



Inhoud

Inhoud	3
.....	
Hoofdstuk 1	
Inleiding	5
1.1 Databeheer: een leemte in de basisopdracht voor de panelenquête	5
1.2 Taakstelling van de databeheerder	6
Hoofdstuk 2	
Verslag werkbezoek IAB, Nürnberg	9
2.1 Vragenlijsten	9
2.2 Dataverzameling	10
2.3 Documentatie en archivering	11
2.4 Vormen van samenwerking	11
2.5 Regulering en documentering van samenwerkingen, dataprotectie	12
2.6 Monitoring van samenwerkingen	13
2.7 Analyses	13
2.7.1 Panelkarakter	13
2.7.2 Imputatie	13
2.7.3 Regionale analyses	14
2.7.4 Cross-culturele analyses	14
2.7.5 Linken met andere databestanden	14
2.8 Publicatie- en disseminatiebeleid, respondentenbinding	15
2.9 Maatschappelijk draagvlak	15
Hoofdstuk 3	
Bestandsbeveiliging in Nederland: OSA, WSA en CBS	16
3.1 Het OSA-arbeidsvraagpaneel	16
3.2 WSA: makelaar in databestanden	18
3.2.1 Rol van het WSA	18
3.2.2 Beveiliging van data: algemeen	19
3.2.3 Beveiliging van bedrijfsdatabanken	20

3.3 On desk en on site werken op CBS-data	22
3.4 Methodieken van bestandsbeveiliging	26
Conclusie.....	28

Hoofdstuk 4

On site werken op PASO-databestanden: databeveiliging via een stand-alone computer.....	29
Conclusies.....	30

Hoofdstuk 5

Monitoringsysteem PASO-samenwerkingsverbanden.....	32
---	-----------

Hoofdstuk 6

Overzicht van de PASO-samenwerkingsverbanden.....	37
6.1 Beknopt overzicht van de PASO-samenwerkingsverbanden	37

Hoofdstuk 7

PASO-samenwerkingsprotocol en bestandsovereenkomsten.....	41
7.1 Protocol PASO-samenwerkingsverbanden	41
7.1.1 Filosofie	41
7.1.2 Samenwerkingsvormen	41
7.2 De bestandsovereenkomst samenwerking PASO - overheid en wetenschappelijke instellingen	44
Bestandsovereenkomst type A (analyses door het PASO-team, rapportering door de gebruiker)	44
Bestandsovereenkomst type B (on site werken op de PASO-bestanden)	48
7.3 Bijlage bij de bestandsovereenkomst	
Model van geheimhoudingsverklaring PASO	53
Verklaring	53

Hoofdstuk 8

Studiedag 'Het Panel Survey of Organisations in Flanders in een breder perspectief' Seminarie donderdag 4 maart 2004 - Hotel Errera, Brussel.....	54
8.1 Inleiding	54
8.2 Doelstellingen van het seminarie	55
8.3 Doelpubliek	56
8.4 Datum en plaats	56
8.5 Agenda	56



Hoofdstuk 1

Inleiding

1.1 | Databeheer: een leemte in de basisopdracht voor de panelenquête

De PASO-enquête genereert als panelsurvey een omvangrijk databestand. Omdat het om een panelenquête gaat neemt dit bestand proportioneel toe in omvang en in complexiteit met elke volgende golf. Eén van de doelstellingen van PASO is overbevraging vermijden. Een van de strategieën daartoe is het databestand zo ruim mogelijk ter beschikking te stellen van overheid en wetenschap om het analysepotentieel ervan maximaal te valoriseren en zo een doelmatige aanwending van de overheidssubsidie voor het project te verzekeren. In het basiscontract van het PASO-onderzoek is deze doelstelling weliswaar expliciet vermeld, er zijn echter geen voorzieningen getroffen om deze doelstelling ook waar te maken. Inmiddels zijn er tijdens de eerste golf van de survey, die plaatsvond in 2002, verschillende scenario's uitgewerkt én ten dele in praktijk gebracht om deze samenwerking in functie van reductie van overbevraging vorm te geven: het gebruik van modules met additionele en eenmalige vragen is daar één van. Daarnaast waren er ook vragen voor additionele analyseopdrachten, met of zonder rapportering door het PASO-team, en vragen van derden om zelf op het bestand te werken.

Omdat de enquête microgegevens bevat over vestigingen zijn bijzondere voorzorgen noodzakelijk om zowel de anonimiteit van de respondenten als de correctheid en objectiviteit van de wetenschappelijke analysemethoden bij gebruik door derden te verzekeren. Het K.U.Leuven-promotorenteam moest daarom alle technische, organisatorische en juridische schikkingen treffen om maximale toegankelijkheid van de data voor overheid en wetenschap te combineren met een noodzakelijke en adequate 'zorg' voor het bestand en de respondenten. Voorstellen van richtlijnen daarvoor werden uitgewerkt in het PASO-protocol en dit is opgevolgd door een ad hoc werkgroep, samengesteld uit PASO en de Vlaamse administratie (EAW en APS). Naast bestandsovereenkomsten en protocollen die deze samenwerkingsvormen ondersteunen (zoals vertrouwelijkheidsafspraken edm.), was het echter ook noodzakelijk een 'fysieke' omgeving te organiseren waarin een correct en bruikbaar databestand met documentatie ter beschikking wordt gesteld van overheid en wetenschap.

De basisfinanciering van het PASO-project liet niet toe deze bijkomende taken in functie van de toegankelijkheid van de data op te nemen. Met andere woorden, voor de doelstelling 'overbevraging vermijden', geconcretiseerd in verschillende samenwerkingsverbanden, waaronder het degelijk ter beschikking stellen van de data voor niet-commerciële derden, waren noch impliciete noch expliciete voorzieningen getroffen. De concretisering van deze opdracht en de draagwijdte ervan kon vooraf ook moeilijk worden ingeschat. Een kader of infrastructuur, kortom mensen of middelen, die het PASO-team daarin konden ondersteunen waren niet aanwezig. Bovendien bleek het, reeds van bij de eerste golf, noodzakelijk de omstandigheden te creëren en de basisvoorwaarden te verzekeren voor een optimale toegankelijkheid van het bestand. De voorgestelde 'basisinvestering' in de uitbouw van een degelijk databeheer werd vanuit die optiek zinvol bevonden. Het project databeheer, ondersteund binnen het Vlaamse VIONA-programma, waarover hier verslag wordt gebracht, maakte dit toch mogelijk.

1.2 | Taakstelling van de databeheerder

In die bijkomende opdracht voor het organiseren van het databeheer voor VIONA werden verschillende taken opgenomen, waarover in dit eindverslag wordt gerapporteerd. De taakstelling van het project luidde als volgt:

1. de kwaliteit en actualisering van het databestand verzekeren voor gebruik door derden. Dit houdt ook in het voldoende documenteren van het bestand in al zijn aspecten (steekproef, weging, itemnonrespons, variabelen, etc.);
2. marktonderzoek naar verschillende methoden van bestandsbeveiliging bij relevante bevragingen/panelenquêtes in Duitsland en Nederland (CBS, OSA, IAB, ...) en die evalueren op hun waarde voor PASO als vestigingsbestand;
3. de meest aangewezen beveiligingsmethoden uitwerken en toepassen;
4. alle andere maatregelen treffen voor het on site werken in een beveiligde omgeving. Dit impliceert het installeren van een stand-alone pc, uitgerust met de nodige programmatuur, die enkel wordt gebruikt voor PASO-analyses;
5. het onderzoeken en behandelen van aanvragen voor specifieke deelbestanden door derden, overheid en wetenschappers;
6. het klaarmaken en overmaken van deze deelbestanden, begeleiden van derden in het gebruik van de data, zowel on site als via andere wegen (e-mail, etc.);
7. de nodige screenings doen van de analyses en publicaties van derden op het respecteren van de anonimiteit van de respondenten;

8. de nodige screenings doen van de analyses en publicaties van derden op het correct gebruik van analysemethoden gegeven de specificiteit van de bestanden en feedback hierover organiseren naar deze derden;
9. een goed monitoringsysteem uitwerken om het gebruik van het bestand door derden te registreren, te evalueren en daarover te rapporteren;
10. voorzien in een permanente ervaringsuitwisseling met APS en de VIONA-administratie inzake deze problematiek;
11. de ervaringen terzake opgedaan met PASO rapporteren om ze ter beschikking te stellen van beide administraties;
12. de koppeling van de PASO-databankthema's met andere bestaande databanken onderzoeken in een specifieke studiedag.

In voorliggend rapport worden de ervaringen met deze verschillende aspecten van de functie 'databeheer' in het kader van het PASO-project gerapporteerd en geëvalueerd.

Dit rapport is geen gewoon onderzoeksrapport. Het brengt de verslagen samen van de verschillende werkzaamheden die zijn ondernomen om het werken op de PASO-data door derden uit te bouwen. Daardoor is het eerder een 'werkingsverslag'. Zo rapporteren de eerste twee hoofdstukken over hoe Nederlandse en Duitse collega's het gebruik van bedrijfsdata door derden organiseren en welke maatregelen zij daartoe nemen. Vervolgens wordt in de drie volgende hoofdstukken getoond hoe PASO dit heeft georganiseerd, hoe de samenwerkingsverbanden zijn 'gemonitord' en welke samenwerkingsverbanden er gedurende het werkingsjaar van het project databeheer zijn gerealiseerd. De laatste twee hoofdstukken omvatten de basisdocumenten die dienen om een samenwerking met derden te omkaderen: een zogenaamd protocol dat het algemene kader toelicht, de bestandsovereenkomst en de geheimhoudingsverklaring. Voor deze laatste documenten werkten de ervaringen van het Nederlandse WSA inspirerend. Tot slot moet worden aangehaald dat een essentieel onderdeel van dit databeheerproject erin bestond een omstandige en permanent geüpdate documentatie van het PASO-databestand ter beschikking te stellen van onderzoekers van zowel binnen als buiten het PASO-onderzoeksteam. Deze uitgebreide codeboeken vormen een volwaardig onderdeel van het databeheerproject maar zijn omwille van hun omvang uitsluitend op cd-rom verkrijgbaar.

We zijn ervan overtuigd dat de ervaringen die in het project zijn opgedaan relevant en nuttig zijn voor zowel de bredere onderzoeksgemeenschap als voor de overheden die het opmaken en ter beschikking stellen van statistieken tot hun opdracht rekenen. We hopen dan ook dat onze ervaringen kunnen bijdragen tot een meer open cultuur van kennisdeling tussen wetenschappers onderling en tussen overheden en wetenschappers. Immers, het is onze overtuiging dat in deze openheid, weliswaar omkaderd door duidelijke afspraken, veel meer mogelijkheden voor kennisdeling schuilen dan vandaag worden

benut: mogelijkheden voor een echte wetenschappelijke meerwaarde door een intensiever en dieper gebruik van beschikbare data, door het toepassen van bredere en aanvullende onderzoeksperspectieven en onderzoeksvragen, door het stimuleren van het kritisch denken over de data en door het bevorderen van de wetenschappelijke dialoog. We hopen dan ook dat onze ervaringen inspirerend mogen werken voor iedereen die beschikt over onderzoeksdata-banken en die niet bang is voor de toets der kritiek.

Monique Ramioul
Projectleider HIVA
PASO-promotor

Hoofdstuk 2

Verslag werkbezoek IAB, Nürnberg

Op 20 en 21 juli 2004 bracht de databeheerder van de PASO-equipe een werkbezoek aan het Institut für Arbeitsmarkt- und Berufsforschung.¹ Dit instituut maakt deel uit van het Bundesagentur für Arbeit, het Duitse federale ministerie van arbeid. De sector 'Bedrijf en werkgelegenheid' (Forschungsbereich 'Betriebe und Beschäftigung') van het IAB coördineert sinds 1993 een groot-schalig bedrijvenpanel (Betriebspanel). Bij de uitwerking van PASO was dit panel een belangrijke inspiratiebron. Het bezoek had tot doel de ervaringen opgedaan tijdens ruim drie jaar PASO-werking te confronteren met 'het Duitse verhaal'. Voor achtergronden bij het IAB-panel verwijzen we naar de notities bij de presentatie die Markus Promberger van het IAB op 4 maart 2004 gaf in het kader van het PASO-seminarie (zie infra).

2.1 | Vragenlijsten

- IAB hanteert het principe 'één golf, één vragenlijst'. Er wordt dus niet zoals in PASO een kortere vragenlijst voorzien voor kleinere vestigingen. Ook zijn er geen verschillen in vraagverwoording tussen industriële en dienstenorganisaties, een onderscheid dat in golf 1 en golf 2 aanleiding gaf tot verschillende versies. Evenmin krijgen quartaire organisaties een ingekorte vragenlijst voorgelegd, zoals in PASO-golf 3 gebeurde. IAB stelt wel vast dat de antwoordkwaliteit wat lager ligt bij kleine vestigingen. Het werken met slechts één vragenlijst impliceert uiteraard een ernstige administratieve en technische vereenvoudiging tegenover het PASO-opzet (= min of meer op maat gesneden enquêtes). Keerzijde is dat de vragen naar inhoud en in hun verwoording algemener moeten zijn. IAB schat dat ongeveer 80% van de gestelde vragen relevant én beantwoordbaar zijn voor alle vestigingen, dus ook de allerkleinste.
- Over de inhoud van de volgende vragenlijst wordt intensief vergaderd. De procedure/routine omhelst onder andere twee conferenties, subconferenties en een reeks IAB-interne vergaderingen. De opdrachtgevers² kunnen vragen voorstellen, bijvoorbeeld om beleidsmaatregelen te evalueren of om

¹ Zie www.iab.de. De informatie is ook deels in het Engels beschikbaar.

attitudes ten opzichte van actuele onderwerpen te meten. Externe onderzoeksgroepen en andere derden hebben in principe geen inspraak in de vragenlijstconstructie. IAB behoudt de eindverantwoordelijkheid over de inhoud van de vragenlijst en over de uiteindelijke vraagverwoording.

- Een onderscheid tussen A-, B- en C-vragen, zoals initieel voorzien in het PASO-concept, wordt niet expliciet gemaakt. De vragenlijst bestaat voor ongeveer de helft uit vragen die jaarlijks worden herhaald. Er zijn ook vraagcycli van 2 en 4 jaar.

2.2 | Dataverzameling

- De IAB-data worden verzameld via face to face interviews. PASO is een webenquête, aangevuld met schriftelijke vragenlijsten (postenquête) bij kleine vestigingen. Uiteraard heeft dit verschil in methode verregaande implicaties voor de unitrespons (in IAB zowel voor panelvestigingen als voor nieuwe respondenten minstens dubbel zo hoog als in PASO), op de itemrespons (geen cijfers voorhanden, maar wellicht ook beter in IAB) en op de kostprijs per respondent (enkele malen hoger in IAB).
- Infratest, een onafhankelijk bureau dat gespecialiseerd is in dataverzameling, neemt het veldwerk voor de IAB-enquête integraal voor haar rekening. Een netwerk van circa 1 000 interviewers wordt ingeschakeld om jaarlijks circa 15 000 organisaties te bevragen. Infratest staat in voor de training en briefing van de interviewers, voor de coördinatie en opvolging van het veldwerk, voor het ingeven en coderen van de data, voor kwaliteitscontrole en datacleaning. Achteraf voert IAB wel extra controles uit op de data en wordt er nog bijkomende cleaning gedaan. Infratest wordt omwille van haar grote methodologische expertise ook actief betrokken in de vragenlijstconstructie. De auteursrechten op de vragenlijsten worden gedeeld.
- Gedurende enkele jaren heeft men ook een klein deel van de steekproef bevestigd via postenquêtes. Dit leverde een respons op van ongeveer 35%. Deze overschakeling was evenwel niet opgezet als experiment of vanuit financiële overwegingen, maar was louter te wijten aan het feit dat Infratest een tijdlang in bepaalde regio's met een tekort aan interviewers kampte.
- Een webbevraging wordt niet overwogen. Men vreest vooral de lage respons bij kleinere vestigingen waar webtoegang en *webskills* niet altijd even vanzelfsprekend zijn. Zolang de nodige middelen voorhanden zijn om een ruime steekproef te realiseren middels een kwalitatief hoogstaande methode van dataverzameling, ziet IAB ook geen reden om uit te kijken naar alternatieven.

² De enquête wordt voor de helft gefinancierd uit eigen middelen (= middelen van het Federale ministerie van Arbeid) en voor de helft door de regeringen van de zestien Bundesländer, die extra steekproefeenheden financieren om regionale analyses mogelijk te maken (infra).

2.3 | Documentatie en archivering

- Per golf wordt een soort van codeboek bijgehouden. Deze documenten zijn echter heel wat minimalistischer van opzet dan de PASO-documentatie. De IAB-codeboeken bevatten minder meta-informatie. Over de cleaning bijvoorbeeld wordt geen informatie gegeven. Dit laatste wordt als een manco ervaren.
- Alle bestanden worden gearchiveerd op een centrale server. Toegang tot de server is beperkt tot het IAB zelf. De databestanden mogen het IAB-gebouw niet verlaten. Medewerkers kunnen van thuis uit niet inloggen op de server.

2.4 | Vormen van samenwerking

- Op het niveau van de *vragenlijst* zijn er geen samenwerkingen mogelijk met IAB. Externen kunnen geen modules aan de bevraging koppelen, zoals in de eerste twee PASO-golven gebeurde.
- Samenwerking op het niveau van de *steekproef* en de *dataverzameling* is voor IAB essentieel. De lokale besturen (Bundesländer) financieren immers een deel van het veldwerk met het oog op het realiseren van regionale deelsteekproeven.
- Op het niveau van de *data* en *analyse* zijn er twee samenwerkingsvormen. Niet-commerciële instellingen konden tot voor kort een vraag stellen aan een helpdesk/hotline. Indien mogelijk, hielp de persoon die de helpdesk bemande de vrager verder met een cijfer, een tabel of een kort rapportje. Deze helpdesk is momenteel afgeschaft, het 'on demand' aanleveren van informatie zal worden opgenomen door de nieuw opgerichte datawarehouse-afdeling binnen IAB (infra). De tweede optie is 'on site' werken; dit is het uitvoeren van analyses op een pc in het IAB-gebouw. Deze optie staat niet open voor iedereen; de opportuniteit van de samenwerking wordt eerst onderzocht. Externen die men niet persoonlijk kent, hebben een aanbeveling nodig (bv. van een promotor). Het werken op de data is gratis. Sommige onderzoekers worden door het IAB uitgenodigd om op de data te werken; hiervoor zijn beurzen (research grants) voorzien. *Remote access*, het werken op de data op een andere locatie dan het IAB-gebouw, is onder geen beding mogelijk, zelfs niet voor IAB-medewerkers. Dit is bij PASO wel mogelijk, zij het onder strikte voorwaarden. Wel kan een externe onderzoeker aan IAB een controlebestand vragen zodat hij of zij in een statistisch pakket (SPSS, SAS, ...) de syntax/code van de gewenste analyses kan uittesten. Een IAB-medewerker laat vervolgens het programma op het eigenlijke bestand lopen en levert de analyse-output aan de externe onderzoeker. In het testbestand zitten de variabelen en een reeks observaties met op toevals-

basis aangemaakte waarden. Jaarlijks krijgt het IAB ongeveer 2 000 aanvragen voor 'samenwerking'. Het betreft hier evenwel meestal concrete vragen voor het opzoeken van één of enkele cijfers. Naar schatting werken er elk jaar een vijftigtal externen op de IAB-data.

2.5 | Regulering en documentering van samenwerkingen, dataprotectie

- Het formaliseren, stroomlijnen en documenteren van samenwerkingsverbanden is een taak van de recent opgestarte datawarehouse-afdeling (Forschungsdatenzentrum). Op dit moment moeten externen een formulier invullen met basisinformatie zoals de naam van de organisatie waarvoor ze werken, informatie over het project en een timing. In het formulier staat ook expliciet dat gebruikers de federale wetgeving op het vlak van dataprotectie moeten naleven. Deze wetgeving stipuleert onder andere dat men op datasets die verzameld zijn door een overheidsinstelling enkel geaggregeerde analyses mag doen, geen analyses op het niveau van individuele cases. Voor het waarborgen van de anonimiteit en een wetenschappelijke benadering van de data valt men dus terug op de wettelijke bepalingen terzake.
- Uit de databestanden die voor *on site* analyses ter beschikking worden gesteld, worden de identificatievariabelen manueel verwijderd. Celfrequenties lager dan vijf verschijnen niet op het scherm. IAB past geen postrandomization³ toe en gebruikt geen gespecialiseerde software om de identificeerbaarheid van de data verder te minimaliseren. Dit wordt ook niet echt nodig geacht, aangezien men enkel personen op de data laat werken die men vertrouwt. Bovendien zijn deze personen, zoals gesteld, gehouden aan de federale wetgeving inzake databescherming.
- IAB behoudt het eerste recht van gebruik van de data en zal het goedkeuren van aanvragen indien nodig uitstellen om voortijdige of gelijktijdige publicaties door derden over eenzelfde thema te vermijden.
- In het algemeen valt op dat IAB er voor opteert de samenwerkingsvoorwaarden en de aanvraagprocedure niet al te zeer te formaliseren. Men bekijkt aanvragen ad hoc en onderhandelt indien nodig over de modaliteiten van de samenwerking. In plaats van voor een strakke, expliciete regulering kiest men voor overleg. Men beseft dat het panel een dure aangelegenheid is en dat een goede toegankelijkheid van de data belangrijk is voor het voortbestaan van de enquête. IAB hanteert de filosofie dat iedereen met een goed voorstel en de nodige credits op de data moet kunnen werken.

³ Het opnieuw aanmaken van de data met dezelfde basiskenmerken. Bij dit systeem veranderen de gemiddelden niet, maar wel de variantie en de significantie.

2.6 | Monitoring van samenwerkingen

- In het jaarverslag wordt een beknopt overzicht gegeven van de samenwerkingsverbanden. Er gebeurt geen meer gedetailleerde documentering of archivering (bv. systematisch bijhouden van publicaties van externen in bibliotheek).
- Het grote aantal samenwerkingen maakt een strikte opvolging van de werkzaamheden van externe gebruikers en een uitgebreide kwaliteitscontrole van analyseresultaten en rapporten in de praktijk onmogelijk. Soms gebeurt het wel dat gebruikers zelf om een controle vragen. Wanneer resultaten of interpretaties gepubliceerd worden waar de IAB-equipe het niet mee eens is, zal men wel overwegen om hierop publiekelijk te reageren, bijvoorbeeld middels een rechtzetting of door een artikel in een volgende editie van het tijdschrift.
- Er is geen handleiding of FAQ uitgewerkt om het werken op de data te faciliteren. Nieuwe gebruikers leren al doende en kunnen met vragen bij de vaste IAB-medewerkers terecht. Men geeft toe dat dit niet de meest efficiënte werkwijze is, noch voor de gebruikers, noch voor het IAB-personeel.

2.7 | Analyses

2.7.1 | Panelkarakter

- IAB geeft Infratest de opdracht in de mate van het mogelijke dezelfde interviewer dezelfde persoon in het panelbedrijf te laten interviewen. Op die manier wordt de vergelijkbaarheid tussen de golven en de betrouwbaarheid van longitudinale analyses bevorderd. Wanneer men in de eerste plaats geïnteresseerd is in cross-sectionele analyses, zo geeft men op IAB toe, is het verstandiger om de respondentkeuze (bv. bedrijfsleider of personeelsdirecteur) af te stemmen op het onderwerp van de bevraging, zoals in PASO gebeurt.
- Panelanalyses die het IAB uitvoerde, gingen onder andere over werkgelegenheid (cf. PASO-dossier over dynamiek op de arbeidsmarkt, waarin de twee eerste golven met elkaar in verband worden gebracht), arbeidstijdregelingen, productiviteit en collectieve arbeidsverhoudingen.

2.7.2 | Imputatie

IAB imputeert regelmatig data uit een vorige golf. Stel dat een organisatie in 2002 een reeks vragen invulde over bijvoorbeeld arbeidstijden en dat dit topic in 2003 niet aan bod kwam, dan worden de variabelen over arbeidstijden anno 2002 toegevoegd aan het bestand van 2003. Imputatie wordt enkel uitgevoerd voor panelbedrijven, uit de meest recente golf en voor

onderwerpen waarvan men weet dat ze op korte termijn doorgaans weinig veranderen. Soms wordt ter controle de vraag gesteld of er op het onderwerp in kwestie belangrijke veranderingen plaatsvonden in vergelijking met een vorig bevragingsmoment. Bij een negatief antwoord is men relatief zeker dat de data geïmputeerd mogen worden.

2.7.3 | Regionale analyses

Zoals eerder reeds vermeld, garandeert de financiering van de Bundesländer dat men representatieve uitspraken kan doen over elk van de deelstaten afzonderlijk. Daalt men nog een niveau, naar steden of stadsregio's, dan is de respons meestal onvoldoende om veralgemeenbare uitspraken te doen. De Bundesländer geven aan plaatselijke universiteiten de opdracht de data te analyseren. Deze rapporten zijn hoofdzakelijk beleidsondersteunend van aard. Ook worden er soms interregionale vergelijkingen gemaakt, vooral dan tussen de vijf nieuwe Länder (= het voormalige Oost-Duitsland).

2.7.4 | Cross-culturele analyses

Er zijn reeds cross-culturele analyses gedaan met de IAB-data. Regelmatig komen buitenlandse onderzoekers aan het IAB werken op de Duitse data, doorgaans om een vergelijking te maken met het eigen land.

2.7.5 | Linken met andere databestanden

Het linken van datasets stuit in Duitsland op een reeks juridische bezwaren. Data mogen niet zonder meer gebruikt worden in combinatie met andere bestanden. De interessante koppeling van het IAB-bedrijvenpanel met de Duitse arbeidskrachtensurvey is bovendien ook om technische redenen moeilijk haalbaar: aan werknemers wordt weliswaar de naam gevraagd van de organisatie waar ze werken, maar deze informatie is in vele gevallen helemaal niet of slechts heel moeilijk te linken met de vestigingen (als juridische entiteiten) die in IAB bevraagd worden. Om de koppeling tussen een werknemers- en een werkgeversbevraging te maken, moet men vertrekken van een design waarin beide vanaf het begin geïncorporeerd worden. De mogelijkheid van de koppeling met een ander vestigingspanel van het IAB, over wervingsgedrag, wordt momenteel onderzocht. Aangezien er met hetzelfde universum gewerkt wordt, zou het mogelijk moeten zijn variabelen in de twee richtingen te imputeren.

2.8 | Publicatie- en disseminatiebeleid, respondentenbinding

- Het IAB en de sector die instaat voor het panel hebben eigen communicatiekanalen en een eigen publicatiebeleid. Het IAB als instituut rapporteert naar de opdrachtgevers (in de vorm van papers) en naar een breder publiek, onder andere via de organisatie van persbriefings. De sector legt meer de nadruk op wetenschappelijke publicaties, gericht op academici en de bedrijfswereld.
- Deelnemers aan de bevraging kregen vroeger een nieuwsbrief opgestuurd, momenteel wordt on line feedback gegeven over de resultaten. De belangrijkste bindingsstrategie of responsmaximalisatietechniek ziet IAB in het feit dat de brief die de respondenten uitnodigt tot medewerking mee wordt ondertekend door gezaghebbende figuren binnen de Duitse werkgeversorganisaties. Hun geloof in en steun aan de IAB-survey wordt als cruciaal ervaren voor de voortzetting van het panel. Met andere vormen van 'beloning', zoals benchmarking of het toekennen van bepaalde voordelen, heeft men geen ervaring.
- Het IAB organiseert tal van activiteiten waarop de onderzoeksresultaten van het instituut worden voorgesteld. De sector 'Bedrijf en werkgelegenheid' richt elk jaar in samenwerking met enkele universiteiten een workshop in waar over de vorige golf/golven wordt gerapporteerd. Deze workshop is echter uitsluitend bedoeld voor panelspecialisten en onderzoekers die actief zijn op het vlak van arbeidsmarkt en organisaties.
- Op de IAB-website worden ruim vijftig rapporten in Engelse vertaling aangeboden. Op die manier worden de resultaten internationaal ruimer verspreid.

2.9 | Maatschappelijk draagvlak

Het maatschappelijk draagvlak voor de IAB-enquête wordt als groot ervaren. Geloof in het nut van de enquête en actieve of passieve ondersteuning van de stakeholders wordt ook als noodzakelijk ingeschat om een dergelijk duur onderzoeksproject te continueren. De vakbonden tonen zich doorgaans geïnteresseerd in de resultaten van de bevraging. Ook met de werkgeversorganisaties is het contact goed. Af en toe wordt er wel eens druk uitgeoefend om een thema niet op te nemen, wanneer men vreest dat de resultaten een negatieve impact hebben op het imago van de bedrijfswereld (bv. het gebrek aan een beleid ten opzicht van oudere werknemers), maar veto's worden er niet gesteld. Ook hier weer geldt dat men via overleg naar een oplossing zoekt die voor alle partijen aanvaardbaar is.

Hoofdstuk 3

Bestandsbeveiliging in Nederland: OSA, WSA en CBS

3.1 | Het OSA-arbeidsvraagpanel

Het *arbeidsvraagpanel* van de Organisatie voor Strategisch Arbeidsmarktonderzoek (OSA, <http://www.uvt.nl/osa>) vindt sinds 1989 tweejaarlijks plaats. Het gaat hier zoals bij PASO om een bevraging van vestigingen. Wel worden enkel vestigingen vanaf vijf werknemers benaderd. Tot en met de zevende golf in 2001 gebeurde de dataverzameling via een mondeling interview, met achterlating van een schriftelijke vragenlijst. In 2003 is uit kostenoverwegingen besloten het face to face interview te vervangen door drie in de tijd gespreide telefonische interviews. Voor meer informatie over het

panel:

http://www.uvt.nl/osa/data/index2/arbeids_vraag_panel.html.

In het algemeen geldt dat OSA erg *terughoudend* is in het beschikbaar stellen van data uit het arbeidsvraagpanel. De samenwerkingsmogelijkheden zijn beperkt tot wetenschappelijk onderzoek en beleidsgericht onderzoek. De OSA-medewerkers beoordelen in eerste instantie zelf of de data geschikt zijn voor de onderzoeksvraag.

De onderzoekers moeten een *gebruikersverklaring* ondertekenen waarin zij onder andere de vertrouwelijke behandeling van de gegevens garanderen. Deze verklaring moet ook onderschreven worden door de directeur van het onderzoeksinstituut. Dit zijn de belangrijkste voorwaarden die in de gebruikersverklaring zijn opgenomen:

- de verstrekte gegevens mogen uitsluitend worden gebruikt ten behoeve van statistische analyse en verwerking met het oog op de uitvoering van bovengenoemd onderzoeksproject door bovengenoemde onderzoeker;
- in alle publicaties en presentaties waarin direct of indirect gebruik wordt gemaakt van gegevens in het OSA-arbeidsvraagpanel is bronvermelding verplicht;
- vanuit het oogpunt zicht te houden op disseminatie verplichten gebruikers zich aan OSA een exemplaar te doen toekomen van alle schriftelijk vastgelegde resultaten van onderzoek (boeken, rapporten, artikels, congresspapers) waarbij OSA-bestanden zijn gebruikt;

- de verstrekte gegevens mogen onder geen enkele voorwaarde worden doorgeleverd aan derden, noch aan derden ter inzage worden gegeven in ruwe of bewerkte vorm, met uitzondering van door OSA goedgekeurde presentatie of publicatie van onderzoeksresultaten;
- OSA is op geen enkele wijze aansprakelijk voor fouten of onjuistheden in de verstrekte gegevens, noch voor fouten of onvolkomenheden in de documentatie betreffende de gegevens;
- OSA is niet aansprakelijk voor eventuele fouten of onjuistheden in de verwerking door derden van OSA-bestanden;
- de gegevens zijn pas voor analyse beschikbaar na ontvangst van alle voor deze overeenkomst benodigde schriftelijke verklaringen.

Sinds enkele jaren is men afgestapt van het *on site werken*. Als redenen geeft men aan dat deze werkwijze 'nogal belastend was voor de gebruikers' en 'als middel om kwaadwillenden te stoppen niet waterdicht'. Voor gebruikers is het inderdaad gemakkelijker in het eigen kantoor en op een vertrouwde pc analyses uit te voeren dan op verplaatsing te werken. Over de (on)veiligheid van on site werken stelt Marcel Kerkhofs van OSA: "Het is bij mijn weten in het verleden bij OSA één keer voorgekomen dat de data tegen de toen geldende afspraken in zijn meegenomen (overigens zonder oneerbare bedoelingen, zuiver voor het gebruikersgemak). Waterdicht on site werken is op zichzelf te realiseren, maar dan moet je een goed geoutilleerde stand-alone pc klaar hebben staan en alle output controleren. Voor ons was dat praktisch niet goed te realiseren en lieten we mensen de facto slechts on site werken om een oogje in het zeil te kunnen houden." Derden kunnen nu dus 'remote', 'on desk' op de OSA-bestanden analyses uitvoeren.

Uit de *datasets* die OSA ter beschikking stelt aan derden, worden een aantal variabelen verwijderd. Andere kenmerken worden op een beperkt niveau van detaillering aangeleverd. Het betreft hier vooral regionale gegevens, de indeling in bedrijfssectoren, financiële gegevens en de bedrijfsgrootte. De beperking van het detailniveau gebeurt manueel. Gevraagd naar de precieze werkwijze, antwoordt Kerkhofs: "Dat is bij ons allemaal handwerk (en ook maatwerk omdat de beperkingen kunnen afhangen van de onderzoeksvraag). Regionale gegevens worden tot een indeling in vijf regio's of twaalf provincies beperkt. Voor de economische activiteit van bedrijven gebruiken wij een 5-digit code, waarvan de gebruikers doorgaans alleen het eerste digit krijgen. Bij de bedrijfsgrootte maken we doorgaans onderscheid in de categorieën 5-9, 10-19, 20-49, 50-99, 100-499 en 500 werknemers of meer." Van de onder andere door CBS ontwikkelde software (MU-ARGUS) en methodieken (postrandomization) wordt met andere woorden geen gebruik gemaakt.

Het OSA-arbeidsvraagpanel is opgenomen in de WSA-catalogus (infra), maar daar wordt voornamelijk doorverwezen naar OSA. Op termijn is het de

bedoeling dat de administratieve afhandeling van het datagebruik ook bij WSA kan plaatsvinden.

Tot nog toe heeft OSA de data enkel beschikbaar gesteld aan onderzoekers die men persoonlijk kent. Het *aantal samenwerkingen is eerder beperkt*. In de afgelopen jaren heeft men naar eigen zeggen gemiddeld vijf verzoeken per jaar gehonoreerd. Ongeveer de helft van de samenwerkingen vonden plaats in het kader van onderzoeksprojecten waarvan OSA zelf opdrachtgever was. In enkele gevallen werd slechts een sterk beperkte versie van de dataset aangeleverd.

3.2 | WSA: makelaar in databestanden

Op 7 juli 2003 bracht een delegatie van het PASO-team, de Vlaamse administratie, Administratie Planning en Statistiek en het Steunpunt Werkgelegenheid Arbeid en Vorming een werkbezoek aan het WSA.

3.2.1 | Rol van het WSA

Het Wetenschappelijk Statistisch Agentschap (WSA, http://www.nwo.nl/nwohome.nsf/pages/NWOP_5GBE4P) heeft als voornaamste doel het stimuleren van secundair gebruik van bestaande databestanden voor wetenschappelijk onderzoek. WSA vervult hierdoor een intermediaire, makelende rol tussen onderzoekers en de beheerders van de databestanden. In Nederland stellen steeds meer dataproducenten hun bestanden via WSA beschikbaar.

De hoofdp opdrachten van het agentschap zijn:

- de acquisitie van relevante bestanden;
- het informeren van gebruikers;
- de vorming van één loket voor het zoeken en bestellen van de data;
- het bevorderen van kennisoverdracht via de organisatie van gebruikersbijeenkomsten.

Voor andere taken, zoals de fysieke archivering en distributie van bestanden wordt samengewerkt met dataproducenten, databeheerders en data-archieven.⁴

WSA is wat men noemt een datamakelaar. WSA, bemand door slechts twee FTE, doet enkel acquisitie van gebruiksrechten op data, dat ze vervolgens via aangepaste contracten en verschillende modaliteiten doorgeven aan faculteiten en onderzoekers. Ze beheren zelf geen databanken. Ze leveren ook geen

⁴ In Nederland bestaan enkele instellingen die zich uitsluitend of hoofdzakelijk met archivering bezighouden. Eén van de belangrijkste archieven voor sociaal-wetenschappelijk onderzoek is het Steinmetzarchief, dat deel uitmaakt van het Nederlands Instituut voor Wetenschappelijke Informatiediensten (NIWI). Zie <http://www.niwi.nl/nl/maatschappijwetenschappen/steinmetzarchief/toon> voor meer informatie.

data. WSA bewaakt de toestemming en het gebruik. De producent zelf blijft verantwoordelijk voor archivering, documentering en distributie. Alle expertise blijft dus bij de onderzoeker/dataproductent zitten en dat is ook noodzakelijk.

De archivering/depotfunctie van data gebeurt door het Centraal Bureau voor de Statistiek (CBS) en (vooral voor niet-CBS-data uit onderzoek) in het Steinmetzarchief. Deze archivering is voorzien in de statistiekwet en nagenoeg een standaardformule in onderzoekscontracten. Er wordt momenteel zelfs op variabelenniveau gearchiveerd en gedocumenteerd. Er is een tendens naar meer actieve archivering. Er wordt momenteel gewerkt naar één loketfunctie tussen Steinmetzarchief, NWO en WSA zodat het voor de gebruiker-onderzoeker gemakkelijker wordt en de verschillende beschikbare databanken zichtbaarder worden.

Een richtcijfer voor de aankoop van het gebruiksrecht is 3% van de totale bevragingenkost.

Het WSA beschikt over een catalogus die on line raadpleegbaar is, via de Dataneth-toepassing (<http://www.dataneth.nl/wsa/search.do?action=search>). Op dit moment biedt Dataneth de mogelijkheid om databestanden te zoeken op de titel van een studie, de gebruikte afkorting of op trefwoord. Er wordt gewerkt aan nieuwe zoekmogelijkheden, zoals het zoeken op variabelenniveau. Van elke studie worden vervolgens de metadata gepresenteerd, zoals het design, een omschrijving van de steekproef, de methodologie en de gebruiksvoorwaarden. Beschikbare bijbehorende documentatie, zoals codeboeken, onderzoeksverantwoording en publicaties kan men downloaden. Databestanden kunnen via de toepassing en onder de geldende voorwaarden on line besteld worden bij het WSA.

De makelaarsfunctie impliceert dat er veel aandacht moet gaan naar een degelijke documentering van de databanken ten behoeve van de gebruikers. Dit kan volgens gestandaardiseerde formats, die door WSA ter beschikking worden gesteld aan de dataproductenten. Die blijven hiervoor verantwoordelijk. Door het feit dat WSA betaalt voor (gebruiksrecht op) de data, kunnen ze ook kwaliteits- en documenteringseisen stellen ten aanzien van de gebruikers. De formats die WSA daarvoor gebruikt zijn tweërlei: enerzijds kwalitatieve info voor een catalogus/website. Maar nu recent ook DDI, dat is een gestandaardiseerd pakket dat een soort data-dictionary is binnen SPSS. Dit archiveert en documenteert op variabelenniveau (eigenlijk de codeboeken), waardoor onderzoekers ook op dit niveau kunnen zoeken. Om dit systeem te promoten is er een internationale organisatie opgericht: het 'Data Liberation Army'.

3.2.2 | Beveiliging van data: algemeen

In het makelen van de data is er een verschil tussen scientific use files en public use files. Deze laatste zijn vaak zo geaggregeerd dat ze vaak voor

onderzoekers niet meer relevant zijn. Scientific use files zijn minder geaggregeerd. Men gaat ervan uit dat toevallige identificatie nooit uit te sluiten is. Daarom sluit men in principe akkoorden af met faculteiten in plaats van met onderzoekers, omdat die een rechtspersoonlijkheid hebben. Zo kunnen faculteiten een soort abonnement krijgen op de Enquête beroepsbevolking van het CBS (microdata!). Toch zegt het WSA dat ‘vertrouwen’ de basis vormt in de uitwisseling van data: vooral vertrouwen bij de producenten in verband met behoorlijk gebruik en het respect voor de privacy.

Om identificatie van individuen in een databestand toch maximaal te vermijden bestaat er zeer nuttige software, bijvoorbeeld Argus dat bestaat voor microdata en voor tabellen. Daarbij worden ad random driedimensionele tabellen gedraaid die dan kunnen worden gecontroleerd. Volgens aggregatie zijn er drie mogelijke uitkomsten: ‘heel herkenbaar’, ‘redelijk herkenbaar’ of ‘minimaal herkenbaar’.

Volgende beveiligingsmogelijkheden werden opgesomd:

- Argus-software (zie supra);
- postrandomisation (zie IAB);
- remote access;
- outputcontrole.

Algemeen stelt men vast dat onderzoekers correct omgaan met de data, vooral met de richtlijnen in verband met het niet voor iets anders gebruiken dan afgesproken, correct omgaan met toevallig geïdentificeerden, niet doorgeven van data aan derden. Ook het tekenen van een contract en een gedragscode werken sterk. Immers met publicaties op basis van gevoelige data kan men toch niet naar buiten komen, omdat men een contract heeft getekend waar serieuze boetes aan vastzitten. Wat onderzoekers vaak verwaarlozen is het opsturen van de eindrapporten en het vernietigen van de data drie jaar na het aanvragen.

Voor PASO zijn vooral de volgende methodieken relevant.

3.2.3 | Beveiliging van bedrijfsdatabanken

WSA is formeel; bedrijfsdata zijn niet te beveiligen. Er bestaat geen systeem waarbij bedrijven niet meer identificeerbaar zijn. Tweede principe is dat de producent van de data de enige is die ze mag bezitten én beheren: ofwel worden de data enkel in een ultrabeveiligde omgeving ter beschikking gesteld bij de producent ofwel krijgt iedereen dezelfde data en zijn ze dus publiek. Er zijn wel een aantal strategieën die men kan toepassen om de privacy te garanderen. Dit is en blijft immers essentieel: er werd sterk benadrukt dat één incident in verband met privacy volstaat om voor jaren medewerking

vanwege respondenten te hypothekeren. De verschillende principes zijn complementair te hanteren.

Een eerste principe dat WSA hanteert, richt zich daarom op *outputcontrole*, controle van het gebruik in plaats van op databeveiliging (wat immers op zich nauwelijks kan). Dat betekent dat elk onderzoeksvoorstel en -contract wordt opgevraagd bij degene die aanvraagt om op de data te kunnen werken. Daarin kan men al heel goed zien wat de onderzoeker van plan is. Opvragen van de onderzoeksvoorstellen werkt ook preventief en is tegelijk ook een handig monitoringsinstrument.

Verder worden alle publicaties ter controle en goedkeuring voorgelegd. Deze goedkeuring richt zich uiteraard niet op de politieke wenselijkheid van uitspraken, wel op kwaliteitscontrole. Die kwaliteitscontrole kan gaan over de statistische maten die men toepast, de opmaak van tabellen, maar ook over interpretatie (bv. conclusies die niet mogen getrokken worden omdat er bias is, omdat er intermediaire variabelen spelen die niet gecontroleerd zijn, omdat de causaliteit niet correct is edm.). De workload van dergelijke controles valt mee. Men heeft een comité dat dit opvolgt, die weet waarnaar te kijken. Ook is er een gedragscode.

Een tweede systeem is *on site*: de onderzoeker werkt op een stand-alone computer die bij de producent on site wordt ter beschikking gesteld. Dit moet in een waterdichte omgeving: geen printerpoort, geen disketteschrijver of zelfs cd-writer, zeker geen netwerkverbinding. Omdat ook dit nog niet volstaat, is het ook essentieel dat de databewaker de logfiles (bv. van SPSS) opvraagt.

Een praktischer systeem dan dit kan ook '*remote access*' zijn: de onderzoeker vraagt aan, bijvoorbeeld via e-mail, en de bestanden worden voor hem/haar klaargemaakt en doorgemailed.

Er is ook *software* die identificeerbaarheid kan zichtbaar maken (website van expert wordt nog doorgegeven): men maakt een soort 'tabellenboom': als er tabellen op een bepaald niveau worden aangemaakt, worden andere tabellen meteen op 'rood' gezet, t.t.z. die kunnen dan niet meer worden aangevraagd/aangemaakt.

Wat ten slotte ook zou kunnen, is de data ter beschikking stellen voor onderzoekers mits *postrandomisation*: de data worden daarbij opnieuw aangemaakt met dezelfde basiskenmerken. Bij dit systeem veranderen de gemiddelden niet, wel de variantie en de significantie, maar voor het oefenen op een bestand is dat minder erg. Dit is ook te combineren met on site of met remote access.

Het spreekt voor zich dat het bestaan van een intermediaire instelling als WSA een belangrijke rol kan spelen in de standaardisering van bestandsbeveiligingsmethodieken en het uitwerken van procedures die de bescherming van de privacy kunnen optimaliseren. Vlaanderen kent vooralsnog geen officiële instantie die als datamakelaar optreedt.

3.3 | On desk en on site werken op CBS-data

Het Centraal Bureau voor de Statistiek is de Nederlandse tegenhanger van het NIS. Het CBS heeft tot taak het verzamelen, bewerken en publiceren van statistieken ten behoeve van praktijk, beleid en wetenschap. Gegevens worden verzameld zowel bij individuen en huishoudens als bij bedrijven en instellingen. Deze informatieverschaffers wordt gegarandeerd dat de verzamelde gegevens uitsluitend worden gebruikt voor statistisch onderzoek.

Op databestanden uit *persoonsgerichte CBS-surveys* kan remote ('on desk') gewerkt worden. Gebruikers kunnen deze bestanden aankopen via het WSA. CBS levert de data dan op een cd-rom aan. Bij de beschikbaarstelling staat de bescherming van de privacy van de respondenten voorop. Er worden verschillende maatregelen toegepast, genomen om de privacy van de respondenten te garanderen.

- *Restricties in de toegang.* Alvorens een onderzoeksorganisatie een bestands-overeenkomst met het CBS kan aangaan, en daarmee toegang krijgt tot de bestanden, vindt er een toetsing plaats door de Centrale Commissie voor de Statistiek (CCS), een zelfstandig bestuursorgaan met een aantal taken in verband met het CBS. De toetsing is eenmalig en toestemming voor gebruik wordt gegeven per organisatie, niet per individuele gebruiker of per bestand. De belangrijkste criteria bij het al dan niet verlenen van de toestemming hebben betrekking op het doel van het gebruik (statistiek en wetenschap) en de fysieke en organisatorische beveiliging van de bestanden binnen de organisatie. De organisatie moet ook een contactpersoon/beheerder aanwijzen die aangesproken kan worden op de beveiliging van de bestanden.
- *Bestandsovereenkomst tussen CBS en gebruiker.* Nadat de toestemming van de CCS is verkregen, sluit de onderzoeksinstelling een bestandsovereenkomst met het CBS. De overeenkomst geeft toegang tot het bestellen van meerdere datasets. De condities die in de bestandsovereenkomst zijn opgenomen komen grotendeels overeen met de voorwaarden die in de PASO-bestandsovereenkomst gestipuleerd zijn. Ze hebben betrekking op onder meer het doel en niveau van de analyses, de bronvermelding en de eigendom van de bestanden. Ook CBS-bestanden mogen niet aan derden worden doorgegeven of op eigen initiatief aan andere bestanden gekoppeld worden. De overeenkomst behelst ook een geheimhoudingsverklaring die door alle onderzoekers en hun directe chef dient ondertekend te worden. De verklaring verplicht de gebruikers tot geheimhouding van elk gegeven omtrent afzonderlijke onderzoekseenheden (personen, huishoudens).
- *Bestandsbeveiliging.* Per survey wordt slechts één standaardbestand aangemaakt en verdeeld. Het in omloop brengen van meerdere versies, met verschillende variabelen of ongelijke niveaus van beveiliging, houdt immers risico's in voor de beveiliging van de data. In paragraaf 3.4 wordt die-

per ingegaan op de technieken die het CBS toepast om bestanden te beveiligen.

Bedrijfsgegevens zijn evenwel nog moeilijker te anonimiseren dan individuele data. Voor *bedrijfsgerichte CBS-bestanden* geldt bij wet dat deze het CBS niet mogen verlaten. Om onderzoekers toch toegang te geven, is binnen het CBS het Centrum voor Research op Economische Microdata (CEREM) opgericht, een on site faciliteit. Idee is dat externe onderzoekers op min of meer onbeveiligde data kunnen werken (enkel directe identificatoren worden verwijderd), maar dat de resultaten gescreend worden voor ze het CBS verlaten.

De procedure die het Cerem hanteert, is sterk uitgewerkt en erg uitgebreid. Naast de formalisering valt op dat van de gebruiker een reeks concrete 'tegenprestaties' worden verwacht: het betalen van een vergoeding, het volgen van een introductie over beveiliging, publicatie van de resultaten in de Ceremreeks, een niet-technische samenvatting en eventueel een voordracht over de onderzoeksresultaten.

- Alvorens een onderzoeksvoorstel in te dienen, voert de onderzoeker *overleg met de projectleider van Cerem*. Het gaat daarbij om doel, wenselijkheid, inpasbaarheid en haalbaarheid van het onderzoek. Mogelijke risico's van belangenverstremming zullen eveneens worden besproken. Indien nodig nemen aan dit overleg, dat eventueel uit meerdere sessies kan bestaan, de beheerder van de 'on site' faciliteit (Voorburg/Heerlen) en de gegevensbeheerder(s) van de desbetreffende statistische divisie(s) deel.
- Afhankelijk van de uitkomst van het overleg wordt een *onderzoeksvoorstel* (waarin doel en methode van onderzoek, benodigde data, tijdsduur/fasering, etc. beschreven zijn) *ingediend* bij de projectleider van Cerem. Dit voorstel wordt beoordeeld door het CBS aan de hand van de volgende punten:
 - kan het onderzoek schade opleveren voor de taak of de positie van het CBS?
 - zijn de data aanwezig om het beoogde onderzoek uit te voeren?
 - hoe groot is het risico van onthulling van uit concurrentieoogpunt gevoelige informatie, zowel tijdens als na afloop van het onderzoek?
- *Binnen vier weken* na ontvangst van het onderzoeksvoorstel deelt de projectleider van Cerem de *beslissing* van het CBS mee aan de indiener van het voorstel. Bij een positieve beslissing gaat deze mededeling vergezeld van een *conceptcontract*. Een negatieve beslissing gaat vergezeld van de motieven daarvoor.
- Het *contract bevat alle afspraken* met betrekking tot het onderhavige onderzoek. In het bijzonder maken deel uit van het contract:
 - het onderzoeksvoorstel;
 - de regels voor externe onderzoekers;
 - de regels betreffende het werken 'on site';

- de regels voor de beveiliging van publicaties die gebaseerd zijn op microdata van bedrijven (nota 302-96-RSM);
- een door de onderzoeker en het hoofd van het instituut waar de onderzoeker werkzaam is te ondertekenen geheimhoudingsverklaring.

In het contract wordt voorts de door de onderzoeker c.q. zijn/haar instituut van herkomst te betalen *vergoeding* vastgelegd. Deze vergoeding zal doorgaans uit twee componenten bestaan:

- een vergoeding voor het gebruiksklaar maken van de benodigde data en het recht op toelichting van de data door materiedeskundigen van het CBS;
- een vergoeding voor het gebruik van de 'on site' faciliteit volgens de daarvoor geldende regels.

Voor assistentie van de zijde van CBS-medewerkers die uitgaat boven toelichting van de data wordt het tarief van 'werk voor derden' in rekening gebracht. De omvang van dergelijke assistentie dient zo mogelijk ook in het contract te worden vastgelegd.

- Na ondertekening van het contract worden de benodigde *data* op het 'on site' netwerk *beschikbaar gesteld* in de vorm van een SPSS system file of een ASCII file. De onderzoeker heeft slechts toegang tot de met hem/haar contractueel overeengekomen data. Deze data bevatten *geen directe identificatoren* (zoals naam en/of adres).
- De onderzoeker regelt in overleg met de beheerder van de 'on site' faciliteit de *periode* waarin de *dagen* (eventueel dagdelen) waarop van deze faciliteit gebruik gemaakt wordt.
- Elke onderzoeker is *verplicht* in de beginfase van het onderzoek een korte *introductie* in het onderwerp 'beveiliging' te volgen.
- Het is de onderzoeker *verboden data*, tussenresultaten, etc., in welke vorm dan ook, *buiten de 'on site' ruimte te brengen*, tenzij met uitdrukkelijke toestemming van de projectleider van Cerem. Het is de onderzoeker *eveneens verboden niet-geautoriseerde personen toe te laten* in de 'on site' ruimte. Een overtreding van één van deze regels leidt tot onmiddellijke beëindiging van het onderzoek, onverminderd de volledige betaling van de contractueel overeengekomen vergoeding. De desbetreffende onderzoeker en eventueel het instituut waar zij/hij werkzaam is, kan worden uitgesloten van toekomstige participatie in Cerem.
- De projectleider van Cerem kan de onderzoeker toestaan, via de beheerder van de 'on site' faciliteit, *eigen programmatuur* en/of data ten behoeve van het onderzoek op het 'on site' netwerk te importeren.
- Op verzoek van de onderzoeker kan de projectleider van Cerem de onderzoeker in contact brengen met *materiedeskundigen* ter toelichting van de data (met betrekking tot steekproefopzet, imputaties e.d.).
- De *onderzoeker rapporteert regelmatig* over de voortgang van het onderzoek aan de projectleider van Cerem.

- Bij het schrijven van een conceptpublicatie volgt de onderzoeker nauwgezet de *regels voor de beveiliging* (nota 302-96-RSM). *Elke conceptpublicatie wordt voorgelegd aan de projectleider* van Cerem. Deze beoordeelt of laat beoordelen of aan de regels voldaan is. De onderzoeker dient zich te realiseren dat met deze procedure enige tijd (één à twee weken) gemoeid kan zijn, en dat het verstandig is bij de planning van vervolgvactiteiten (bijvoorbeeld de indiening voor een congres, het houden van een externe voordracht, een promotie) hiermee rekening te houden.
- *Publicatie* vindt in eerste instantie plaats in de reeks Cerem-rapporten (en eventueel tegelijkertijd in de rapportenreeks van het instituut van herkomst van de onderzoeker). Het manuscript wordt daartoe in de voorgeschreven lay-out aangeleverd. De onderzoeker ontvangt maximaal 35 exemplaren gratis. Tegen meerkosten kunnen extra exemplaren worden geleverd. Het staat de onderzoeker daarna vrij een dergelijk rapport vervolgens aan te bieden aan een congres, een tijdschrift of een bundel, dan wel het te gebruiken als onderdeel van een door hem/haarzelf te vervaardigen groter geheel (boek, proefschrift, etc.). Zowel de oorspronkelijke publicatie als de op basis daarvan vervaardigde vervolgpublishaties dienen melding te maken van het feit dat het onderhavige onderzoek bij het CBS c.q. Cerem plaatsgevonden heeft en dienen een *'disclaimer'* te bevatten ('The views expressed in this paper are those of the author and do not necessarily reflect the policies of Statistics Netherlands' dan wel het Nederlandse equivalent hiervan).
- Daarnaast wordt van elke onderzoeker verwacht dat hij/zij *desgevraagd* over de resultaten van het onderzoek een *voordracht* houdt in de reeks RNO-discussiebijeenkomsten (gewoonlijk op vrijdagmiddag) en dat hij/zij een beknopte *niet-technische samenvatting* schrijft (onder andere ten behoeve van het jaarverslag van Cerem). In dit stuk dient met name de maatschappelijke en/of wetenschappelijke relevantie van het onderzoek aan de orde te komen.
- Na voltooiing van het onderzoek worden de gebruikte *data* en de eventueel door de onderzoeker geïmporteerde of vervaardigde *programmatuur* (set up's e.d.) van het 'on site' netwerk *verwijderd*. De onderzoeker kan verzoeken om bewaring van de data en/of de programmatuur voor vervolgonderzoek. Hiervoor kunnen kosten in rekening worden gebracht.
- Er wordt een *register* bijgehouden waarin de naam (namen) van de onderzoeker (s), het onderzoeksproject, de publicatie(s), en de beschikbaar gestelde databestanden vermeld staan. Dit register is toegankelijk voor berichtgevers van het CBS en wordt gepubliceerd in het jaarverslag van Cerem.

3.4 | Methodieken van bestandsbeveiliging

In het kader van het Europese CASC-project (Computational Aspects of Statistical Confidentiality, <http://neon.vb.cbs.nl/casc>) is onderzoek verricht naar de ontwikkeling en implementatie van methoden van bestandsbeveiliging ('statistical disclosure control'). Twee softwarepakketten werden ontwikkeld: MU-ARGUS voor het beveiligen van microbestanden en TAU-ARGUS voor het beveiligen van tabellen. Beide programma's zijn van de CASC-website te downloaden. Omdat we vooral geïnteresseerd zijn in de beveiliging van datasets gaan we in wat volgt dieper in op MU-ARGUS.⁵

Simpel gesteld valt de beveiliging van datasets uiteen in twee stappen: het opsporen en het remediëren van veiligheidsrisico's. Eerst moet vastgesteld worden welke gegevens (variabelen of combinaties van variabelen) onveilig zijn. Vervolgens worden deze gegevens zodanig 'bewerkt' dat de kans op identificatie geminimaliseerd wordt. De manipulatie van de data mag de bruikbaarheid echter niet al te zeer verminderen. Kunst is zowel het risico op openbaring of onthulling ('disclosure') als het informatieverlies aanvaardbaar laag te houden.

Databeveiliging houdt de minimalisering in van het risico dat derden uit een dataset gevoelige informatie over individuele respondenten kunnen afleiden. Identificatie⁶ van een individuele respondent is mogelijk aan de hand van een sleutel, dit is een combinatie van waarden op identificerende variabelen. Als een bepaalde combinatie in de populatie uniek of zeldzaam is, en de dataset bevat één (of enkele) respondenten met deze combinatie, dan is het risico groot dat de respondent(en) geïdentificeerd kunnen worden. 'Man, boekhouder, Antwerpen' is een sleutel die minder risico's op identificatie inhoudt dan bijvoorbeeld 'vrouw, vrachtwagenchauffeur, Herstappe'. Informatie over het aantal eenheden in de populatie met een bepaalde sleutel is echter vaak niet voorhanden en moet dan ook geschat worden aan de hand van de dataset. Om het onderscheid te kunnen maken tussen veilige en onveilige waarden, moet de databeveiligaar drempelwaarden bepalen. Waarden of combinaties van waarden worden als veilig beschouwd als de geschatte frequentie van het voorkomen van de waarde of de combinatie in de populatie (of de eigenlijke frequentie in het bestand) boven de drempelwaarde ligt. Ligt de (geschatte) frequentie onder de drempelwaarde, dan wordt de waarde of combinatie als onveilig beschouwd en dient men maatregelen te nemen.

- *Globale hercodering (global recoding)*. Verschillende waarden van een variabele worden samengevoegd. Dit gebeurt niet enkel voor observaties met onveilige waarden of combinaties, maar voor de hele dataset. Stel bijvoor-

⁵ Zie http://neon.vb.cbs.nl/casc/deliv/mu_manual_31.pdf voor een uitgebreide handleiding (met achtergronden) bij MU-ARGUS.

⁶ In het Engels spreekt men van *re-identification* of heridentificatie. De respondent is inderdaad reeds geïdentificeerd, namelijk door de dataverzamelaar.

beeld dat het aantal universiteitsprofessoren en/of het aantal hogeschool-docenten in een bestand onder de drempelwaarde ligt, dan kan men deze waarden samenvoegen tot bijvoorbeeld 'docenten en professoren'. De kans op identificatie aan de hand van combinaties met beroep wordt door deze ingreep kleiner.

- *Onderdrukking van waarden (local suppression)*. Eén of meerdere waarden in een onveilige combinatie worden onderdrukt, dit wil zeggen vervangen door een ontbrekende waarde. Zo kan in de combinatie 'vrouw, vrachtwagenchauffeur, Herstappe' één of meer van de drie waarden op missing worden gezet om identificatie van deze respondent te bemoeilijken. In tegenstelling tot globale hercodering laat men de andere waarden op dezelfde variabele ongemoeid. Het is dus niet omdat in onze voorbeeldcombinatie bijvoorbeeld Herstappe wordt onderdrukt, dat voor alle andere observaties ook de woonplaats op missing wordt gezet.
- *Herocodering 'naar beneden' of 'naar boven' (top and bottom coding)*. Bij ordinale en metrische variabelen kunnen de laagste of de hoogste waarden samengevoegd worden in een nieuwe categorie. Een voorbeeld van *top coding* is het toewijzen van alle gezinnen met meer dan vijf kinderen aan een categorie 'zes of meer'.
- *Afronden (rounding)*. Bij metrische variabelen kunnen zeldzame waarden afgerond worden om de kans op identificatie te verkleinen.
- *Postrandomisering (postrandomization)*. Deze techniek houdt in dat (bepaalde) waarden op een categorische variabele vervangen worden door andere waarden, niet systematisch zoals bijvoorbeeld bij globale hercodering, maar via een vooraf bepaald en dus gekend lotingschema (*probability mechanism*). De gebruikers van de dataset krijgen informatie over het lotingschema zodat ze hun schattingen kunnen corrigeren. Zo kunnen drie inwoners uit Hasselt na postrandomisering als woonplaats bijvoorbeeld Brugge, Bierbeek en Zottegem krijgen. Het spreekt voor zich dat de gemeente-informatie in een dergelijk postgerandomiseerd bestand door een kwaadwillende gebruiker niet meer voor identificatiedoelinden kan gebruikt worden. Meer informatie over de exacte werkwijze, voor- en nadelen vindt u op http://neon.vb.cbs.nl/casc/related/Sdp_98_2.pdf.

Deze technieken voor databeveiliging zijn alle in MU-ARGUS geïntegreerd. Eenvoudig gesteld geeft de gebruiker aan welke variabelen hij als identificerende variabelen beschouwt (met vier gradaties: directe identificatoren, identificerende variabelen, meer identificerende variabelen en meest identificerende variabelen) en wat de drempelwaarden zijn voor hun gecombineerde frequenties. Het pakket geeft dan een overzicht van de onveilige combinaties. Hierop kan men dan één of meerdere van de hierboven beschreven *disclosure control procedures* toepassen, volgens parameters die men zelf kan instellen. Eerst worden de technieken toegepast die het detailniveau verlagen (globale hercodering, hercodering naar beneden

of naar boven, afrondingen) of die de waarden veranderen (postrandomisering). Vervolgens wordt berekend of en hoeveel onveilige combinaties er overblijven. In deze combinaties kan men waarden laten onderdrukken.

Het pakket laat ook toe de impact van de toegepaste technieken te evalueren. Via zogenaamde protectierisicomodellen kan men de veiligheid van beveiligde bestanden inschatten. Naast de veiligheid wordt ook de bruikbaarheid van de beveiligde data gekwantificeerd, via een inschatting van het informatieverlies. De impact die de bestandsbeveiligingstechnieken hebben op de analytische kracht van de data is inderdaad een reële bekommernis. Men is immers niets met erg goed beveiligde bestanden als de analyses onbetrouwbare of ongeldige resultaten opleveren.

Conclusie

Niettegenstaande de verdiensten van het MU-ARGUS-pakket lijkt toepassing van deze software in het kader van PASO-samenwerkingsverbanden ons op dit moment niet aangewezen. Het onderzoek naar de implicaties op de bruikbaarheid van de beveiligde data is, hoewel reeds gevorderd, verre van geruststellend. Eén van de ontwikkelaars van MU-ARGUS, Anco Hundepool (CBS), geeft overigens toe dat de software meer geschikt is voor het beveiligen van 'sociale data' (gegevens verzameld bij individuen) dan voor economische data. Het is mede om deze reden dat CBS voor bedrijvendata on site faciliteiten (Cerem, supra) in het leven heeft geroepen.

We kiezen ervoor de on site faciliteiten voor PASO verder uit te bouwen. Het gebruikskomfort voor de onderzoeker ligt weliswaar iets lager (men werkt niet in de eigen vertrouwde omgeving), maar daartegenover staan meer en directere ondersteuningsmogelijkheden. Ons lijkt de anonimiteit van het PASO-bestand bij on site werken (inclusief supervisie op de werkzaamheden en screening van de resultaten achteraf) minstens in dezelfde mate gegarandeerd als bij remote werken op een softwarematig gemanipuleerd bestand. Bovendien is de wetenschappelijke waarde van analyses op een versleuteld bestand steeds aan een zekere twijfel onderhevig.

De optie die de PASO-equipe dus voorstelt is on site analyses op een bestand waaruit enkel de directe identificerende variabelen verwijderd zijn. De gebruiker kan in een gecontroleerde omgeving op het best mogelijke bestand werken. Hij of zij is er wel toe gehouden de data enkel voor het vooropgeschreven doel aan te wenden en wetenschappelijk correct te rapporteren. Misbruiken worden in de praktijk bemoeilijkt of onmogelijk gemaakt door de technische set-up van de stand-alone computer, en zijn achteraf, via inspectie van de logbestanden en de analyseresultaten, objectief detecteerbaar. Het feit dat de gebruiker vooraf op de hoogte wordt gebracht van de screeningsprocedure heeft op zich reeds een ontradend effect.

Hoofdstuk 4

On site werken op PASO-databestanden: databeveiliging via een stand-alone computer

Eén van de mogelijke vormen van samenwerking om overbevraging van organisaties te reduceren, is geïnteresseerde wetenschappers laten werken op de PASO-data. Dit kan in de eerste plaats 'on site'. Van deze optie werd onder meer maar niet alleen gebruik gemaakt door studenten die in het kader van hun licentiaatsthesis gerichte analyses wilden uitvoeren op de databestanden van de eerste en/of de tweede golf. Bij deze samenwerkingen achtte de PASO-equipe het weinig opportuun om het bestand in fysieke vorm aan de gebruiker over te maken. Er werden een reeks maatregelen uitgewerkt om de beveiliging van het bestand te maximaliseren, zonder evenwel de toegankelijkheid van de data te compromitteren.

- 'On site' betekent concreet een leegstaand lokaal in één van de HIVA-gebouwen. In dit lokaal is op elk moment controle van de gebruiker mogelijk. Er is behalve een pc ook een telefoon aanwezig, zodat de gebruiker snel en kosteloos een PASO-medewerker om hulp kan vragen.
- De wegschrijfmogelijkheden van de computer zijn beperkt om 'diefstal' van het bestand of een deel ervan zo goed mogelijk uit te sluiten. Vandaar dat men spreekt over een stand-alone pc: een computer die 'op zich staat' en speciaal geconfigureerd is om dataoverdracht en communicatie met andere systemen (in theorie) onmogelijk te maken. Concreet werden volgende voorzieningen getroffen:
 - de pc is niet verbonden met het internet, noch aangesloten op een intern netwerk, wat elektronische transfer van bestanden via e-mail, ftp of een-der welk ander protocol onmogelijk maakt;
 - de pc bevat een cd-rom (nodig om het bestand op de pc te plaatsen), maar met dit apparaat kunnen geen cd-r's of cd-rw's geschreven worden;
 - het disktestation werd niet verwijderd, zodat de gebruiker analyse-resultaten naar diskette kan wegschrijven en niet verplicht is tabellen en dergelijke meer manueel over te schrijven. De databestanden zelf zijn evenwel te groot om op diskette op te slaan;

- compressieprogramma's zoals Winzip en Winrar werden van de pc verwijderd om te vermijden dat (deel)bestanden gecomprimeerd op diskette weggeschreven worden;
- het geïnstalleerde besturingssysteem WinNT kent geen ondersteuning van USB-poorten; gebruik van een memory-stick of een ander USB-apparaat is dus evenmin mogelijk.
- We verwijderen variabelen die individuele vestigingen identificeerbaar kunnen maken uit de databestanden die op de stand-alone computer worden geplaatst. Het gaat hier in de eerste plaats om de naam en coördinaten van de organisatie bevraagd in het eerste deel van de vragenlijst onder de titel 'identificatie'. Maar ook een aantal andere variabelen, zoals de vraag achteraan naar eventuele opmerkingen bij de vragenlijst, zijn weggelaten. Afhankelijk van de onderzoeksvragen van de gebruiker wordt beslist om het integrale bestand (met uitzondering van de identificatievariabelen) of een deelbestand ter beschikking te stellen. Een deelbestand bevat enkel de vragenlijstonderdelen waarin de gebruiker geïnteresseerd is. De praktijk van het weglaten van identificatievariabelen sluit de mogelijkheid van een toevallige identificatie van een organisatie niet uit (bv. op basis van sector, grootte en enkele meer inhoudelijke kenmerken), maar ook in dit geval is de gebruiker gebonden door de bestandsovereenkomst die hij heeft ondertekend en die hem verbiedt uitspraken te doen over individuele cases.
- De bestanden worden, naar voorkeur van de gebruiker, beschikbaar gesteld als SAS- en/of als SPSS-bestand. Ook tot het codeboek, onmisbaar voor een correct en efficiënt gebruik van de data, krijgt de gebruiker toegang.

Conclusies

- Het on site werken met een stand-alone computer waarborgt tegelijkertijd een gedegen bescherming van de data en goede ondersteuningsmogelijkheden voor de gebruiker. Op beide vlakken zijn evenwel nog verbeteringen denkbaar. Zo kunnen we overwegen om de logbestanden van het statistisch pakket bij te houden. Op die manier kan exact nagegaan worden welke bewerkingen en analyses er gebeurd zijn. Het gebruikscomfort kan verder verhoogd worden door het opstellen van een lijst met 'veel gestelde vragen' (FAQ) als praktische aanvulling op het codeboek (bv. hoe weegfactor activeren in SPSS, bijzondere aandachtspunten in SAS-programmatuur).
- Het toelaten van on site werken op de data brengt een relatief grote tijdsinvestering van de onderzoekers met zich mee: afspraken maken met betrekking tot het gebruik van het lokaal, het bestand klaarmaken en overzetten naar de stand-alone pc, begeleiding en controle van de gebruiker, eventuele 'nazorg'.
- In elk geval leert het databeheerproject dat het on site werken op de data de meest verkieslijke vorm van samenwerking met derden is, en alleszins

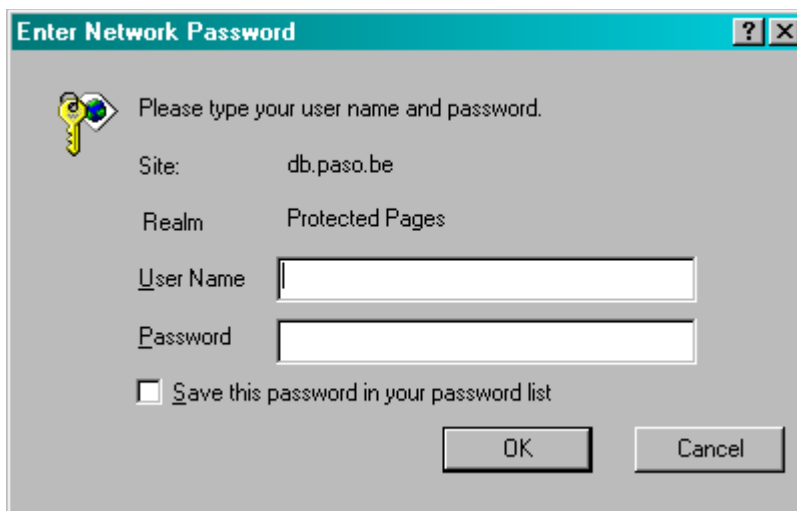
de voorkeur verdient boven het ter beschikking stellen van data of (deel)-bestanden aan derden. Dit laatste is vanuit privacy- en kwaliteitsbepalingen niet wenselijk.

Hoofdstuk 5

Monitoringsysteem PASO-samenwerkingsverbanden

Eén van de opdrachten van het databeheerproject was het uitwerken van een 'monitoringsysteem om het gebruik van het bestand door derden te registreren, te evalueren en daarover te rapporteren'.

Voor deze opdracht werd een aparte website aangemaakt (<http://db.paso.be>). Enkel gebruikers met een geldige login (combinatie van gebruikersnaam en paswoord) krijgen toegang tot de website. Er zijn twee logins: één voor bezoekers (met enkel leesrechten) en één voor PASO-medewerkers (lees- en schrijfrechten).



Enter Network Password

Please type your user name and password.

Site: db.paso.be

Realm: Protected Pages

User Name

Password

Save this password in your password list

OK Cancel

Na het inloggen komt men op de startpagina terecht.

Address  http://db.paso.be/

PASO-databeheer

- Vragenlijsten **golf 1, golf 2** (PASO-site)
- Codeboek **golf 1, golf 2** (Word)
- Overzicht **rubrieken, thema's en vragen** golf 1 en golf 2 (Excel)
- Overzicht van de **PASO-samenwerkingsverbanden**
De databank bevat momenteel informatie over 25 samenwerkingen.
Hiervan zijn er 15 afgerond, 5 actueel en 5 in aanvraag.

Hier kan men de vragenlijsten, de codeboeken en een overzicht in spreadsheetvorm van de vragenlijsten, rubrieken, thema's en vragen raadplegen. Het belangrijkste is echter de link naar de databasetoepassing, die tot doel heeft de samenwerkingsverbanden met derden te registreren, te documenteren en op te volgen.

Address  http://db.paso.be/lijsjt.php

Overzicht van de PASO-samenwerkingsverbanden

Afgeronde samenwerkingen

organisatie	contactpersoon	project	begin	einde
Administratie Planning en Statistiek (APS)	Thierry Vergeynst	module "Economische indicatoren" (golf 1)		
Centrum voor Bevolkings- en Gezinsstudie (CBGS)	Walter Van Dongen	"Werken aan het combinatiemodel. Beschikbare faciliteiten van het combinatiemodel in Vlaamse organisaties"		
Faculteit Economie en bedrijfskunde, UG	Aimé Heene	module "Maatschappelijk verantwoord ondernemen" (golf 2)		
Hoger Instituut voor de Arbeid (HIVA)	Miet Lamberts	"Krapte op de arbeidsmarkt, kansen voor vrouwen? Een sectorale invalshoek"		
Vereniging van Vlaamse Kamers voor Handel en Nijverheid	Kris Vanherpe	regionale rapporten		jun 03

De hoofdpagina van de toepassing geeft in tabelvorm een overzicht van alle samenwerkingsverbanden. De samenwerkingen staan geordend volgens hun status: afgerond, actueel of aangevraagd. Klikte men op een rij, dan krijgt men een steekkaart te zien met informatie over:

- de *organisatie* waarmee werd samengewerkt;

- de *persoon/personen* met wie binnen de organisatie werd samengewerkt, respectievelijk promotoren/projectleiders en medewerkers/onderzoekers;
- het *project* van de partnerorganisatie waarbinnen de samenwerking met PASO tot stand kwam;
- de *vorm* van de samenwerking: wat hield de samenwerking in?;
- de inhoud van de samenwerking: op welke *thema's* had de samenwerking betrekking?;
- de *status* van de samenwerking: aangevraagd, actueel of afgerond;
- de *looptijd* van de samenwerking: begin- en einddatum (bij benadering);
- registratie en *documentering* van de samenwerking: projectvoorstel, bestandsovereenkomst, output van de samenwerking (onderzoeksrapporten e.d.), enz.

http://db.paso.be/fiche.php?id=8

SAMENWERKING MET

Faculteit Toegepaste Economische Wetenschappen, UA

[overzicht](#) | [updaten](#)

ORGANISATIE

NAAM
Faculteit Toegepaste Economische Wetenschappen, UA

COÖRDINATEN
Universiteit Antwerpen
Faculteit TEW
Prinsstraat 13
B-2000 Antwerpen
België
Tel. +32-3-220 47 96
Fax +32-3-220 47 99

WEBSITE
www.tew.ua.ac.be

CONTACTPERSOON

NAAM

Opgeladen documenten kan men met een simpele muisklik downloaden.

DOCUMENTATIE

omschrijving	bestandsnaam	opmerkingen	verwijder
rapporten Antwerpen, Antwerpen-Waasland, Limburg	kamerrapporten_antw_antwaaas_limb.zip	gezipt	*
rapporten Vlaams-Brabant, Oost- en West-Vlaanderen	kamerrapporten_vlbr_ovl_wvl.zip	gezipt	*

[bestand toevoegen](#)

Bezoekers met schrijfrechten kunnen informatie over een nieuwe samenwerking toevoegen, de gegevens van eerder ingevoerde samenwerkingen updaten en bestanden toevoegen.

Organisatie

naam

coördinaten

url

Contactpersoon 1

naam

email

telefoon

functie

Er is ook een link voorzien naar een script dat de basisgegevens van alle samenwerkingen in lijstvorm onder elkaar plaatst. Op deze wijze kan eenvoudig een up-to-date overzicht gegenereerd worden.

Beknopt overzicht van de PASO-samenwerkingsverbanden op 27-08-2004

1. *Administratie Planning en Statistiek (APS)*
 - o **project:** module "Economische indicatoren" (golf 1)
 - o **vorm:** module vragenlijsten golf 1 ('Economische indicatoren'); analyses door PASO-team op deze module
 - o **status:** afgerond
2. *Administratie Werkgelegenheid*
 - o **project:** Strategisch Plan Werkgelegenheid 2003
 - o **vorm:** ad hoc analyses door het PASO-team op golf 1 en golf 2
 - o **status:** afgerond
3. *Centrum voor Bevolkings- en Gezinsstudie (CBGS)*
 - o **project:** "Werken aan het combinatiemodel. Beschikbare faciliteiten van het combinatiemodel in Vlaamse organisaties"
 - o **vorm:** module vragenlijst golf 1 en analyses golf 1
 - o **status:** afgerond
4. *Departement Toegepaste economische wetenschappen, K.U.Leuven*
 - o **project:** paper over de factoren die het gebruik van een HR-Scorecard bepalen
 - o **vorm:** analyses op golf 2
 - o **status:** afgerond

Het grote voordeel van een dergelijk interactief monitoringsysteem is dat zowel opdrachtgevers als medewerkers op elk moment op de site terecht kun-

nen voor een actuele stand van zaken. De informatie in de database kan ook makkelijk in eender welk format op het scherm gebracht worden: in de vorm van tabellen, lijsten, gedetailleerde steekkaarten, enz. Dit voorkomt lay-out-werk bij (tussentijdse) rapporteringen.

Er kan aan gedacht worden deze databasetoepassing als uitgangspunt te nemen voor een on line registratiesysteem voor kandidaat-gebruikers. Geïnteresseerde derden zouden zich op de site volledig kunnen informeren over mogelijkheden om op de PASO-data te werken: wat is de procedure (protocol en bestandsovereenkomst), welke vragen zijn opgenomen in welke vragenlijsten, enz.? Vervolgens kan de geïnteresseerde on line een aanvraag doen. Eventuele documentatie, zoals het projectvoorstel, kan opgeladen worden naar de server. Voordeel van een dergelijk systeem is dat de gebruiker zelf de velden invult, en dit dus niet meer door een PASO-medewerker hoeft te gebeuren. Na het invullen van het aanvraagformulier zou de aanvrager een op maat gesneden bestandsovereenkomst (naam, project e.d. ingevuld) op het scherm krijgen, die hij of zij moet afprinten, ondertekenen en opsturen. Een dergelijke werkwijze laat het PASO-team toe de aanvragen en bestandsovereenkomsten zowel in digitale vorm (output van de on line registratie) als in fysieke (papieren) vorm te archiveren.

Hoofdstuk 6

Overzicht van de PASO-samenwerkingsverbanden

6.1 | Beknopt overzicht van de PASO-samenwerkingsverbanden

Het volgend overzicht geeft beknopt weer welke samenwerkingsovereenkomsten er zijn afgesloten tot en met eind september 2004. Er dient te worden opgemerkt dat deze mogelijkheden zijn afgebouwd op het einde van 2004 omdat er op dat moment geen voldoende zekere vooruitzichten waren op een verlenging van het PASO project. Dit bracht met zich mee dat verschillende aanvragen van voornamelijk wetenschappers om te werken op de PASO-data moesten geweigerd worden omwille van het ontbreken van de nodige omkadering.

1. Administratie Planning en Statistiek (APS):
 - project: module 'economische indicatoren' (golf 1);
 - vorm: module vragenlijsten golf 1 ('economische indicatoren'); analyses door PASO-team op deze module;
 - status: afgerond.
2. Administratie Werkgelegenheid:
 - project: Strategisch Plan Werkgelegenheid 2003;
 - vorm: ad hoc analyses door het PASO-team op golf 1 en golf 2;
 - status: afgerond.
3. Centrum voor Bevolkings- en Gezinsstudie (CBGS):
 - project: 'Werken aan het combinatiemodel. Beschikbare faciliteiten van het combinatiemodel in Vlaamse organisaties';
 - vorm: module vragenlijst golf 1 en analyses golf 1;
 - status: afgerond.
4. Departement Toegepaste economische wetenschappen, K.U.Leuven:
 - project: HR-outsourcing, kans of bedreiging?;
 - vorm: analyses op golf 2;
 - status: afgerond.

5. DIVA, Dienst Informatie Vorming en Afstemming, Ministerie van de Vlaamse Gemeenschap:
 - project: realisatie van het 'Jaarboek levenslang en levensbreed leren 2003';
 - vorm: analyses op golf 2;
 - status: afgerond.
6. Els Bieseman (UA):
 - project: licentiaatsthesis over stand van zaken van e-HRM in Belgische organisaties;
 - vorm: analyses op golf 2;
 - status: afgerond.
7. Faculteit Economie en bedrijfskunde, UG:
 - project: module 'Maatschappelijk verantwoord ondernemen' (golf 2);
 - vorm: module vragenlijst golf 2;
 - status: afgerond.
8. Hannelore Harts (UA):
 - project: licentiaatsthesis 'Nieuwe organisatieconfiguraties en 'human resource management' in Vlaanderen';
 - vorm: analyses op golf 1;
 - status: afgerond.
9. Hoger Instituut voor de Arbeid (HIVA):
 - project: 'Krapte op de arbeidsmarkt, kansen voor vrouwen? Een sectorale invalshoek';
 - vorm: module vragenlijst golf 1 en analyses op golf 1;
 - status: afgerond.
10. Karolien Vanbrabant (K.U.Leuven):
 - project: licentiaatsthesis 'De afstemming van arbeid en gezin in Vlaanderen. Het voorkomen van combinatiemaatregelen en het effect op de prestatieindicatoren en vrijwillig verloop';
 - vorm: analyses op golf 1;
 - status: afgerond.
11. Steunpunt Werkgelegenheid, Arbeid en Vorming (WAV):
 - project: 'Jaarboek arbeidsmarkt Vlaanderen, editie 2003';
 - vorm: analyses op golf 1;
 - status: afgerond.
12. Stijn Tavernier (K.U.Leuven):
 - project: licentiaatsthesis 'Het gebruik van strategieën ter flexibilisering van de arbeid, een analyse voor het Vlaamse bedrijfsleven';

- vorm: analyses op golf 1;
 - status: afgerond.
13. VDAB:
- project: 'Benchmark Competentiecentra';
 - vorm: analyses door het PASO-team op golf 2;
 - status: afgerond.
14. Veerle Loeckx (K.U.Leuven):
- project: licentiaatsthesis 'De rol van de human resource manager en de human resource architectuur in de praktijk';
 - vorm: analyses op golf 2;
 - status: afgerond.
15. Vereniging van Vlaamse Kamers voor Handel en Nijverheid:
- project: regionale rapporten;
 - vorm: analyses door PASO-team op golf 1;
 - status: afgerond.
16. Departement Toegepaste economische wetenschappen, K.U.Leuven:
- project: 'Evaluatie van de opleidingscheques-maatregel';
 - vorm: analyses op golf 2;
 - status: actueel.
17. Faculteit Toegepaste economische wetenschappen, UA:
- project: 'De quartaire sector';
 - vorm: analyses op golf 1 en 2;
 - status: actueel.
18. Kelly Vereeken (K.U.Leuven):
- project: licentiaatsthesis 'De aanwezigheid van allochtonen op de Vlaamse werkvloer: een analyse op basis van kwantitatieve gegevens';
 - vorm: analyses op golf 2;
 - status: actueel.
19. Landelijke Bedienden Centrale (LBC):
- project: toeleveren achtergrondmateriaal vormingsdagen LBC-metaal;
 - vorm: analyses door het PASO-team op golf 1 en golf 2 omtrent de Vlaamse metaalindustrie;
 - status: actueel.
20. Quintessence:
- project: 'Competentiemanagement';
 - vorm: analyses door PASO-team op golf 2;
 - status: actueel.

21. Vlerick Leuven Gent Management School:
 - project: verwerking PASO-module 'Mainstreaming van maatschappelijk verantwoord ondernemen (MVO)';
 - vorm: analyses op golf 2;
 - status: actueel.

22. Cevora:
 - project: 'Feasibility-studie voor de uitbreiding van de PASO-enquête voor heel België naar sectoren die door CEVORA vertegenwoordigd worden';
 - vorm: feasibility-studie;
 - status: aangevraagd.

23. Nino Di Giusto (K.U.Leuven):
 - project: licentiaatsthesis rond competentie management;
 - vorm: analyses op golf 2;
 - status: aangevraagd.

24. Vacature:
 - project: reeks 'Talentcentra';
 - vorm: analyses door PASO-team; tabellen m.b.t. aantal hooggeschoolden in de bedrijven, aantal tewerkgestelde migranten, vormingsinvesteringen van bedrijven en aantal starters;
 - status: aangevraagd.

25. VOKA:
 - project: innovatieindicatoren;
 - vorm: analyses op golf 3;
 - status: aangevraagd.

26. Kenniscentrum Toerisme en Horeca:
 - project: 'Foto van de sector';
 - vorm: analyses op drie PASO-golven en START, enkel voor toerisme en horeca;
 - status: analyses worden uitgevoerd.

■

Hoofdstuk 7

PASO-samenwerkingsprotocol en bestandsovereenkomsten

De ervaringen die zijn opgebouwd in het project in verband met het laten werken op de PASO-data door derden, hebben aanleiding gegeven tot een evaluatie en beperkte bijsturing van het zogenaamde samenwerkingsprotocol en de samenwerkingsovereenkomsten die als leidraad dienen voor de betrokken partijen bij de samenwerking, aan de ene kant het PASO-team en aan de andere kant de onderzoekers die van de data gebruik willen maken voor hun wetenschappelijk werk. De voornaamste aanpassingen houden in dat aan derden nooit een databestand noch een uittreksel daarvan worden doorgegeven. Met andere woorden, enkel aangemaakte tabellen kunnen de K.U.Leuven ‘verlaten’. Voorts kan de betrokkene eveneens ter plekke werken op de data op de voorziene infrastructuur (een stand-alone pc), zoals toegelicht in hoofdstuk 4.

7.1 | Protocol PASO-samenwerkingsverbanden

7.1.1 | Filosofie

Het PASO-team streeft een zo ruim mogelijk gebruik van de data en een zo breed mogelijke valorisatie van de onderzoekresultaten na, evenwel zonder toegevingen te doen op vlak van privacybescherming. De gebruiker verbindt zich via de overeenkomst en de geheimhoudingsverklaring tot het respecteren van juridische, wetenschappelijke en deontologische regels. U vindt alle details in de modellen van overeenkomst (zie paragraaf 7.2). Het PASO-team waakt over de naleving van deze normen, onder andere via de controle van de output die uit de samenwerking voortvloeit. Op alle onderzoeksverslagen wordt voorafgaand aan publicatie een kwaliteitscontrole doorgevoerd.

7.1.2 | Samenwerkingsvormen

Het PASO-team voorziet vijf vormen van samenwerking met externen. We beschrijven wat elke samenwerkingsformule inhoudt, welke procedure gevolgd dient te worden, wat de voorwaarden zijn en wie voor de samenwerking in

aanmerking komt. Bij dit laatste maken we een onderscheid tussen 'publieke gebruikers' (beleidsverantwoordelijken, overheidsdiensten), 'wetenschappelijke gebruikers' (individuele onderzoekers of onderzoeksinstellingen) en 'andere niet-commerciële derden' (individuen en niet-publieke non-profit-organisaties).

1. *Samenwerking op niveau van de vragenlijst*

Wat? Onder bepaalde voorwaarden kan een *module* (= samenhangend geheel van vragen over één onderwerp of deelgebied) aan de PASO-vragenlijst gekoppeld worden. Overbevraging kan op deze manier tegengegaan worden. Zowel in de eerste als in de tweede PASO-golf werden twee externe modules in de bevraging geïntegreerd.

Hoe? De *opportuniteit* van het integreren van de voorgestelde module in de vragenlijst wordt door het PASO-team *beoordeeld*. De lengte en thematische strekking van de hoofdvragenlijst zijn hierbij belangrijke aandachtspunten. Daar deze vorm van samenwerking rechtstreeks ingrijpt op het concept van de bevraging dient de leverancier van de module een reeks technische en inhoudelijke richtlijnen na te leven. Deze richtlijnen zijn opgenomen in een *draaiboek*. De aanvrager staat in voor de ontwikkeling van de module en de daarbij horende *kosten*. Tevens moet de gebruiker kunnen garanderen dat de moduledata ook daadwerkelijk geanalyseerd zullen worden.

Wie? Publieke en wetenschappelijke gebruikers.

2. *Samenwerking op niveau van de steekproeftrekking*

Wat? Het financieren van *bijkomende steekproefeenheden*, complementair aan de PASO-steekproef, al dan niet selectief ten aanzien van bepaalde segmenten uit de populatie (sectoren, microbedrijven, ...). Op die manier kunnen beide partijen beschikken over een ruimere steekproef. Deze piste heeft in de praktijk enkel zin als ze samen gaat met de toevoeging van een module en als er afspraken gemaakt worden omtrent de verwerking en analyse van de gegevens.

Hoe? De opdrachtgever en het PASO-team sluiten een *contract* af. Hierin worden alle modaliteiten van de samenwerking opgenomen. De extra steekproefeenheden worden aangerekend tegen de *geldende eenheidsprijs*.

Wie? Publieke en wetenschappelijke gebruikers, andere niet-commerciële derden.

3. *Analyses door het PASO-team, rapportering door de gebruiker*

Wat? Het PASO-team levert op aanvraag *cijfermateriaal* aan derden aan. Hierbij kan gedacht worden aan percentages, correlaties of aan de resultaten van complexere analyses. De cijfers worden in 'naakte' vorm aangeleverd, de interpretatie en rapportering gebeurt door de aanvrager.

Hoe? De gebruiker dient de *bestandsovereenkomst* (type A), inclusief geheimhoudingsverklaring in te vullen en te ondertekenen. Verder wordt een *uurprijs* aangerekend voor de analyses.

Wie? Publieke en wetenschappelijke gebruikers, andere niet-commerciële derden.

4. *Analyses én rapportering door het PASO-team*

Wat? Het PASO-team voert in opdracht van een externe organisatie *analyses* uit en werkt de resultaten uit tot een *onderzoeksrapport*. De inhoud, aard en omvang van het rapport, alsook de inbreng van de opdrachtgever, worden vooraf besproken.

Hoe? De opdrachtgever en het PASO-team sluiten een *onderzoeksovereenkomst* af. Hierin worden alle modaliteiten van de samenwerking opgenomen. De kostprijs wordt berekend *zoals voor andere onderzoeksopdrachten*.

Wie? Publieke en wetenschappelijke gebruikers, andere niet-commerciële derden.

5. *On site werken op de PASO-bestanden*

Wat? De gebruiker krijgt *toegang tot één of meerdere geanonimiseerde PASO-(deel)bestanden* op een stand-alone pc. Deze pc wordt door het PASO-team 'on site', dit wil zeggen in een lokaal van de K.U.Leuven ter beschikking gesteld. Alle identificatievariabelen zijn uit het bestand verwijderd en de wegschrijfmogelijkheden zijn ingeperkt om misbruik te voorkomen. De PASO-databeheerder houdt toezicht op het gebruik van de on site faciliteit. De logbestanden van het gebruikte statistische pakket kunnen worden geïnspecteerd.

Hoe? De gebruiker dient de *bestandsovereenkomst* (type B), inclusief geheimhoudingsverklaring in te vullen en te ondertekenen. Verder wordt een *vergoeding overeengekomen* voor begeleiding van de gebruiker door een PASO-medewerker.

Wie? Wetenschappelijke gebruikers.

7.2 | De bestandsovereenkomst samenwerking PASO - overheid en wetenschappelijke instellingen

Bestandsovereenkomst type A (analyses door het PASO-team, rapportering door de gebruiker)

naam organisatie contractpartner, vertegenwoordigd door naam vertegenwoordiger, verder genoemd 'gebruiker'

EN

de PASO-promotoren van de K.U.Leuven, vertegenwoordigd door naam vertegenwoordiger, verder genoemd het PASO-team

zijn als volgt overeengekomen:

Artikel 1: Doel van verstrekking van het bestand

1.1 'gebruiker' deelt voorafgaandelijk aan deze overeenkomst volgende informatie aan het PASO-team mee:

- de naam van het onderzoeksproject;
- de gebruiker van het onderzoeksproject;
- een beknopte omschrijving van de inhoud en gebruikte onderzoeksmethoden van het project;
- een omschrijving van de planning van het project;
- de naam van de projectleider en van de onderzoekers;
- de beoogde resultaten;
- het publicatieplan voor het onderzoeksproject.

1.2 Een kopie van de overeenkomst van het onderzoeksproject wordt in de bijlage van dit contract opgenomen.

1.3 Het PASO-team bezorgt een kopie van deze overeenkomst, inclusief de bijlagen aan VIONA, Administratie Werkgelegenheid.

Artikel 2: Terbeschikkingstelling van de gegevens

2.1 Het PASO-team stelt de gegevens ter beschikking van gebruiker onder de vorm van tabellen.

2.2 De beslissing over de opportuniteit van het verstrekken van de gegevens en over de manier waarop dit gebeurt (on site werken of aanlevering tabellen) berust bij het PASO-team.

Artikel 3: Wijze van verstrekking

- 3.1 gebruiker stuurt een aanvraag tot verkrijgen van tabellen aan het PASO-team. Daarbij wordt precies opgesomd welke tabellen worden aangevraagd. Een gedetailleerde lijst van de te gebruiken variabelen (codeboek) wordt ter beschikking gesteld door het PASO-team.
- 3.2 Het PASO-team verstrekt de gegevens uiterlijk één maand na ontvangst van de concrete bestelling.
- 3.3 gebruiker betaalt het PASO-team het verschuldigde bedrag uiterlijk één maand nadat gebruiker de verstrekte gegevens en een factuur van het PASO-team ontvangen heeft.

Artikel 4: Auteursrechten

- 4.1 Het auteursrecht en alle andere intellectuele eigendomsrechten op de verstrekte gegevens berusten bij de PASO-promotoren van de K.U.Leuven.

Artikel 5: Kwaliteitsbewaking

- 5.1 Het PASO-team draagt zorg voor de kwaliteit van de verstrekte tabellen.
- 5.2 Het PASO-team meldt eventuele onvolkomenheden in de verstrekte tabellen en in de daarbij geleverde documentatie zo snel mogelijk aan gebruiker.
- 5.3 Het PASO-team herstelt, voor zover redelijkerwijs binnen het bereik van zijn mogelijkheden, fouten in de ter beschikking gestelde tabellen en de daarbij geleverde documentatie.

Artikel 6: Beperkingen op gebruik

- 6.1 gebruiker gebruikt de door het PASO-team verstrekte tabellen uitsluitend ten behoeve van bewerking in het kader van de uitvoering door gebruiker van het onder 1.1 en 1.2 van artikel 1 vermeldde onderzoeksproject.
- 6.2 Elke tabel waardoor de verstrekte gegevens direct ten grondslag liggen aan antwoorden van individuele ondernemingen of organisaties is uitgesloten.
- 6.3 gebruiker onthoudt zich ervan de gegevens, met uitzondering van (via artikel 8 goedgekeurde) onderzoeksresultaten bedoeld in artikel 1, geheel of gedeeltelijk aan derden door te leveren.

Artikel 7: Toegang tot verstrekte gegevens

- 7.1 gebruiker ondertekent de geheimhoudingsverklaring die integraal deel uitmaakt van deze overeenkomst en in bijlage is opgenomen.
- 7.2 gebruiker ziet erop toe dat uitsluitend door gebruiker aangewezen personen, belast met de uitvoering van het betrokken onderzoeksproject in de organisatie zelf kennis nemen van de verstrekte gegevens en daaruit af te leiden informatie, anders dan de goedgekeurde onderzoeksresultaten bedoeld in artikel 1.
- 7.3 gebruiker ziet erop toe dat deze personen zich verbinden tot geheimhouding van hetgeen hen bij de uitvoering van het onderzoek ter kennis is gekomen, op basis van deze geheimhoudingsverklaring.

Artikel 8: Rapportage op basis van verstrekte gegevens

- 8.1 Alvorens over te gaan tot publicatie wordt de uiteindelijke tekst voorgelegd aan het PASO-team. Hierbij wordt de voorstelling en interpretatie van de tabellen afgetoetst aan de reeds bestaande publicaties. Op basis hiervan kunnen in samenspraak voorstellen tot aanpassingen worden gedaan.
- 8.2 Het PASO-team voert de in 8.2 bedoelde toetsing binnen de maand uit. Indien het PASO-team niet binnen de maand reageert, wordt de vorm van rapportage van de onderzoeksresultaten geacht te zijn goedgekeurd. Indien dit nodig blijkt ter waarborging van geheimhouding of niet herkenbaarheid van individuele elementen, bepaalt het PASO-team de vorm aan waarin de resultaten voor presentatie in aanmerking komen.
- 8.3 Bij elke rapportering van onderzoeksresultaten vermeldt gebruiker duidelijk dat deze resultaten verkregen zijn door middel van bewerking van PASO-gegevens die zijn verzameld in opdracht van VIONA, Vlaamse Regering.
- 8.4 Van elke rapportage waarin onderzoeksresultaten op basis van de door het PASO-team verstrekte gegevens zijn opgenomen, stelt gebruiker twee exemplaren aan het PASO-team ter beschikking, waarvan één exemplaar is bestemd voor VIONA, Administratie Werkgelegenheid.
- 8.5 Publicaties op basis van PASO-data vermelden uitdrukkelijk dat de bespreking van de gegevens en de publicaties enkel de visie van de auteurs vertolken en niet de visie van het PASO-team, de VIONA of de Vlaamse Regering.
- 8.6 De rechten van de eerste publicatie op basis van de recentste PASO-golf zijn steeds voorbehouden aan VIONA - Vlaams minister van Werkgelegenheid en Toerisme.
- 8.7 Het publicatieplan op basis van PASO-gegevens zoals omschreven in het onderzoeksproject, zie artikel 1, wordt in nauwe samenspraak met het

PASO-team besproken. Om de resultaten van een volgende PASO-bevragingsgolf niet in het gedrang te brengen, kan het PASO-team eventueel beslissen dat in de twee maanden voorafgaand aan het opstarten van een (nieuwe) PASO-bevragingsronde geen resultaten mogen worden gepubliceerd.

Artikel 9: Duur beschikbaarheid van gegevens

9.1 Uiterlijk twaalf maanden na de beëindiging van het onderzoeksproject (zoals beschreven in artikel 1) waarvoor de gegevens zijn verstrekt, vernietigt gebruiker de verstrekte gegevens en alle op basis hiervan verkregen niet-gepubliceerde gegevens, tenzij gebruiker vóór het verstrijken van deze termijn een gemotiveerd verzoek om verlenging van deze termijn doet aan het PASO-team.

Deze vernietiging zal gebruiker schriftelijk aan het PASO-team bevestigen.

Artikel 10: Overdracht rechten

10.1 gebruiker kan zijn rechten uit deze overeenkomst nooit aan derden overdragen. Geen enkele vorm van onderaanneming van het wetenschappelijk onderzoek op de PASO-data zoals beschreven in artikel 1 is toegelaten.

Artikel 11: Duur en geldigheid van overeenkomst

11.1 Deze overeenkomst is aangegaan voor de duur van het onderzoeksproject zoals bedoeld in artikel 1.

11.2 Ook na beëindiging van de overeenkomst blijft het bepaalde in artikel 8 van kracht ten aanzien van de weergave van onderzoeksresultaten waarbij gebruik is gemaakt van door het PASO-team verstrekte gegevens.

11.3 Het PASO-team behoudt zich het recht voor om bij niet-nakoming van de bepalingen van deze overeenkomst onmiddellijk de overeenkomst te beëindigen.

Opgemaakt te Leuven op datum, in drie exemplaren, waarvan het PASO-team er een bezorgt aan de VIONA, Administratie Werkgelegenheid.

Voor gebruiker

naam

Voor het PASO-team

naam

voor Naam organisatie contractpartner

Bestandsovereenkomst type B (on site werken op de PASO-bestanden)

Naam organisatie contractpartner, vertegenwoordigd door **naam vertegenwoordiger**, verder genoemd 'gebruiker'

EN

de PASO-promotoren van de K.U.Leuven, vertegenwoordigd door **naam vertegenwoordiger**, verder genoemd het PASO-team

zijn als volgt overeengekomen:

Artikel 1: Doel van verstrekking van het bestand

1.1 'gebruiker' deelt voorafgaandelijk aan deze overeenkomst volgende informatie aan het PASO-team mee:

- de naam van het onderzoeksproject;
- de gebruiker van het onderzoeksproject;
- een beknopte omschrijving van de inhoud en gebruikte onderzoeksmethoden van het project;
- een omschrijving van de planning van het project;
- de naam van de projectleider en van de onderzoekers;
- de beoogde resultaten;
- het publicatieplan voor het onderzoeksproject.

1.2 Een kopie van de overeenkomst van het onderzoeksproject wordt in de bijlage van dit contract opgenomen.

1.3 Het PASO-team bezorgt een kopie van deze overeenkomst, inclusief de bijlagen aan VIONA, Administratie Werkgelegenheid.

Artikel 2: Terbeschikkingstelling van de gegevens

2.1 Het PASO-team stelt gegevens ter beschikking van gebruiker op het niveau van individuele elementen in de vorm van een geanonimiseerd bestand met de daarbij horende documentatie. Deze gegevens kunnen on site worden bewerkt. Hiervoor wordt er een stand-alone computer voorzien in een lokaal van het HIVA (Parkstraat 47). De PASO-databank mag onder geen enkele vorm van deze computer worden afgehaald.

2.2 De beslissing over de opportuniteit van het verstrekken van de gegevens en over de manier waarop dit gebeurt (on site of aanreiken van tabellen) berust bij het PASO-team.

Artikel 3: Wijze van verstrekking

- 3.1 opdrachtgever stuurt een aanvraag tot gebruik van het PASO-bestand. Daarbij wordt precies opgesomd welke variabelen zullen worden gebruikt. Een gedetailleerde lijst van de variabelen (codeboek) wordt ter beschikking gesteld door het PASO-team.

Artikel 4: Auteursrechten

- 4.1 Het auteursrecht en alle andere intellectuele eigendomsrechten op de verstrekte gegevens berusten bij de PASO-promotoren van de K.U.Leuven.

Artikel 5: Kwaliteitsbewaking

- 5.1 Het PASO-team draagt zorg voor de kwaliteit van de verstrekte gegevens.
- 5.2 Het PASO-team meldt eventuele onvolkomenheden in de verstrekte gegevens en in de daarbij geleverde documentatie zo snel mogelijk aan gebruiker.
- 5.3 Het PASO-team herstelt, voor zover redelijkerwijs binnen het bereik van zijn mogelijkheden, fouten in de ter beschikking gestelde gegevens en de daarbij geleverde documentatie.
- 5.4 opdrachtgever verbindt zich tot het gebruik van wetenschappelijke criteria voor de keuze van methoden om op de data te werken en tot het verstrekken van informatie over de gehanteerde procedures die de wetenschappelijke correctheid van de analyses moeten garanderen.

Artikel 6: Beperkingen op gebruik

- 6.1 opdrachtgever gebruikt de door het PASO-data uitsluitend ten behoeve van statistische analyse en bewerking in het kader van de uitvoering door opdrachtgever van het onder 1.1 en 1.2 van artikel 1 vermelde onderzoeksproject.
- 6.2 Elke toepassing waardoor de verstrekte gegevens direct ten grondslag liggen aan antwoorden van individuele ondernemingen of organisaties is uitgesloten.
- 6.3 opdrachtgever onthoudt zich ervan de gegevens als zodanig of in bewerkte vorm, met uitzondering van (via artikel 8 goedgekeurde) onderzoeksresultaten bedoeld in artikel 1, geheel of gedeeltelijk aan derden door te leveren, ter inzage te geven dan wel ter verwerking of ter gebruik te verstrekken.

Artikel 7: Toegang tot verstrekte gegevens

- 7.1 opdrachtgever ondertekent de geheimhoudingsverklaring die integraal deel uitmaakt van deze overeenkomst en in bijlage is opgenomen.

- 7.3 opdrachtgever ziet erop toe dat deze personen zich verbinden tot geheimhouding van hetgeen hen bij de uitvoering van het onderzoek ter kennis is gekomen, op basis van deze geheimhoudingsverklaring.
- 7.4 In het geval opdrachtgever de beschikking krijgt over een deelbestand zoals bepaald in artikel 2, draagt opdrachtgever zorg voor een adequate fysieke en logische beveiliging van de gegevens. Daartoe treft opdrachtgever ten minste voorzieningen ter voorkoming van:
- gebruik van de gegevens anders dan in deze overeenkomst is bepaald;
 - toegang tot of inzage in de verstrekte gegevens van andere dan in 7.2 bedoelde personen.
- 7.5 opdrachtgever toont het PASO-team op diens verzoek deze voorzieningen.

Artikel 8: Rapportage op basis van verstrekte gegevens

- 8.1 opdrachtgever geeft de op basis van de door het PASO-team verstrekte gegevens verkregen onderzoeksresultaten uitsluitend weer in de vorm van overzichten waaraan geen gegevens omtrent afzonderlijk herkenbare ondernemingen of organisaties kunnen worden ontleend.
- 8.2 opdrachtgever rapporteert geen onderzoeksresultaten verkregen op basis van de verstrekte gegevens in welke vorm ook - met inbegrip van voordrachten en congresbijdragen - dan na aan het PASO-team de onderzoeksresultaten te hebben voorgelegd ter toetsing op geheimhouding of niet-herkenbaarheid van individuele elementen uit het gegevensbestand.
- 8.3 Het PASO-team voert de in 8.2 bedoelde toetsing binnen de maand uit. Indien het PASO-team niet binnen de maand reageert, wordt de vorm van rapportage van de onderzoeksresultaten geacht te zijn goedgekeurd. Indien dit nodig blijkt ter waarborging van geheimhouding of niet herkenbaarheid van individuele elementen, bepaalt het PASO-team de vorm waarin de resultaten voor presentatie in aanmerking komen.
- 8.4 opdrachtgever rapporteert geen onderzoeksresultaten verkregen op basis van de verstrekte gegevens in welke vorm ook - met inbegrip van voordrachten en congresbijdragen - dan na aan het PASO-team de onderzoeksresultaten te hebben voorgelegd ter toetsing op de wetenschappelijke kwaliteitscriteria waarnaar wordt verwezen in paragraaf 5.4 van artikel 5.
- 8.5 Het PASO-team voert de in 8.4 bedoelde toetsing binnen de maand uit. Indien het PASO-team niet binnen de maand reageert, wordt de vorm van rapportage van de onderzoeksresultaten geacht te zijn goedgekeurd.

- 8.6 Bij elke rapportage van onderzoeksresultaten vermeldt opdrachtgever duidelijk dat deze resultaten verkregen zijn door middel van bewerking van PASO-gegevens die zijn verzameld in opdracht van VIONA, Vlaamse Regering.
- 8.7 Van elke rapportage waarin onderzoeksresultaten op basis van de door het PASO-team verstrekte gegevens zijn opgenomen, stelt opdrachtgever twee exemplaren aan het PASO-team ter beschikking, waarvan één exemplaar is bestemd voor VIONA, Administratie Werkgelegenheid.
- 8.8 Publicaties op basis van analyses met de PASO-databank vermelden uitdrukkelijk dat de bespreking van de gegevens en de publicaties enkel de visie van de auteurs vertolken en niet de visie van het PASO-team, de VIONA of de Vlaamse Regering.
- 8.9 De rechten van de eerste publicatie op basis van de recentste PASO-golf zijn steeds voorbehouden aan VIONA - Vlaams Minister van Werkgelegenheid en Toerisme.
- 8.10 Het publicatieplan op basis van PASO-gegevens zoals omschreven in het onderzoeksproject, zie artikel 1, wordt in nauwe samenspraak met het PASO-team besproken. Om de resultaten van een volgende PASO-bevragingsgolf niet in het gedrang te brengen, kan het PASO-team eventueel beslissen dat in de twee maanden voorafgaand aan het opstarten van een (nieuwe) PASO-bevragingsronde geen resultaten mogen worden gepubliceerd.

Artikel 9: Duur beschikbaarheid van gegevens

- 9.1 Uiterlijk twaalf maanden na de beëindiging van het onderzoeksproject (zoals beschreven in artikel 1) waarvoor de analyses worden gebruikt, vernietigt opdrachtgever de verstrekte gegevens en alle op basis hiervan verkregen niet-gepubliceerde gegevens, tenzij opdrachtgever vóór het verstrijken van deze termijn een gemotiveerd verzoek om verlenging van deze termijn doet aan het PASO-team.
- Deze vernietiging zal opdrachtgever schriftelijk aan het PASO-team bevestigen.

Artikel 10: Overdracht rechten

- 10.1 opdrachtgever kan zijn rechten uit deze overeenkomst nooit aan derden overdragen. Geen enkele vorm van onderaanneming van het wetenschappelijk onderzoek op de PASO-data zoals beschreven in artikel 1 is toegelaten.

Artikel 11: Duur en geldigheid van overeenkomst

11.1 Deze overeenkomst is aangegaan voor de duur van het onderzoeksproject zoals bedoeld in artikel 1.

11.2 Ook na beëindiging van de overeenkomst blijft het bepaalde in artikel 8 van kracht ten aanzien van de weergave van onderzoeksresultaten waarbij gebruik is gemaakt van door het PASO-team verstrekte gegevens.

11.3 Het PASO-team behoudt zich het recht voor om bij niet-nakoming van de bepalingen van deze overeenkomst onmiddellijk de overeenkomst te beëindigen.

Opgemaakt te Leuven op **datum**,
in drie exemplaren, waarvan het PASO-team er een bezorgt aan de VIONA,
Administratie Werkgelegenheid.

Voor **gebruiker**

naam

Voor het PASO-team

naam

voor **Naam organisatie contractpartner**

7.3 | Bijlage bij de bestandsovereenkomst Model van geheimhoudingsverklaring PASO

Model van de verklaring van gebruiker tot naleving van de voorwaarden verbonden aan de bestandslevering.

Verklaring

A. Ondergetekende, naam, functie

wonende te adres

verklaart zijn/haar werkzaamheden betreffende het/de in artikel 1 van de overeenkomst genoemde onderzoeksproject(en) te zullen verrichten onder de volgende voorwaarden:

1. Hij/zij verplicht zich tot geheimhouding van elk gegeven omtrent afzonderlijke ondernemingen of organisaties dat hem/haar bij de uitvoering van de analyses op de PASO-data in het kader van dit onderzoeksproject bekend wordt.
2. Hij/zij zal van al hetgeen hem/haar bij deze analyses ter kennis komt niet openbaren dan met voorafgaande schriftelijke toestemming van het PASO-team en volgens de modaliteiten bepaald in artikel 8 van de overeenkomst.
3. Ingeval van discussie omtrent de uitleg van bovengenoemde voorwaarden beslist het PASO-team. Deze verklaring blijft van kracht, ook na beëindiging van bovenbedoelde werkzaamheden.

.....

(Plaats)

.....

(Datum)

.....

(Handtekening)

B. Ondergetekende, naam, functie

van naam instelling, organisatie

aanvaardt hierbij de verantwoordelijkheid voor de naleving van de hierboven vermelde voorwaarden waaronder

de heer/mevrouw naam, medewerk(st)er van bovengenoemde instelling zijn/haar in deze verklaring omschreven werkzaamheden zal verrichten.

.....

(Plaats)

.....

(Datum)

.....

(Handtekening)

Hoofdstuk 8

Studiedag 'Het Panel Survey of Organisations in Flanders in een breder perspectief'

Seminarie donderdag 4 maart 2004 - Hotel Errera, Brussel

Een laatste opdracht in het kader van het project databeheer was het organiseren van een gespecialiseerde workshop om PASO te vergelijken met andere gelijkaardige projecten en om een evaluatie van de verschillende doelstellingen na drie jaar PASO mogelijk te maken.

Hieronder volgt een beschrijving en het programma van de workshop aan de hand van de bekendmakingsfolder. De verschillende sprekersbijdragen en de aanwezigheidslijst zijn opgenomen op de cd-rom die integraal deel uitmaakt van dit rapport.

8.1 | Inleiding

PASO is een longitudinale panelstudie van de vraagzijde van de arbeidsmarkt in Vlaanderen. In dit in 2000 opgestarte project wordt jaarlijks een staal van ongeveer 10 000 organisaties bevraagd. In 2001 werd de PASO-enquête, gefinancierd met het VIONA-programma, en de op te starten PSIC-bevraging, gefinancierd in het kader van het Steunpunt Ondernemerschap, Ondernemingen en Innovatie (STOOI), geïntegreerd. Het VIONA-onderzoeksproject loopt tot en met 30 november 2004. De financiering in het kader van het STOOI loopt ten einde in 2006. De eerste bevragingronde ging op 18 april 2002 van start, de tweede golf in maart 2003, de derde in maart 2004.

Het doel van dit onderzoek, zoals destijds geformuleerd, was tweeledig:

- beantwoorden aan de steeds toenemende behoefte aan accurate informatie over evoluties aan de vraagzijde van de arbeidsmarkt;
- het tegengaan van bevragingmoeheid (overbevraging) bij organisaties.

Het belang van de eerste doelstelling kan nauwelijks worden overschat. De output van het project is omvangrijk en de vragen naar bijkomende analyses en data komen uit zeer uiteenlopende hoek. Op dit punt zijn er nog veel onbenutte mogelijkheden.

De tweede doelstelling moet op korte termijn geherdefinieerd worden. ‘Overbevraging’ heeft namelijk uiteenlopende oorzaken en er zijn veel actoren op de markt (Europese, federale, regionale, academische, commerciële, etc.) die buiten de invloed van het PASO-project staan. De ervaringen van de afgelopen jaren hebben echter wel geleerd dat PASO als basis kan fungeren voor zeer uiteenlopende samenwerkingsverbanden.

8.2 | Doelstellingen van het seminarie

Het seminarie speelt in op twee vragen.

Vooreerst is er de kwestie van de precieze plaats van een organisatiepanel in het statistieklandschap van een regio zoals Vlaanderen. Uit internationaal vergelijkend onderzoek blijkt dat het systematisch en longitudinaal verzamelen van microdata op niveau van vestigingen uit alle sectoren van de economie als het sluitstuk kan worden beschouwd van het statistieksysteem van een land of regio. Dergelijk vraagzijdeonderzoek vervult immers een essentiële complementaire rol ten aanzien van administratieve gegevens waarover de overheid beschikt. Deze rol vervullen stelt echter hoge eisen op vlak van het survey design (o.m. het longitudinale karakter en de representativiteit). Een afstemming met andere pijlers van het statistiekbeleid, zoals de koppeling aan administratieve bestanden en een beleid inzake administratieve vereenvoudiging voor bedrijven, is daarnaast nodig.

Ten tweede stelt zich de vraag hoe PASO kan bijdragen aan de oriëntering van het beleid. In functie van een adequaat economisch en arbeidsmarktbeleid is een aanvulling van macro-economische data met micro- en meso-economische gegevens cruciaal. Zo kunnen verbanden worden onderzocht tussen initiatieven op vlak van overheidsbeleid en ontwikkelingen in het beleid van organisaties, en tussen de beleidskeuzes van organisaties en arbeidsmarktontwikkelingen. Essentieel is dat de data die worden gegenereerd beleidsrelevant zijn en optimaal worden gevaloriseerd. De geloofwaardigheid én gebruiksmogelijkheden op de langere termijn staan of vallen evenwel met de kwaliteit van de wetenschappelijke onderbouw. Een organisatiepanel moet het vertrouwen van de nationale en internationale wetenschappelijke community blijven verdienen, wat heel wat inspanningen vergt. Beide dimensies respecteren, beleidsrelevantie en wetenschappelijk karakter, betekent ook een compromis zoeken tussen continuïteit en flexibiliteit.

De inleidingen en discussies van het seminarie kunnen bijdragen tot een beter inzicht in de mogelijkheden voor optimalisering van het design, de organisatie én de valorisatie van dit vraagzijdeonderzoek. Dit kan leiden tot een aantal bijsturingen, zodat het project ook in de toekomst op een wetenschappelijk verantwoorde manier aan de toenemende informatiebehoefte van steeds meer verschillende stakeholders kan tegemoet komen.

Het seminarie wordt georganiseerd door VIONA en de PASO-equipe. Van bij het begin kon PASO terugvallen op de ervaring van twee gerenommeerde buitenlandse panels, namelijk het IAB-panel uit Duitsland en het Vraagzijde-panel van OSA uit Nederland. Beide instituten werden bereid gevonden om een bijdrage te leveren over hun ervaring met een aantal cruciale aspecten van organisatiepanels. Ook een vertegenwoordiger van één van de federale opdrachtgevers van het Duitse panel zal een inleiding verzorgen.

8.3 | Doelpubliek

Het seminarie is een gesloten workshop waarop vertegenwoordigers worden uitgenodigd van de Vlaamse administraties werkgelegenheid, economie en statistiek, de relevante Steunpunten en de leden van de VIONA-stuurgroep en de PASO-visiegroep.

8.4 | Datum en plaats

Donderdag 4 maart 2004 – Hotel Errera, Koningstraat 14, 1000 Brussel (vlakbij het Centraal station).

8.5 | Agenda

De presentaties van deze studiedag zijn als PDF of powerpoint presentaties terug te vinden op: <http://www.paso.be/seminarie2004.htm>.

Dagvoorzitter: Geert Sterck, VIONA-coördinator

09:30 – 10:00	Inleiding <i>Monique Ramioul, PASO-promotor</i> PASO in het internationaal landschap van organisatiepanels
10:00 – 12:45	Organisatiepanels als sluitstuk voor een regionaal statistieksysteem
10:00 – 10:30	<i>Marcel Kerkhofs, OSA - Nederland</i> De evolutie van het OSA vraagzijdepanel-design en nieuwe uitdagingen
10:30 – 11:00	<i>Markus Promberger - IAB - Duitsland</i> Panelanalyse: welke meerwaarde tegen welke prijs? (Engelstalige bijdrage)
11:00 – 11:15	Koffiepauze
11:15 – 11:45	<i>André Zijdenbos - Survey@ - Nederland</i> Responsmaximalisatietechnieken bij organisatieonderzoek

	11:45 - 12:15	<i>Luc Sels, PASO-promotor</i> Lessen uit drie golven PASO en te maken keuzes
	12:15 - 12:45	Discussie
12:45 - 14:00	Lunch	
14:00 - 16:45	PASO vanuit beleidsperspectief	
	14:00 - 14:30	<i>Thierry Vergeynst - Administratie Planning en Statistiek</i> Het Vlaams statistiekbeleid in de toekomst
	14:30 - 14:50	<i>Peter van der Hallen, Steunpunt WAV</i> De impact van statistiek op het arbeidsmarktbeleid: ervaringen van 10 jaar Steunpunt WAV
	14:50 - 15:10	<i>Mark Andries - VOKA</i> Welke indicatoren willen organisaties?
	15:10 - 15:25	Koffiepauze
	15:25 - 15:45	<i>Freek De Witte, Administratie Werkgelegenheid – Economie</i> PASO en het Vlaams beleid inzake Economie en Werkgelegenheid
	15:45 - 16:15	<i>Geert Van Hootegem, PASO-promotor</i> Dimensies in organisatievernieuwing en hoe ze te onderzoeken
	16:15 - 16:45	Discussie
16:45	Conclusies Fons Leroy, Kabinet van Vlaams minister van Werkgelegenheid en Toerisme	