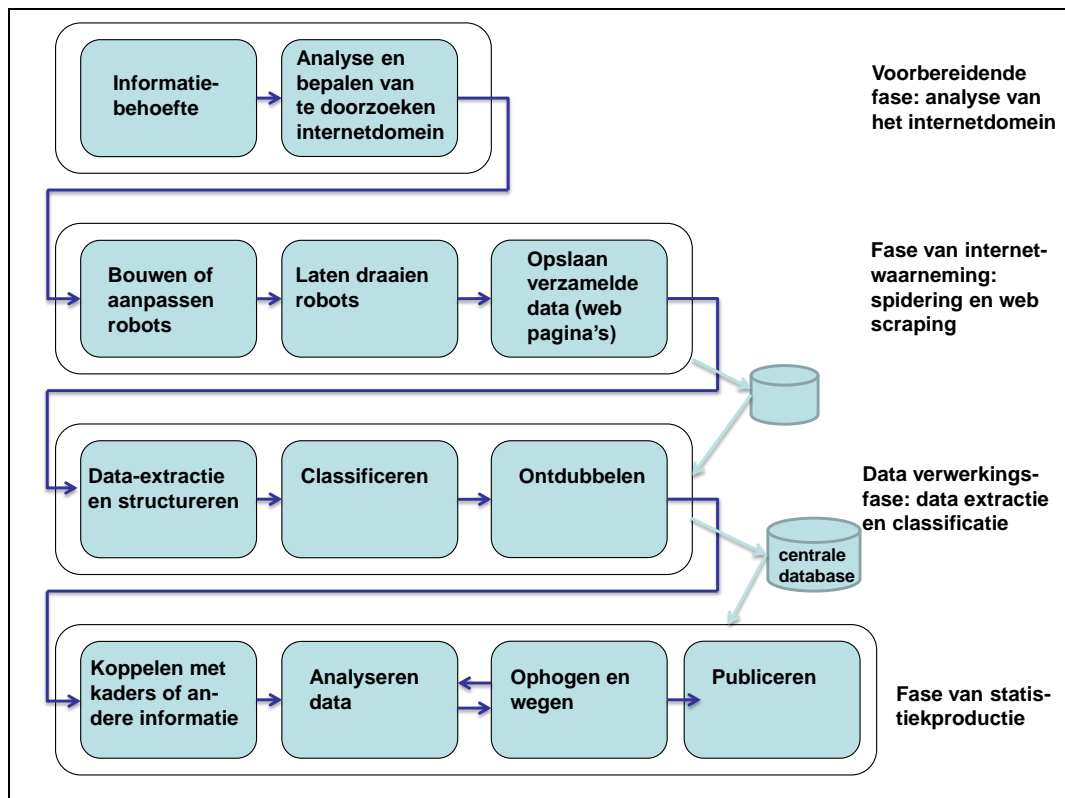
De Structuurenquête is gebaseerd op een gentleman's agreement met de Europese Unie. Maar met name bij het UWV is behoefte aan de informatie uit deze statistiek. De laatste Structuurenquête is in 2008 gehouden. In 2010 is de statistiek uitgesteld om te bezien of er alternatieven waren voor de huidige uitvraag, die veel administratieve lasten met zich meebrengt³. Als één van de alternatieven wordt internetwaarneming gezien. Daarbij was het geen

³ De Structuurenquête telt op zich niet mee in de officiële berekening van de enquêtedruk, omdat er geen sprake is van een verplichte statistiek.

probleem dat met internetwaarneming niet alle hierboven genoemde variabelen kunnen worden gemeten. Hoewel voor sommige variabelen proxy's kunnen worden gebruikt.

3. Ten slotte, wordt er bij de Nationale Rekeningen (NR, conjunctuur) een *vacature-indicator* gepubliceerd die maandelijkse indicaties geeft over de ontwikkelingen in de vacaturemarkt. Deze indicator is afgeleid uit verschillende variabelen van de maandelijkse conjunctuuronderzoeken die door het CBS en het Economisch Instituut voor de Bouw worden gehouden. De vacature-indicator geeft een kwalitatieve indicatie van de richting waarin de vacatures zich naar verwachting van de ondernemers zullen ontwikkelen (zie verder Ruth en Wekker, 2009 en 2011).

Figuur 1. Het maken van statistiek op basis van internetrobots (Heerschap en Kuipers, 2012).



3.2 Internet en vacatures

Het werven van personeel via het adverteren met of beter gezegd het *posten* van vacatures heeft altijd al op internet plaatsgevonden. De meeste bedrijven, die op internet actief zijn, plaatsen de bij hen beschikbare vacatures op hun website. Een belangrijke ontwikkeling aan het eind van de vorige eeuw is de opkomst van job boards geweest. Deze job boards bieden specifieke vacaturesites aan, waarop bedrijven en instellingen hun vacatures tegen betaling kunnen plaatsen. Het aantal van dit soort sites in Nederland is in de loop van tijd sterk gegroeid, namelijk tegen de 600 (Intermediair, 2011 en Dees, 2012A). De belangrijkste zijn o.a. Werk.nl (van het UWV), Monsterboard.nl, NationaleVacaturebank.nl, Jobtrack.nl, Stepstone.nl en Intermediair.nl. Naast deze algemene vacaturesites, zijn er ook kranten, die hun vacature-aanbod op internet plaatsen en sites die zich op specifieke doelgroepen richten, zoals JobsInFinance en Villamedia (journ-

listiek). Door het grote aanbod van vacatures op verschillende plekken zijn er services (zogenoemde aggregate-sites) ontstaan die voor de werkzoekende in één keer alle of veel vacaturesites doorzoeken. Voorbeelden zijn AskJim.nl, CareerJet.nl en JobRapido.nl.

Met de snelle opkomst van sociale media in het laatste decennium is het niet verbazingwekkend dat ook via deze kanalen steeds vaker personeel wordt geworven. LinkedIn is het bekendste voorbeeld, maar er wordt ook steeds meer gebruik gemaakt van andere bekende sociale media zoals Facebook, Hyves, Twitter en Youtube. Een nieuwe ontwikkeling van de laatste paar jaar is de toename en daarmee het belang van gratis vacaturesites, zoals Banenmatch. De laatste twee ontwikkelingen, sociale media en gratis vacaturesites, hebben er toe bijgedragen dat de traditionele job boards het moeilijk hebben (zie Boerman, 2013). Ten slotte zijn er organisaties die het Internet afzoeken naar vacatures en de gestandaardiseerde informatie daarvan opslaan in een database met variabelen zoals postcode, opleiding, beroep, url e.d. De opgeslagen informatie wordt aangeboden voor gebruik door derde partijen, onder meer om marktonderzoek te doen en ondersteuning te verlenen aan werkzoekenden (o.a. het UWV), maar ook om statistieken van te maken. Workinglife.nl en Textkernel (Jobfeed) zijn dergelijke bedrijven. In dit onderzoek is gebruik gemaakt van de gegevens van Textkernel.

Tabel 1. Marktbereik internet als wervingskanaal, 2013 (bron: Ruis en Grijpstra, 2013).

Totaal en sectoren	Internet in %
Totaal	88
Per Sector	
Gezondheidszorg	94
Welzijn	93
Vervoer en opslag	92
Openbaar bestuur en onderwijs	92
Informatie en communicatie	91
Overige dienstverlening	90
Onroerend goed en overige zakelijke diensten	90
Financiële dienstverlening	89
Horeca	89
Specialistische zakelijke diensten	88
Bouwnijverheid	86
Groothandel	85
Detailhandel	83
Industrie	80
Agrarische sector	62

Nadere analyse van de van Textkernel verkregen data geven aan dat het niet zonder meer mogelijk is om internetwaarneming één op één over te nemen voor het maken van kwalitatief goede statistieken. Aspecten, die spelen bij het waarnemen van vacatures op internet, zijn onder meer:

- Een deel van de beschikbare vacatures bij bedrijven en instellingen wordt niet op internet geplaatst. Er zijn verschillen per bedrijfstak. Dit ontbrekende deel lijkt in de loop van de jaren echter wel een heel stuk kleiner te zijn geworden. Zo gaf onderzoek in 2006 (Grijpstra e.a., 2010) nog aan dat de gemiddelde dekkingsgraad van internetvacatures ten op-

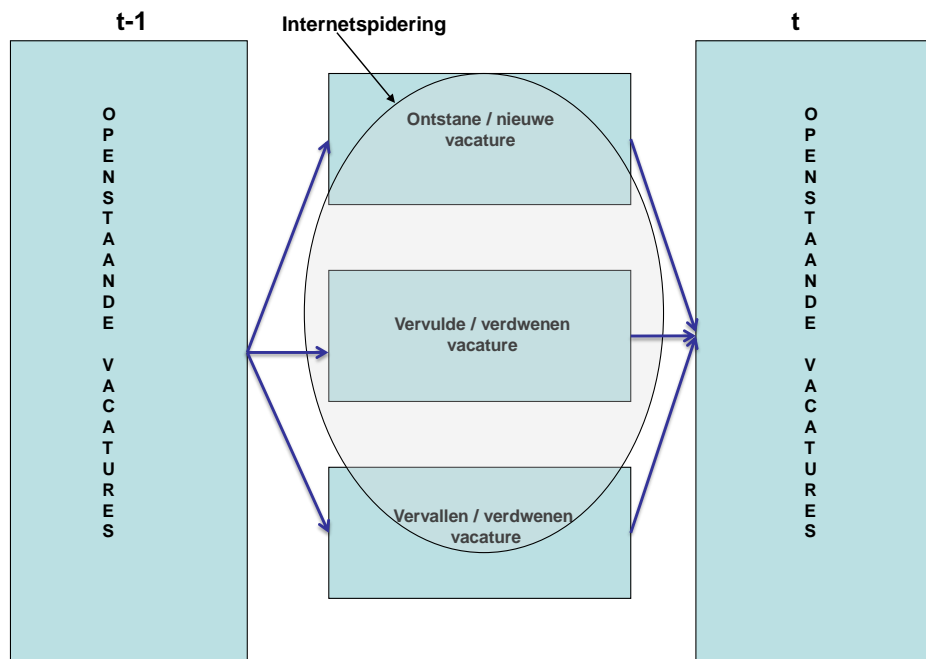
zichte van alle vacatures niet groter was dan 40%. In onderzoek uit 2013 (Ruis en Grijpstra, 2013) wordt aangegeven dat dit nu 88% is. Zie tabel 1.

- Een essentiële stap in het proces is het vaststellen van het aantal *unieke* vacatures op internet, of wel het detecteren van dezelfde vacature die op meer dan één plek op internet is gepost. Dit proces van ontdebelen is geen sinecure. Het is niet alleen van belang hoe uit de tekst van de op verschillende plekken geposte internetvacatures wordt geconstateerd dat er sprake is van een dubbele vacature, ook is het van belang hoe ver men in de tijd teruggaat met deze controle. Zo werd bij Jobfeed in eerste instantie 3 weken teruggegaan om vacature-teksten met elkaar te vergelijken. De overstap naar een periode van 6 weken terug lijkt betere resultaten op te leveren. Een extra complicatie vormt het groeiende belang van gratis vacature-sites en sociale media, waardoor dezelfde vacature op nog meer plekken op internet te vinden is, vaak met een (net iets) andere omschrijving.
- Als vacatures worden waargenomen bij uitzendbureaus is er meestal te weinig informatie beschikbaar, zoals de naam van de werkgever, het adres e.d., om de vacature goed te identificeren. Koppeling met bijvoorbeeld Kamer van Koophandelnummer of classificatie naar regio en SBI is door het ontbreken van gegevens dan meestal niet goed mogelijk. Dat maakt het moeilijk om deze vacatures te betrekken bij het ontdebelen en uiteindelijk het maken van statistieken. Volgens Grijpstra e.a. (2010) gaat het hier om grote aantallen.
- Een ander aspect betreft de zogenaamde bulkvacatures, waarbij in één advertentie wordt gevraagd naar bijvoorbeeld 3 programmeurs. Het is niet gemakkelijk uit de tekst te halen dat er in feite sprake is van 3 in plaats van 1 vacature. De indruk is dat het slechts gaat om een relatief klein aantal vacatures (Grijpstra e.a., 2010).
- Het aantal op internet waar te nemen vacatures kan worden beïnvloed door strategieën van job boards en laksheid van bedrijven en instellingen. Zo is het mogelijk dat vacatures nadat zij al zijn vervuld nog op internet blijven staan. Het aantal vacatures dat een job board heeft, bepaalt bijvoorbeeld de plek die zo'n service inneemt op de lijst met belangrijkste vacaturesites (Dees, 2012B). Ook worden er termijnen gehanteerd dat een vacature op een betaalde job board staat, bijvoorbeeld een maand of zes weken (zie Van Velzen e.a., 2011).
- Het aantal op internet waargenomen vacatures kan ook worden beïnvloed door ontwikkelingen die geheel losstaan van het daadwerkelijke aantal vacatures. Zo kan een geconstateerde groei (of daling) van het aantal op internet waargenomen vacatures ook het gevolg zijn van het toenemende (of afnemend) gebruik door bedrijven en instellingen van internet als wervingskanaal of een verschuiving tussen wervingskanalen. Het daadwerkelijke aantal vacatures hoeft in feite niet te zijn gestegen (of gedaald). Ook wijzigingen in methoden kunnen dit soort ruis opleveren, zoals een verbetering in het spideren om vacatures op internet te vinden, te koppelen aan KvK-nummer of te ontdebelen. Een wijziging in de methodiek kan leiden tot een wijziging in de cijfers.
- Aansluitend bij het vorige punt, is het zo dat bedrijven en instellingen, die een (internet)vacature niet kunnen vervullen, deze soms weghalen en op een later tijdstip opnieuw plaatsen, vaak met een iets gewijzigde tekst. De vraag is dan of en wanneer er sprake van een nieuwe vacature?
- Op internet gaat het duidelijk om externe vacatures. Bij de vacaturestatistieken worden, naast externe vacatures, ook interne vacatures meegenomen. Dit maakt een vergelijking tussen internetwaarneming en de vacaturestatistieken minder gemakkelijk. Er is sprake van definitieverschillen.

- Bij de Kwartaalstatistiek Vacatures gaat het om openstaande vacatures aan het einde van het kwartaal. Openstaande vacatures worden bij internetwaarneming vastgesteld op basis van het feit dat ontstane vacatures nog niet van internet verdwenen zijn. Vacatures worden als ontstaan gekwalificeerd op het moment dat zij op internet zijn gevonden. Dus niet op het moment dat zij geplaatst zijn. Het onderscheid tussen vervallen en vervulde vacatures kan via internetwaarneming niet worden vastgesteld. Daar kan alleen worden gezien of een vacature niet meer op internet staat, dus verdwenen is.

Op dit moment is er onvoldoende zicht op de omvang en richting van en dus de invloed van deze aspecten op de ontwikkeling van het aantal daadwerkelijke vacatures. Er zijn (nog) geen mogelijkheden om deze invloeden als parameters mee te nemen bij de analyse, bijvoorbeeld in een later stadium bij de ophoging of weging naar representatieve aantallen.

Figuur 2. Kader vacatures en internetwaarneming.



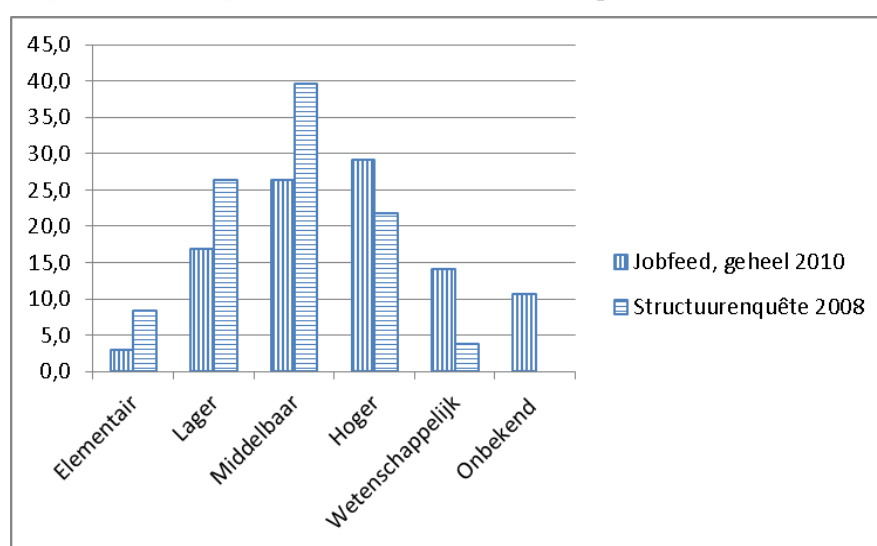
4 Belangrijkste resultaten onderzoek 2011 en 2012

Vooraf zij opgemerkt dat bij de analyses de internetwaarneming (Jobfeed) vergeleken is met de Kwartaalenquête Vacatures en in beperkte mate met de Structuurenquête 2008. Deze CBS-statistieken zijn als “standaard” gebruikt. De pretentie is hier echter niet om deze statistieken als “foutloos” te bestempelen. Ook bij de traditionele wijze van waarnemen kunnen zich problemen voordoen, zoals non-response, kwaliteit van invullen, onduidelijkheid of het gaat om externe én interne vacatures e.d. Ook de tijdigheid is een issue. Wel is het zo dat de doelpopulatie goed is afgebakend, op een verantwoorde wijze een steekproef is getrokken en daarmee de marges van gepresenteerde cijfers gekwantificeerd kunnen worden. Intentie is ook niet om een uitspraak te doen welke benadering, internetwaarneming of surveyonderzoek, beter is.

4.1 Eerste fase: onderzoek 2011

Het CBS heeft in 2011 de beschikking gekregen over drie leveringen uit Jobfeed, de database met internetvacatures van Textkernel. Twee via het UWV en één rechtstreeks van Textkernel. Dit betrof een beperkt aantal kwartalen en bedrijfstakken (4^{de} kwartaal 2009 en het jaar 2010). Alleen bij de laatste levering was een koppeling, via het Kamer van Koophandelnummer, aan de CBS-waarneming mogelijk.

Grafiek 1. Percentuele verdeling van vacatures op basis van typen beroepen in de data van Jobfeed (internet, geheel 2010) en de Structuurenquête CBS (2008) (Heerschap et al, 2012)



De belangrijkste statistische bevindingen van de analyses op deze bestanden waren (zie Heerschap e.a., 2012):

- Dat in vergelijking met de Structuurenquête eind 2008, de vacatures aan de bovenkant van de arbeidsmarkt zijn oververtegenwoordigd en vacatures aan de onderkant van de arbeidsmarkt zijn ondervertegenwoordigd in de Jobfeed data van 2010. Zie grafiek 1. Er zijn grote verschillen tussen bedrijfstakken: voor sommige bedrijfstakken zijn de verschillen duidelijk groter dan voor andere. Aangetekend zij wel dat de tijdstippen, eind 2008 en 2010, waartussen vergeleken is, redelijk ver uit elkaar liggen. Echter ook het onderzoek van Ruis en Grijpstra (2013) geeft aan dat vacatures waarbij een lagere opleiding

wordt gevraagd minder snel op internet komen dan vacatures waarbij een middelbare of hogere opleiding wordt gevraagd. Daarbij gaat het om dekkingsgraden van respectievelijk 85% en 91%.

- De ontwikkelingen en de niveaus van de vacaturetotalen van Jobfeed 2010 op macroniveau sluiten niet aan bij de Kwartaalstatistiek 2010. Dit betekent dat de uitkomsten van de reguliere Kwartaalstatistiek niet kunnen worden “voorspeld” door bijvoorbeeld extrapolatie van de uitkomst van één waargenomen Kwartaalstatistiek met vervolgens de kwartaalontwikkelingen in de Jobfeed gegevens.
- Als het gaat om het aantal vacatures op microniveau, dus per bedrijfseenheid, zijn de verschillen tussen de internetwaarneming en de CBS-waarneming meestal aanzienlijk. Hierbij kan voor een deel de wijze van koppelen tussen de KvK-nummers van Jobfeed en Bedrijfseenheden van het CBS, de statistische eenheid van de vacaturestatistieken, een rol gespeeld hebben. Een andere verklaring zou kunnen zijn dat er sprake is van definitiever verschillen: bij de Kwartaalstatistiek gaat het om externe én interne vacatures en bij de internetdata alleen om externe vacatures.
- Hoewel de uitkomsten van het CBS en Jobfeed op onderdelen fors verschillen, zijn er wel overeenkomsten bij bedrijfseenheden waar geen vacatures zijn ontstaan. Bedrijfseenheden waar geen vacatures zijn ontstaan volgens de Kwartaalstatistiek 2010, hebben meestal ook geen ontstane vacatures volgens Jobfeed 2010, en vice versa. Op grond van de Jobfeed data kan dus redelijk goed voorspeld worden welke bedrijfseenheden geen vacatures hebben in de Kwartaalstatistiek.
- Hierbij aansluitend, kunnen de kwartaalontwikkelingen voor de Kwartaalstatistiek dan ook goed worden benaderd door de kwartaalontwikkelingen van de groep van bedrijfseenheden, die op grond van Jobfeed één of meerdere vacatures hebben, te zetten op de primaire waarneming. Een kansrijke strategie is dan ook om één van de vier kwartalen de volledige Kwartaalstatistiek uit te sturen en om in de drie overige kwartalen alleen vragenlijsten uit te sturen aan de bedrijfseenheden die volgens Jobfeed minimaal één vacature hebben. De ontwikkeling van het aantal vacatures van de bedrijven met minimaal één Jobfeed vacature wordt dan gebruikt als proxy voor de ontwikkeling van het totaal aantal vacatures. Deze strategie zou een forse vermindering van de enquêtedruk betekenen, aangezien slechts een deel van alle bedrijfseenheden één of meerdere Jobfeed vacatures heeft. De strategie zou echter tot vertekende uitkomsten, met name bij verdere detaillering, kunnen leiden en vereist daarom nader onderzoek.
- Zoals in hoofdstuk 3 al is aangegeven, bevat de waarneming van vacatures via internet en de structurering daarvan (nog) “hiaten” en ruis, die op dit moment moeilijk te kwantificeren zijn.

Hiermee was eind 2011 de conclusie dat de gestelde doelstellingen, d.w.z. de vervanging van de waarneming bij de Structuurenquête Vacatures en een beperking van de waarneming bij de Kwartaalstatistiek Vacatures door internetwaarneming, niet goed haalbaar waren, althans met de resultaten van de tot dan toe uitgevoerde analyses. Voor verder onderzoek was er bij de onderzoekers sterke behoefte om over een uitgebreidere dataset over een langere periode te beschikken. Daarbij lagen de beste kansen bij het op microniveau koppelen van de Jobfeed data aan de CBS-waarneming van de Kwartaalstatistiek vacatures.

4.2 Tweede fase: onderzoek 2012

Om meer gedegen onderzoek te kunnen doen naar de bruikbaarheid van internetwaarneming voor de vacaturestatistieken is begin 2012 dan ook een dataset van Textkernel aangekocht met

alle op het Nederlandse deel van internet waargenomen vacatures van 2010 en 2011. Dubbele postings van vacatures waren door Textkernel vooraf al uit de dataset verwijderd. In het onderzoek zijn de vacatures die via uitzendbureaus zijn geplaatst niet gebruikt. Doordat de data per vacature ook het KvK-nummer bevatte zijn de internetdata, via het ABR⁴, gekoppeld aan de CBS-waarneming van de Kwartaalstatistiek op basis van Bedrijfseenheid. Dit was op zich een lastige exercitie.

Om het onderzoek beter te structureren zijn vooraf een achttal onderzoekscenario's gedefinieerd, die zich vooral hebben gericht op het beperken van de uitvraag bij de Kwartaalstatistiek Vacatures (zie Daalmans en Smeets, 2012). Daarbij gaat het om:

- 1) Het direct vervangen van de primaire waarneming van de Kwartaalstatistiek Vacatures door de waarneming van Jobfeed (scenario's 1 en 2). In deze scenario's wordt bij de Kwartaalstatistiek Vacatures nog maar één kwartaal primair waargenomen. De andere drie kwartalen worden samengesteld door extrapolatie op basis van de kwartaalontwikkelingen van de Jobfeed-gegevens.
- 2) Het niet uitzenden van vragenlijsten naar (bepaalde groepen van) bedrijven en instellingen in de steekproef van de Kwartaalstatistiek Vacatures die bij Jobfeed geen vacatures hebben (scenario's 3 - 5). Het schatten van het aantal vacatures e.d. gaat dan op basis van bedrijven en instellingen met één of meer vacatures in de Jobfeed data.
- 3) Het gebruik van Jobfeed data als hulpinformatie bij het schatten van het aantal vacatures met behulp van andere methoden dan de huidige directe Horvitz-Thompson schatter (scenario's 6-8).

Met als uiteindelijke hoofdvraag of het mogelijk is om minder "vragenlijsten" uit te sturen voor de Kwartaalstatistiek Vacatures dan nu het geval is met primaire waarneming, zonder dat dit al te zeer ten koste gaat van de kwaliteit.

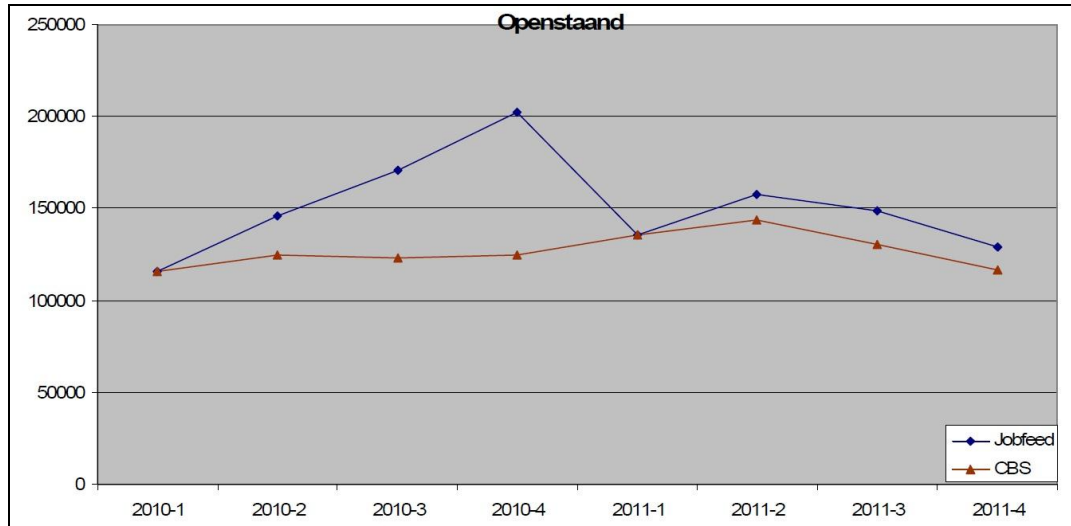
De belangrijkste resultaten van dit onderzoek zijn (zie Smeets en Daalmans, 2012):

- Dat het nog steeds niet goed mogelijk blijkt om op basis van primaire waarneming bij het eerste kwartaal de ontwikkeling van vacatures van de overige drie kwartalen te baseren op ontwikkelingen in de Jobfeed data (scenario's 1 en 2). Het aantal vacatures op internet stijgt sneller (of daalt minder snel) dan het aantal vacatures volgens de Kwartaalstatistiek in dezelfde periode. Daarbij zij wel aangetekend dat de vergelijking tussen de extrapolatie op basis van Jobfeed data en de Kwartaalstatistiek voor 2011 aanzienlijk beter waren dan die voor 2010. Mogelijk komt dat deels doordat in het begin van 2010 het aantal openstaande vacatures niet nauwkeurig kan worden vastgesteld omdat daarvoor onvoldoende gegevens beschikbaar waren (waren niet in de levering meegenomen). Een andere oorzaak kan zijn dat het gebruik van internet als wervingskanaal verder is toegenomen tussen 2010 en 2011, maar vooral dat Textkernel zijn methoden heeft verbeterd. Hiermee kan de invloed van geconstateerde "hiaten" en ruis zijn afgenomen.

Een belangrijk nadeel van scenario's 1 en 2 is dat er sprake kan zijn van een trendbreuk tussen kwartaal 4 en kwartaal 1 van het volgende jaar wanneer er weer sprake is van de traditionele primaire waarneming. Die trendbreuk is des te groter naarmate de ontwikkeling van het aantal vacatures per kwartaal op internet niet in de pas loopt met de jaar op jaar ontwikkeling via de primaire waarneming. Zie grafiek 2 als voorbeeld.

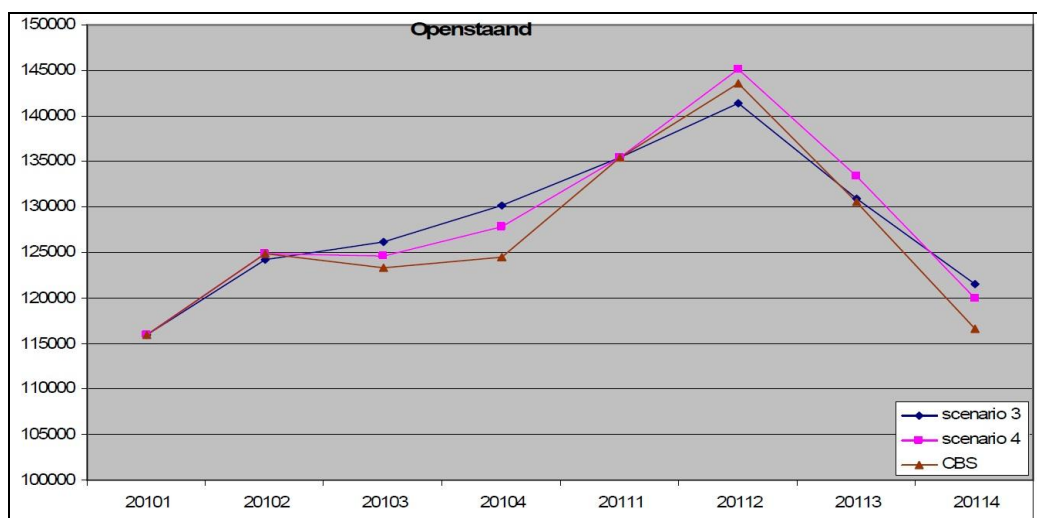
⁴ Het Algemene Bedrijvenregister van het CBS.

Grafiek 2. Aantallen openstaande vacatures gebaseerd primaire waarneming (CBS) en aantallen openstaande vacatures gebaseerd op extrapolatie op basis van ontwikkelingen in de Jobfeed data (Jobfeed), startend bij de een primaire waarneming van het eerste kwartaal (Smeets en Daalmans, 2012).



- Ook in deze tweede fase van het onderzoek wordt bevestigd dat de scenario's (3 en 4), waarbij geen vragenlijsten worden uitgestuurd naar (bepaalde groepen van) bedrijven, die geen vacatures hebben in de Jobfeed data, op dit moment de beste mogelijkheid biedt om internetwaarneming in te zetten. De totale besparingen aan "vragenlijsten" ligt tussen de 14% en 45%. Wel is het zo dat de resultaten vooral positief zijn voor het totaal aantal openstaande vacatures e.d. op macroniveau, zie grafiek 3 als voorbeeld, maar minder voor meer gedetailleerde cijfers, zoals een uitsplitsing naar grootteklasse en SBI.

Grafiek 3. Aantallen opstaande vacatures gebaseerd op bedrijven met één of meer vacatures in de Jobfeed data (scenario's 3 en 4) en Kwartaalstatistiek (CBS) (Smeets en Daalmans, 2012).



- Een potentiële mogelijkheid is om Jobfeed data te gebruiken als hulpinformatie bij andere schattingsmethoden dan de huidige directe Horvitz-Thompson schatter. Het gebruik van twee van de drie onderzochte alternatieve methoden (scenario's 6-8) leveren mogelijkheden op voor beperking van de uitvraag bij de Kwartaalstatistiek, namelijk 4% en 6% bij respectievelijk de quotiëntschatting en de GREG. In vergelijking met scenario's 3 en 4, is de beperking van de uitvraag relatief klein. Ook gaat het hier om complexe methoden.
- Ook in deze tweede fase van het onderzoek bleef de behoefte bij de onderzoekers groot om meer informatie te hebben over de bestaande "hiaten" en ruis in de internetwaarneming en gehanteerde verwerkingsmethoden. Anders is het moeilijk om onderscheid te kunnen maken tussen allerlei niet relevante invloeden op de ontwikkeling van internetvacatures en de werkelijke ontwikkeling van het aantal openstaande vacatures e.d. Bij scenario's 3 en 4 speelt dit minder, omdat daar die versturende invloeden het minste effect hebben.

Tabel 2. Overzicht van onderzochte scenario's.

	Scenario's							
	1	2	3	4	5	6	7	8
Potentiële lastendrukvermindering bij maximale afwijking van 5% met de Kwartaalstatistiek	10%	0%	45%	14%	0%	4%	6%	0%
Complexiteit van methode	-/+	-/+	-/+	-/+	-/+	++	++	++
Specifieke aspecten	Grote kans op trendbreuk bij de overgang kw 4 naar kw 1 van het volgende jaar	Er zijn grote verschillen op microniveau tussen de cijfers van de Kwartaalstatistiek en Jobfeed data	Minder goed toepasbaar op meer gedetailleerde data	Minder goed toepasbaar op meer gedetailleerde data	Minder goed toepasbaar op meer gedetailleerde data, maar nauwkeuriger als scenario 3	Hier gaat het om de Quotiënt schatter	Hier gaat het om de GREG	Hier gaat het om kleine domeinschatters
Toepasbaar	nee	nee	ja	ja	nee	mogelijk	mogelijk	nee

Scenario 1: Jobfeed gegevens worden "direct" gebruikt ter vervanging van de primaire waarneming van de Kwartaalstatistiek. Kwartaal 1 wordt primair waargenomen en de overige drie kwartalen worden, op macroniveau, geschat door extrapolatie op basis van ontwikkelingen in de Jobfeed data.

Scenario 2: Zie scenario 1, maar dan toegepast op individuele bedrijfseenheden (microniveau).

Scenario 3: Jobfeed gegevens worden indirect gebruikt om te bepalen welke bedrijfseenheden wel of geen vragenlijst krijgen. Bedrijven krijgen alleen een vragenlijst als zij in de Jobfeed gegevens minimaal één opstaande, ontstane of verdwenen vacature hebben.

Scenario 4: Zie scenario 3, maar dan alleen toegepast op grootteklasse 4-7.

Scenario 5: Zie scenario 3, alleen het bepalen of een bedrijf wel of geen vragenlijst krijgt wordt gebaseerd op het wel of niet hebben van vacatures e.d. in eerste kwartaal van de Kwartaalstatistiek. Er wordt dus geen gebruik gemaakt van Jobfeed gegevens.

Scenario 6: Jobfeed data worden gebruikt als hulpinformatie bij het gebruik van de quotiëntschatting voor de cijfers van de Kwartaalstatistiek.

Scenario 7: Zie scenario 6, maar in dit geval wordt de GREG gebruikt als schatter.

Scenario 8: Zie scenario 6, maar in dit geval wordt gebruik gemaakt van kleine domeinschatters.

Koppeling van Jobfeed data met CBS waarneming

Voorafgaande aan de analyses zijn de data van Jobfeed van 2010 en 2011 gekoppeld aan de waarneming van de Kwartaalstatistiek. Op basis van het ABR zijn de data van Jobfeed via KvK-nummers gekoppeld aan Bedrijfseenheden, die de statistische eenheid vormen van de Kwartaalstatistiek. Het koppelp proces was een lastige en tijdrovende klus. De oorzaak daarvan ligt vooral in het feit dat Bedrijfseenheden met bijbehorende KvK-nummers in de loop van de

tijd kunnen ontstaan, splitsen, fuseren en overliden, waardoor de koppeling via bijbehorende KvK-nummers minder eenduidig is. Het is van belang dat het KvK-nummer met de juiste datum aan een Bedrijfseenheid wordt gekoppeld. Daarnaast is het zo dat Jobfeed met behulp van algoritmes vanuit de tekst van de internetvacature, met onder andere naam en adres, een KvK-nummer afleidt. Dit hoeft niet altijd goed te gaan. Textkernel heeft de wens uitgesproken om dit proces te verbeteren. De daarvoor benodigde uitwisseling met CBS-data levert echter juridische problemen (geheimhouding) op.

Om dit koppelproces in de toekomst zo veel mogelijk te vereenvoudigen wordt voorgesteld om eerst bij het CBS de steekproef te trekken. Op basis van de getrokken Bedrijfseenheden via het ABR de bijbehorende KvK-nummers af te leiden en deze vervolgens aan Jobfeed aan te bieden. Zij kunnen dan aangeven of betreffende KvK-nummers wel of geen vacatures hebben. Ten behoeve van het gaafmaakproces bij het CBS zou dit ook kunnen worden uitgebreid met het aantal openstaande vacatures per KvK-nummer.

Juridische aspecten

Juridische aspecten spelen hier een ondergeschikte rol. Internetvacatures betreffen in principe geen persoonsgegevens. Alleen als er sprake is van eenmanszaken (die vacatures plaatsen), is er sprake van persoonsgegevens, zoals naam, adres e.d.

Wel is het van belang dat bij een eventueel proces van uitwisseling van gegevens tussen Textkernel en het CBS (KvK-nummers) dit juridisch goed moet worden afgedekt. Het leveren van KvK-nummers door het CBS aan Textkernel levert Textkernel in principe geen extra inzicht op over de bedrijven en instellingen. Zij bezitten zelf ook de betreffende KvK-nummers. Het enige inzicht dat zij krijgen is welke bedrijven het CBS in de steekproef van de Kwartaalstatistiek heeft zitten.

Gaat het CBS zelf spideren dan is het gewenst dat men zich houdt aan de afspraken zoals die gemaakt zijn bij het spideren van woningen en prijzen (o.a. openheid en herkenbaarheid van CBS-robots).

4.3 Eindconclusies

De mogelijkheden om internetwaarneming in te zetten ter vervanging of ter ondersteuning van de waarneming bij de Kwartaalstatistiek Vacatures en/of de Structuurenquête Vacatures van het CBS zijn vooral nog beperkter gebleken dan bij de start van het onderzoek was gedacht.

Daarbij spelen een aantal factoren een rol. Hier worden de belangrijkste genoemd. Hoewel de dekkingsgraad sinds 2010 sterk lijkt te zijn toegenomen, worden, ten eerste, niet alle vacatures op internet geplaatst. Dit kan per bedrijfstak verschillen. Er lijkt bovendien een bias (hoe groot is onduidelijk) te zitten in de vacatures die wel op internet worden geplaatst. Vacatures, waarbij een hogere opleiding wordt gevraagd, lijken oververtegenwoordigd en vacatures, waarbij een lagere opleiding wordt gevraagd, lijken ondervertegenwoordigd. Ook dit kan verschillen per bedrijfstak. Een tweede factor zijn de vacatures die op internet bij uitzendbureaus worden waargenomen. Door de vaak beperkte informatie over deze vacatures kunnen deze niet altijd aan de juiste bedrijfstak en/of regio worden toebedeeld, maar ook het ontdebelen van deze groep vacatures is minder gemakkelijk. Het gaat hier om forse aantallen vacatures. Bij de analyses van de gegevens zijn deze vacatures buiten beschouwing gelaten. Hoewel waarschijnlijk minder groot van omvang, speelt, ten derde, het fenomeen bulkvacatures een rol, waarbij meer dan één vacature wordt gepresenteerd in dezelfde internetvacature. En, ten slotte, spelen bij internetwaarneming allerlei invloeden (ruis), die geheel losstaan van de daadwerkelijke ontwikkeling van het aantal vacatures. Te noemen zijn het toenemend gebruik van internet als wervingska-

naal, strategieën van bedrijven of job boards om vacatures langer op internet te laten staan dan nodig, opkomst van nieuwe en verschuiving tussen bestaande wervingskanalen, zoals sociale media en gratis job boards, en het verbeteren van het spideren en de wijze van ontdubbeling bij Jobfeed (methoden).

Deze “hiaten” en ruis in de waarneming maken het op dit moment moeilijk in te schatten hoe representatief de aantallen internetvacatures en de ontwikkeling daarvan zijn voor de daadwerkelijke aantallen en ontwikkeling van vacatures. Dit zal ook verschillen per bedrijfstak. Er zijn onvoldoende gegevens beschikbaar om de omvang en richting van deze factoren goed te kwantificeren. Dat is echter wel nodig om internetwaarneming op een verantwoorde wijze bij de vacaturestatistieken van het CBS te kunnen inzetten.

Opgemerkt moet worden dat de meting van en de ontwikkeling in het vacatureaanbod in 2011 via data van Jobfeed beter aansluiten bij de cijfers van de Kwartaalstatistiek Vacatures dan in 2010 het geval was. Een belangrijke factor daarvoor kan zijn dat de invloed van de genoemde “hiaten” en ruis de afgelopen jaren is afgenomen door onder meer een hogere dekkingsgraad van de internetvacatures en een verbetering van de methoden om onder meer internetvacatures te spideren en te ontdubbelen. De statistische vraag is dan of deze betere fit tussen internetwaarneming en de Kwartaalenquête Vacatures in 2011 toeval is of niet?

In 2012 zijn ook acht scenario's onderzocht om met behulp van internetinformatie de uitvraag voor de Kwartaalenquête Vacatures te beperken. De beste mogelijkheden voor een reductie van de uitvraag lijken te liggen bij het zogenaamde nul-scenario, waarbij geen vragenlijsten worden uitgezonden naar die bedrijven die in de Jobfeed data geen vacatures hebben. Er is namelijk een relatie geconstateerd tussen bedrijven die bij de Kwartaalenquête Vacatures opgeven geen vacatures te hebben en waar Jobfeed ook geen vacatures op internet waarneemt, en vice versa. Afhankelijk van de variant die wordt gekozen kan dat oplopen tot een reductie van het aantal vragenlijsten van 14% tot 45%, respectievelijk bij scenario 4 en scenario 3. Bij scenario 3 krijgen alle bedrijven in de steekproef die 0 jobfeed vacatures hebben geen vragen over vacatures. Bij scenario 4 wordt dit alleen gedaan bij bedrijven in de grootteklasse 4-7. Met dit laatste scenario lijkt de kwaliteit van de Kwartaalstatistiek beter gewaarborgd.

Voor de scenario's 3 en 4 wordt in deze notitie een business case gepresenteerd (zie hoofdstuk 5). De geschatte eenmalige kosten zijn € 10.000 en de structurele kosten op jaar basis € 22.000 voor met name het verkrijgen van de data van Jobfeed. Daartegenover staan als baten: een beperking van de uitvraag, mogelijk een betere kwaliteit van de statistiek en vooral kansen om het imago van het CBS te verbeteren. Er wordt vanuit gegaan dat er één jaar wordt schaduwgedraaid.

Een nadeel blijft het feit dat de waarneming van de Kwartaalstatistiek is gekoppeld aan die van de statistiek Ziekteverzuim. De enquêtedrukvermindering uit zich dan ook alleen in minder vragen en niet in minder formulieren. Als de waarneming van de Kwartaalstatistiek wordt losgekoppeld van de waarneming van de statistiek Ziekteverzuim, heeft dit alleen zin als ook voor de statistiek Ziekteverzuim een alternatieve databron kan worden gevonden. Anders leidt het waarschijnlijk juist tot een gepercipieerde lastendrukverzwaring.

Ook scenario's 6 en 7, waarbij alternatieve schattingsmethoden worden gebruikt, kunnen leiden tot een beperking van de uitvraag bij de Kwartaalstatistiek Vacatures. In vergelijking met het nul-scenario, gaat het hier om complexe methoden en is de vermindering van de uitvraag relatief klein (4-6%).

Hoewel de resultaten van het onderzoek minder positief zijn uitpakend dan gehoopt, is het niet zo dat daarmee internetwaarneming van vacatures niet meer bij het CBS in beeld is. Integen-

deel, er is echter meer onderzoek nodig. Om goede uitspraken te kunnen doen is een langere reeks data nodig, bijvoorbeeld van Jobfeed data van 2012 en 2013. Dat maakt de tot nu behaalde resultaten beter generaliseerbaar voor andere perioden dan de nu onderzochte jaren.

Naast de vervanging van de waarneming bij de Kwartaalenquête Vacatures en de Structuur-enquête Vacatures, kan internetdata ook worden ingezet voor onder meer: hulpinformatie in het verwerkingsproces, het produceren van snelle bèta-indicatoren en het detailleren van de cijfers, bijvoorbeeld naar regio⁵ of andere variabelen (zie Structurenquête). Verder kan ook nog gedacht worden aan nieuwe informatie, bijvoorbeeld door het koppelen met andere statistieken (o.a. Polis-administratie).

Zelf waarnemen op internet

Bij de start van het onderzoek eind 2010 is de gedachte geweest om zelf te gaan spideren bij de top 100-200 bedrijven. Gezien de resultaten van de analyses op de Jobfeed data tot nu toe en de eventuele kosten voor het verkrijgen van de data van derden, komt dit alternatief weer in beeld. Men dient zich wel te realiseren, dat het niet alleen gaat om het waarnemen op internet, maar vooral ook om het proces daar achter, lopend van data extractie en het ontdebellen tot het standaardiseren van de gegevens en het koppelen aan Kamer van Koophandelnummers. Juist de expertise van Textkernel. Ook het zelf waarnemen kost (veel) geld, zeker als het gaat om het ontwikkelen van kennis en het opdoen van ervaring.

Zet men beide alternatieven naast elkaar dan ontstaat het volgende beeld, waarbij opgemerkt moet worden dat de geschatte eenmalige en structurele jaarlijkse kosten bij de eigen internet-waarneming aan de lage kant zijn. Vooral het traject na de internetwaarneming zal resources vereisen. Dit kan mogelijk relatief laag worden gehouden als puur wordt gefocust op de websites van de top 100-200 grootste bedrijven.

Men dient zich ook te realiseren, dat het domein van waarnemen op internet bij vacatures dynamisch is, doordat bedrijven en instellingen in de tijd splitsen, fuseren, ontstaan en overlijden, met alle gevolgen van dien voor de bijbehorende KvK-nummers en de waar te nemen websites. Dat ligt geheel anders dan bij het spideren van bijvoorbeeld woningen.

Beperkte eigen internetwaarneming	Gebruik van Jobfeed data (scenario's 3 en 4)
Kennis nog te ontwikkelen (lange doorlooptijd)	Kennis direct al beschikbaar
Kennis in te zetten op andere terreinen	Kennis blijft extern
Bepalen van en inzicht in (eigen) methoden	Beperkte invloed op de gehanteerde methoden
Onafhankelijk van een derde partij	Afhankelijk van een derde partij
Geschatte eenmalige kosten > € 200.000 (2 vte., aanpassen en ontwikkelen van systemen)	Eenmalige kosten € 10.000 (zie hoofdstuk 5)
Geschatte structurele kosten > € 40.000 (0,2 vte)	Structurele kosten € 22.000 (zie hoofdstuk 5). Worden de internetdata ook gebruikt voor publicatie dan lopen de kosten op naar € 100.000

De in de tabel genoemde structurele kosten betreffen het opleveren van de ruwe data. Dus het omvat het bijhouden van het te spideren internetdomein (welke websites van welke KvK's), het bouwen en onderhouden van de robots, het extraheren van de gewenste data, (eventueel) het

⁵ Bijvoorbeeld via iteratief proportioneel fitten (ipf) op basis van internetdata of banen/werkzame personen. Op vergelijkbare wijze kan dat ook voor opleiding en beroep (o.a. Polis-administratie).

ontdubbelen, het classificeren en standaardiseren en het eventueel afleiden van Kamer van Koophandelnummers, inclusief SBI. En niet het analyseren van de data, het ophogen en wegen en het maken van de output.

In zekere mate betreft het een afweging tussen het opdoen en bezitten van kennis over het waarnemen van vacatures met internetrobots tegenover de duidelijk hogere kosten van de eigen waarneming ten opzichte van het gebruik van Jobfeed data. Daarbij moet rekening worden gehouden met de voortdurende (technologische) ontwikkeling van internet en het gebruik van dit kanaal als wervingsbron. Je bent niet in één keer klaar, maar je moet voortdurend bijblijven. Wel is het zo dat met eigen waarneming geïnvesteerd wordt in kennis en ervaring, die ook op andere terreinen kunnen worden ingezet, bijvoorbeeld als het gaat om textmining, standaardiseren en classificeren.

5 Business case nul-scenario

Het scenario dat op dit moment het beste haalbaar lijkt, is het scenario waarbij naar de in de steekproef getrokken bedrijven en instellingen die bij de Jobfeed geen vacature hebben geen “vragenlijst” te sturen. Of anders gezegd, de schatting van het aantal vacatures e.d. vindt plaats op basis van in de steekproef getrokken bedrijven die één of meer vacatures hebben in de Jobfeed data. Scenario 4, waarbij dit alleen wordt gedaan voor grootteklassen 4-7⁶, waarborgt de kwaliteit van de uitkomsten het beste. Maar er kan ook worden gekozen voor scenario 3.

Het proces van het nul-scenario (onderzochte scenario's 3 en 4) ziet er als volgt uit:

1. Per kwartaal wordt een steekproef getrokken (standaard).
2. Op basis van het ABR worden de bijbehorende KvK-nummers afgeleid (nieuw). Van belang is dat ook de periode van de geldigheid van de koppeling tussen Bedrijfseenheid en KvK-nummer wordt meegegeven.
3. De set met KvK-nummers wordt aan Textkernel geleverd. Deze koppelen de gewenste gegevens met unieke vacatures en bijbehorende data aan de KvK-nummers en zenden deze data aan het CBS retour (nieuw). Van belang is dat dit tijdig gebeurt om de gewenste einddatum van publicatie te halen.
4. Vervolgens worden de gegevens van de KvK-nummers geaggregeerd naar Bedrijfseenheden (nieuw).
5. Bedrijfseenheden die volgens Jobfeed geen vacatures hebben krijgen een vragenlijst over ziekteverzuim, maar zonder vacaturevragen (nieuw: aanpassen van het systeem, twee sets enquêteformulieren).
6. Bij het rappelleren moet er onderscheid blijven tussen bedrijven met en zonder vragen over vacatures (standaard, maar ook aanpassen brieven, aanpassen systeem).
7. Controleren en verwerken van de wel uitgezonden vragenlijsten (standaard). Mogelijk kan hier Jobfeed data gebruikt worden als controle data. Dat zou betekenen dat van alle aan Textkernel gestuurde KvK-nummers vacature aantallen moeten worden verkregen en niet alleen het vlaggetje “wel of geen vacatures”.
8. Tenslotte, het ophogen van gaafgemaakte data naar gewenste output (standaard).

Tabel 3 geeft een overzicht van de te maken eenmalige en structurele kosten voor dit scenario. Doordat de Kwartaalstatistiek gekoppeld is aan de statistiek Ziekteverzuim is, zoals eerder gezegd, hier geen sprake van besparingen van uit te zenden enquêteformulieren, brieven, postzegels⁷ e.d. Dit blijft gelijk. Er is dus in alle gevallen sprake van een negatief financiële business case. De te maken kosten komen bovenop de kosten voor de huidige Kwartaalstatistiek/Ziekteverzuim. Vandaar elke keer het bedrag van € 0 in de kolom bij de “huidige waarne- ming”.

Tegenover de hiergenoemde kosten staan de volgende baten:

⁶ De gedachte is dat het toepassen van deze methode bij de hele kleine bedrijven en instellingen (grootteklasse 1-3) zal leiden tot een vertekend beeld van deze groep.

⁷ Een groot deel komt ook elektronisch binnen en er wordt gebeld.

- Lastendrukvermindering bij de Kwartaalstatistiek Vacatures als het gaat om het aantal in te vullen vragen door bedrijven en instellingen die geen vacatures hebben in de Jobfeed data. Afhankelijk van het gekozen scenario tussen 14-45%.
- Kansen om het imago van het CBS te verbeteren, als het gaat om het beeld rond de inspanningen die het CBS verricht om de enquêtedruk te verminderen.
- Een betere kwaliteit van de Kwartaalstatistiek als Jobfeed data ook worden gebruikt als controle bij het gaafmaakproces.

Tabel 3. Eenmalige en structurele kosten voor scenario 's 4 of 3.

	Huidige waarneming	Nul-scenario	Opmerkingen
Eenmalig			
1. Aanpassing automatiseringssysteem CBS	0	€ 5.000	Probleem is dat er geen goed inzicht is bij Dataverzameling in het systeem an sich. Kosten zijn inclusief testen.
2. Aanpassing systeem Jobfeed	n.v.t.	€ 3.000	
3. Aanpassen brieven/formulieren	n.v.t.	€ 2.000	
Totale eenmalige kosten	0	€ 10.000	
Structureel			
4. Aanschaf data op jaarbasis	0	€ 20.000	Het betreft hier de extra data van Jobfeed.
5. Afleiden en aggregeren van KvK's naar Bedrijfseenheden	0	€ 2.000	
Totale structurele kosten per jaar	0	€ 22.000	

Uiteindelijk is het aan het management om een beslissing te nemen om dit scenario wel of niet uit te voeren. Als het wel het geval is dan zal zeker voor een jaar schaduw moeten worden gedraaid.

Een optie is, om naast Jobfeed, te kijken naar alternatieve aanbieders van data over internetvacatures. Ook zal men mogelijk te maken krijgen met een aanbestedingstraject. Tot slot lijkt het verstandig in deze samen op te trekken met het UWV.

Zoals al eerder aangeven is het feit dat de uitvraag van de Kwartaalstatistiek Vacatures gekoppeld is aan de uitvraag van de statistiek Ziekteverzuim een bottleneck om de voordelen van het nul-scenario ten volle te benutten. Hierdoor uit de beperking in de lastendruk zich alleen in minder vragen en niet in minder enquêteformulieren.

6 Aanbevelingen voor een vervolg

Met de resultaten van dit onderzoek is de bijdrage van het Programma Impact ICT en daarmee Prima afgerond. Samen met de twee CBS-interne onderzoeknotities (Daalmans en Smeets, 2012 en Smeets en Daalmans, 2012), vormt deze notitie de afrondende rapportage van dit onderzoek aan het ministerie van Economische Zaken (en Prima).

Verder onderzoek of implementatie van scenario's zal plaatsvinden op basis van eigen CBS-middelen. In principe is daarvoor in 2013 geld gereserveerd in de beleidsreserve van het CBS.

Het onderzoek van de afgelopen twee jaar heeft nog onvoldoende beeld gegeven om goed te kunnen inschatten of internetwaarneming op een verantwoorde wijze kan worden ingezet bij de vacaturestatistieken van het CBS. Het ontbreekt vooral aan voldoende gegevens over de omvang en richting van geconstateerde "hiaten" en ruis in de waarneming via internet. Gebleken is dat dit soort onderzoek een lange adem vereist. Internetdata als aanvulling op of als ondersteuning van bestaande waarneming lijkt in eerste instantie meer voor de hand te liggen. Wel is het natuurlijk goed mogelijk om (snelle) bèta-indicatoren of andere statistieken te maken en publiceren over de vacaturemarkt, die puur gebaseerd zijn op internetwaarneming, maar waarvan men weet dat deze minder betrouwbaar zijn.

Voor de korte termijn worden de volgende vervolgstappen voorgesteld:

1. *De aanschaf van Jobfeed data voor minstens nog één jaar (2012 en mogelijk 2013).* Hiermee kan worden beoordeeld of de betere fit tussen de ontwikkeling in vacatures van de internetdata en de Kwartaalenquête Vacatures van 2011 zich handhaaft. Zoals aangegeven, lijkt de dekkingsgraad van internetvacatures de laatste jaren sterk te zijn toegenomen. Bovendien zijn de methoden van Jobfeed om vacatures op internet waar te nemen en te verwerken verbeterd. Hiermee kan het zijn dat de invloed van de "hiaten" en ruis in de internetwaarneming kleiner zijn geworden dan uit dit onderzoek naar voren is gekomen. Dit vraagt om nieuw overleg met Textkernel. Het is daarbij ook van belang te kijken naar de koppeling via het KvK-nummer. Dat heeft in 2012 veel werk gekost.
2. *Onderzoek om meer inzicht te krijgen op de "hiaten" in en de ruis op de internetwaarneming.* Lukt het niet om hierop beter grip te krijgen dan blijft het moeilijk internetwaarneming, buiten het nul-scenario om, in te zetten voor het maken of ondersteunen van CBS-waardige vacaturestatistieken. Een start zou kunnen zijn om een klein diepteonderzoek uit te voeren naar de verschillen op microniveau tussen het aantal vacatures bij bedrijven en instellingen op internet en het aantal vacatures dat zij opgegeven bij de Kwartaalstatistiek. Waarom zijn die verschillen vaak zo groot en in hoeverre speelt de waarneming bij de Kwartaalstatistiek of de koppeling op basis van KvK-nummers daarbij een rol? Er zijn ongetwijfeld bedrijven en instellingen die hieraan willen meewerken.
3. *Aansluitend op het vorige punt, het benutten van extra informatie.* Bijvoorbeeld meer recente gegevens over het percentage vacatures dat bedrijven en instellingen op internet zetten (zie wel Ruis en Grijpstra, 2013). Er kan ook gedacht worden om bij één van de uitzendingen van de Kwartaalstatistiek meer (hulp)informatie te vragen, die gebruikt kan worden bij het samenstellen van de andere drie kwartalen op basis van Jobfeed data.
4. *Eén jaar schaduwdraaien met het nul-scenario, en dan met name variant 4 (grootteklasse 4-7).* Het lijkt verstandig de beslissing over deze stap pas te nemen als de stap bij het eerste pijltje, dat wil zeggen: nog een jaar Jobfeed data, is afgerond.

5. *Verder onderzoek naar andere mogelijkheden om internetwaarneming in te zetten bij de vacaturestatistieken van het CBS.* Voorbeelden zijn: als hulpinformatie bij het verwerkingsproces, voor detaillering van de cijfers (o.a. de Structuurenquête Vacatures) en als snelle bèta-indicatoren als aanvulling op de huidige output.
6. Bij de vacaturestatistieken is men gestart met *onderzoek naar de behoefte aan cijfers over de vacaturemarkt.* Op basis van die behoefte kan vervolgens gekeken worden naar de bronnen, die nodig zijn om die cijfers te maken. Daarbij kan internetwaarneming een rol spelen. Onderdeel daarvan zou ook moeten zijn om na te gaan in hoeverre gebruikers van de vacaturestatistieken geïnteresseerd zijn in snellere en meer gedetailleerde cijfers, zogenaamde bèta-indicatoren, over de vacaturemarkt, maar dan puur gebaseerd op internetwaarneming (vergelijk de huidige vacature indicator van NR)? Ook al weet men dat de representativiteit te wensen overlaat.
Hierbij behoort ook de vraag in hoeverre het CBS het als haar taak ziet om dit soort bèta-indicatoren te produceren? Is het antwoord op deze vragen ja en gaat men verder met Jobfeed data, dan dient men zich te realiseren dat aan zo'n variant een aanzienlijk hoger kostenplaatje hangt. Dat zal voor de Jobfeed data liggen rond de € 100.000 per jaar (exclusief BTW). Mogelijke alternatieven zijn: het hierin gezamenlijk optreden met het UWV en om te kijken naar andere aanbieders van internetdata over vacatures. Bij het laatste alternatief speelt ook de kwaliteit en continuïteit van die aanbieders een rol.
7. En, ten slotte, kan onderzocht worden of het zinvol is om *zelf internetwaarneming te verrichten*, bijvoorbeeld bij de websites van de top 100-200 bedrijven en instellingen. De verwachting is echter dat de kosten van deze optie structureel aanzienlijk hoger zullen uitpakken dan het verkrijgen van internetdata van Textkernel. Voordelen zijn wel: het verkrijgen van kennis en ervaring op dit terrein (w.o. bijvoorbeeld textmining), het kunnen bepalen van de gehanteerde methoden en geen afhankelijkheid van een derde partij.

Voorgesteld wordt om evenals vorig jaar te starten met een brainstormsessie met deskundigen om het onderzoek te structureren. Dat heeft in 2012 goed uitgepakt. Bij de brainstormsessie van 2012 zijn alleen CBS'ers uitgenodigd. Voorstel is om dit nu uit te breiden met externe deskundigen. Niet om behoeften te peilen aan cijfers, maar om te kijken naar oplossingen voor de methodologische problematiek. Te denken valt aan deskundigen op de terreinen van de arbeidsmarkt, spidering, methodologie / econometrie, big data en het ABR.

Ten slotte, zij nog opgemerkt dat de keuze voor vervolgonderzoek bepaald wordt door de doelen die worden nagestreefd. Is het doel alleen lastendrukvermindering? Wil men ook kijken naar snellere of meer gedetailleerde cijfers aansluitend op de behoeften. Of "beperken" we ons tot de informatiebehoefte van Eurostat? En als laatste is kwaliteitsverbetering van de vacaturestatistieken een issue?

7 Referenties

- Barendswaard, P. (2009), Internetrobots en wetgeving (Databankenwet, CBS-wet), 2-10-2009, interne nota CBS.
- Binnenlands Bestuur (2012), Nieuwe portal maakt CBS cijfers makkelijker vindbaar, 20-01-2012.
- Boerman, P. (2013), De tragische teloorgang van Monsterboard, in Management Team. 10 januari 2013.
- Daalmans, J. (2012), Steekproefadvies 2012 vacature- en ziekteverzuimstatistiek, interne nota CBS
- Daalmans, J en M. Smeets (2012), Inzet van internetwaarneming bij de vacaturestatistieken; Onderzoekscenario's Jobfeed (PPM-2012-04-04-JDAS-MSET), 4 april 2012, interne nota CBS.
- Drees, M. (2012A), Job board landschap verandert snel, kansen voor recruiters, column op het Internet, 23 januari 2012.
- Drees, M. (2012B), De opkomst en ondergang van Monsterboard, column op het Internet, 14 januari 2012.
- Dialogic (2008), Go with the data flow, 13 mei 2008.
- Grijpstra, D.H., G. Mulder en M. Folkeringa (2010), Eindverslag pilot Vacaturemeter, september 2010.
- Heerschap N. (2010A), Inzet van internetrobots bij de Vacature-enquête, Project Initiatie Document (klein), Versie 1.1, interne nota CBS.
- Heerschap, N. (2010B), Programma Impact ICT 2011 en 2012.
- Hoefkens R., (2012), Vermindering Lastendruk vacatures en ziekteverzuim, Project Initiatie document, versie 1.0p2, interne nota CBS
- Hoekstra, R., O. ten Bosch en F. Hartevelde (2010), Automated data collection from web sources for official statistics: first experience, 25 april 2010.
- Intermediair (2011), Vacaturesites de belangrijkste bron voor werkzoekenden, 30-08-2011
- Ruis, A. en D. Grijpstra (2013), *Vacatureonderzoek Kenniscentra*, Panteia, april 2013
- Ruth Van, F.J. en R. Wekker (2009), *Monthly employers' sentiment indicators; doing more with business survey data*, Centraal Bureau voor de Statistiek discussionpaper 09009. Url: <http://www.cbs.nl/NR/rdonlyres/1459ED14-E741-4384-94B4-1252AA4B11F7/0/200909x10pub.pdf>

- Ruth Van, F.J. en R. Wekker (2011), *Nieuwe Vacature-indicatoren; Doel, Achtergrond en Samenstellingen*, Rapport, Centraal Bureau voor de Statistiek. Url: <http://www.cbs.nl/NR/rdonlyres/3A3DB463-D4E1-4BC2-A279-F3189C402DF5/0/2011nieuwevacatureindicatoren.pdf>
- Smeets, M. en J. Daalmans (2012), Vermindering lastendruk vacaturestatistiek: innovatieve waarneming op het Internet (PPM-2012-06-21-MSET-JDAS), 29 mei 2012, interne nota CBS.
- Ten Bosch, O. en D. Windmeijer, Eindrapport internet robots woningmarkt over 2011 en 2012, CBS-rapport, 2013 (nog te publiceren).
- Velzen M. van, J.A. Daalmans, N. Heerschap, D. ter Steege (2011), Statistieken van de vacatures op internet: een vooronderzoek, Interne nota DMV-2011-07-28-MVLN-JDAS-NHRP-DSGE CBS Den Haag, , interne nota CBS.