

Benaming van de instelling

Federale Overheidsdienst Economie, K.M.O., Middenstand en Energie (Algemene Directie Statistiek en Economische informatie)

Benaming van het project

Microcensus 2006

Acroniem

Microcensus 2006

Beschrijving van het project en situering binnen het kader van de instelling

De FOD Economie, K.M.O., Middenstand en Energie - Algemene Directie Statistiek en Economische informatie (ex Nationaal Instituut voor de Statistiek), hierna genoemd Directie Statistiek - organiseert al sinds 1846 zogenaamde « Algemene Volkstellingen » die naast hun functie voor het bepalen van de bevolkingsomvang en samenstelling vooral een wetenschappelijk karakter hebben. De Directie Statistiek verwerkt en verspreid de resultaten in de vorm van tabellen en publicaties maar ook als bronmateriaal waarop onderzoeksinstellingen van de overheid, de academische wereld en andere belanghebbenden beroep kunnen doen bij de studie van een brede waaier aan terreinen van socio-economische relevantie. Tot op vandaag is de census de belangrijkste bron van socio-economische informatie in ons land, en voor een hele reeks gegevens uit de tellingen is er geen andere bron voorhanden.

Voor de toekomst kondigt zich echter een grote omwenteling aan: bij beslissing op de zitting van 4 mei 2000 heeft de ministerraad erin toegestemd om op 1 oktober 2001 een Algemene Socio-economische enquête te organiseren, maar tegelijk aan de minister van Economie de opdracht gegeven om een voorstel uit te werken tot aanpassing van de wet van 4 juli 1962 betreffende de openbare statistiek teneinde de verplichting op te heffen om de tien jaar een klassieke exhaustieve telling te houden, en alternatieven te ontwikkelen om deze gegevens te verzamelen zonder beroep te moeten doen op een exhaustieve enquête, meer bepaald door bestaande of op te richten administratieve gegevensbanken te gebruiken.

Het organiseren van het studiewerk rond de alternatieven voor de klassieke telling werd toevertrouwd aan de Hoge Raad voor de Statistiek en inmiddels verrichten een vijftal universitaire onderzoeksteams al sinds twee jaar specifiek onderzoek op diverse deelterreinen van deze problematiek.

Naast dit studiewerk is het echter van het grootste belang dat met de praktische uitbouw en invulling van de statistische databanken een aanvang wordt genomen, zodat zowel valorisering van vroeger en lopend onderzoekswerk als de continuïteit van de census naar 2011 toe gewaarborgd blijven. Daarom wordt er voorgesteld om een « Administratieve Proeftelling 2006 » te organiseren, die enerzijds moet toelaten meer technische expertise en logistiek op te bouwen rond de overgang naar administratieve tellingen, maar anderzijds ook de ambitie heeft precies in de helft van de klassieke tienjaarlijkse tijdssprong van de census, al een eerste ronde van bruikbare socio-economische gegevens van administratieve oorsprong op te leveren. Deze proeftelling zal gebeuren op basis van een steekproef van 20% van de referentiepersonen en de leden die tot hun huishouden behoren. Op die manier kan het volledige Belgische grondgebied met de nodige resolutie bereikt worden. Door raadpleging van de grote administratieve databanken (Rijksregister, Kruispuntdatabank Sociale Zekerheid, Kruispuntdatabank Ondernemingen, Kadastrale Databank van de Algemene Administratie van de Patrimoniumdocumentatie, de databanken rond onderwijs en afgeleverde diploma's, ...) zal een maximaal aantal van de klassieke censusdomeinen gedekt worden.

De timing in 2006 is om meerdere redenen opportuun:

- het tijds punt precies halverwege de Enquête 2001 en de geplande opvolging in 2011: gezien een van de doelstellingen rond de overgang naar administratieve tellingen ook het verkorten van de tussentijdse cyclus was, is dit tegelijk een aanzet tot een mogelijk vijfjaarlijks tellingssysteem.
- Er is voorafgaand aan de administratieve telling voldoende tijd om onderzoekswerk dat terzake verder loopt te incorporeren in dit project.
- Tegelijk laat de voorbereidingstijd van zowat een jaar toe om de cruciale technische en organisatorische stappen te nemen om de proeftelling te doen slagen: het opzetten binnen de Directie Statistiek van de noodzakelijke statistische databanken; het treffen van regelingen met de beheerders van de administratieve databanken, enz.
- Het nemen van een vijfjaarlijkse stap t.a.v. de vorige telling draagt bij tot het praktische gebruik en de referentiewaarde van de resultaten.

Het project dat hier wordt voorgesteld past in een bredere visie rond de ontwikkeling van het toekomstig Belgisch statistisch apparaat. De grotere rol die wordt toebedeeld aan de administratieve databanken zal ook bijkomende eisen stellen voor de inzameling van aanvullende grootschalige steekproefenquêtes. Een goede exploitatie van administratieve data laat toe om de belangrijke demografische en sociale evoluties in kaart te brengen. De belangrijkste voordelen van het gebruik van de administratieve gegevens zijn hierbij de grote precisie en het exhaustieve karakter. Tegelijk heeft een dergelijke statistiek ook belangrijke beperkingen. De data weerspiegelen een zuiver administratieve werkelijkheid die niet steeds in overeenstemming is met de sociale realiteit en laten niet toe om dieper te graven naar causale verbanden. Vandaar het toegenomen belang van de verdere uitbouw en ondersteuning van grote nationale steekproefenquêtes.

Naast de bestaande enquêtes die al lang door de Directie Statistiek worden gevoerd, werd ook gestart met een totaal nieuwe panelstudie: SILC (Statistics on Income and Living Conditions). De Directie Statistiek nam ook het initiatief om de gemeenschappen en gewesten te betrekken bij een voorstudie voor het opzetten van een mogelijke nieuwe demografische enquête: GGS (Gender and Generation Survey). Voor de uitwerking van de SILC-enquête werd door de Directie Statistiek al aanzienlijke inspanningen geleverd door investeringen in mensen en materiaal.

Het behoort tot de klassieke gewoonten bij steekproefonderzoek om de resultaten aan grootschalige representatieve bronnen te toetsen. Tot op heden speelt de "volkstelling" of "Socio-Economische Enquête" deze rol. De nieuwe wijze van gegevensverzameling zal de referentiewaarde van de census echter ook in vraag stellen. Daarom is er dringend bijkomend studiewerk nodig, gebaseerd op gegevens die al volgens het nieuwe systeem werden verzameld. Indien ook de administratieve census in de toekomst een vergelijkbare centrale rol wil blijven spelen in het Belgische statistische apparaat is het van belang tijdig met het verkennen van de nieuwe situatie te beginnen.

Tot slot verwijzen we naar de Europese dimensie in de verzameling van socio-economische statistiek. Deze dimensie wint nog steeds aan belang. Verscheidene grote steekproefenquêtes kaderen in bredere Europese projecten of worden voor rapportering aan de Europese statistiekinstellingen gebruikt. De arbeidskrachtenenquête (Labour Force Survey) is hiervan wellicht het meest klassieke voorbeeld. Maar de Europese dimensie is ook voor de overgang naar administratieve tellingen van belang. Een aantal landen, vooral in Scandinavië, maken nu al succesvol gebruik van administratieve databanken voor statistische doeleinden. Bovendien bereiden vele anderen zich op een dergelijke overgang voor of onderzoeken de mogelijkheden van deze formule. Binnen deze landen werd al een brede variëteit aan mogelijkheden onderzocht, uitgetest en geïmplementeerd, en werd de statistische infrastructuur op diverse wijzen aangepast om deze verzamelwijze van informatie mogelijk te maken en ook te laten renderen voor andere functies van openbaar belang. Uiteraard doet dit onderzoeksvoorstel eveneens beroep op deze expertise, en moet de integratie van resultaten afkomstig uit databanken en enquêtes ook in dit Europese perspectief bestudeerd worden.

Uit te voeren opdracht:

a. door de onderzoeksploeg

De taken van de onderzoeksploeg vallen uiteen in een viertal domeinen:

- 1) Het opstellen van een lijst van variabelen en hun specificaties (tijdstip, formaat, ...) die bij de diverse administratieve beheerders van databanken aangevraagd zullen worden, samen met de meta-informatie nodig om deze informatie te exploiteren en op hun validiteit en betrouwbaarheid te controleren. Dit behelst ook het nemen van de nodige maatregelen bij de data-aanvraag en de steekproeftrekking om het harmoniseren van de diverse steekproeven uit de respectieve databanken voor te bereiden.
- 2) De Directie Statistiek assisteren en adviseren bij de exploitatie van deze gegevens. Dit behelst alle deelactiviteiten vanaf de transactie van de gegevens in aangeleverde vorm tot en met de productie van afgewerkte statistieken: het selecteren van definitief te gebruiken bronnen en variabelen, het koppelen van informatie van verschillende oorsprong, waarborgen van de vertrouwelijkheid, controles en kwaliteitsanalyse (o.a. volledigheid, dubbeltelling, interne consistentie, tijdsindex van opmeting, ...). Ook de procedures voor het opkuisen van de gegevens worden door de onderzoeksploeg vastgelegd. Dit alles moet gebeuren rekening houdende met de specificiteit van de informatie afkomstig uit administratieve statistische bronnen: d.w.z. met het feit dat het gaat om gefragmenteerde informatie, met diverse tijdsindex en een verzamelwijze met het oog op administratieve eerder dan statistische doeleinden.
- 3) Het systematisch rapporteren en in kaart brengen van het volledige proces, de datastromen, de kwaliteitsanalyse en alle aspecten die van belang zijn voor de dekkingsgraad, de betrouwbaarheid, de actualiteitsgraad, de vergelijkbaarheid en andere kwaliteitsaspecten die de administratieve verzamelwijze van socio-economische informatie met zich meebrengt, alsook het formuleren van aanbevelingen om de mogelijke negatieve effecten op de kwaliteit van het eindresultaat te minimaliseren of tegen te gaan.
- 4) Een wetenschappelijke publicatie ter valorisering en verspreiding van de resultaten van de administratieve proeftelling 2006 realiseren, echter rekening houdende met de onzekere kwaliteit van het eindproduct. Bij voldoende hoge kwaliteit wordt aan de formule van een zeer beperkte « Volkstellingsmonografie » gedacht, met aandacht voor de socio-economische verschuivingen sinds 2001. Indien de kwaliteit en vergelijkbaarheid met 2001 onvoldoende gewaarborgd is zal het eindproduct eerder een kritische bronnenstudie zijn, waarbij de nieuwe gegevens voornamelijk met andere bronnen binnen hetzelfde tijdsbestek vergeleken worden.
- 5) De ploeg zal nagaan in hoeverre de gegevens van het project in een Europese en/of internationale context kunnen worden vergeleken en geharmoniseerd. De ploeg stelt hiervoor een rapport op.

b. door de openbare instelling

1. voorafgaand aan het eigenlijke onderzoek

Voorafgaand aan het eigenlijke onderzoek zullen de Directie Statistiek en de Directie Informatie- en Communicatietechnologie (ICT) de nodige voorbereidingen treffen rond de transactie van gegevensbestanden met de voornaamste beheerders van relevante administratieve bronnen. Het gaat hier o.m. om het Kadaster (FOD-Financiën); het Rijksregister, de Kruispuntbank Sociale Zekerheid en de gewestelijke instanties die instaan voor de onderwijsstatistieken. Met deze instellingen bestaan doorgaans al goede samenwerkingsverbanden en worden ook nu al administratieve

gegevens uitgewisseld. Het is echter van belang in deze samenwerkingsverbanden structureel het element van de administratieve census binnen te brengen, en hieraan het nodige overleg te koppelen. Daarom zullen de betrokken instanties uitgenodigd worden om de problematiek rond gegevensuitwisseling voor statistische doeleinden te bespreken en voor een samenwerking rond censusstatistiek op langere termijn de nodige structuren uit te werken.

De Directie Statistiek verbindt er zich toe een verslag op te maken rond de stand van zaken m.b.t. de databanken die in voorbereiding zijn om in de toekomst longitudinaal de bevolking, het onderwijspeil, het woningenpark enz. op te volgen (Project DEWOLO). Daarin komen ook de activiteiten binnen de Directie Statistiek en ICT aan bod die al in het verleden werden ondernomen om koppeling en exploitatie van alternatieve bronnen voor te bereiden of te evalueren (Socio-economische informatie uit Kruispuntbank Sociale Zekerheid, Arbeidskrachtenenquête, adresinformatie uit Rijksregister, Kadaster, enz...).

Ten slotte neemt de Directie Statistiek de nodige administratieve en logistieke stappen om ervoor te zorgen dat het nodige personeel aanwezig is en vrijgesteld wordt om de verdere uitbouw van deze databanken onmiddellijk te kunnen opstarten wanneer het project aanvangt. Dit vraagt aan de Directie Statistiek en ICT samen ongeveer 5 FTE aan gekwalificeerd personeel. Een interne projectaanvraag P890 m.b.t. deze activiteiten is al ingediend.

2. in de loop van het project

In de loop van het project zal de indiener verantwoordelijk zijn voor:

- het organiseren van de gesprekken met de respectieve beheerders van administratieve databanken, afvaardigen van betrokkenen binnen de instelling en van de onderzoeksploeg;
- het behandelen van de aanvragen voor informatie-uitwisseling en de technische aspecten rond uitwisselingsprotocols en formaten alsook de opvolging van de dossiers rond aanvragen;
- het beheren van de administratieve informatie voor statistisch gebruik, het maken van de nodige koppelingen en het toeleveren van geanonimiseerde steekproeven voor evaluatie door de onderzoeksploeg;
- het publiceren van het eindresultaat van de werkzaamheden, in een vorm die zoals hoger aangegeven in functie van de kwaliteit van de informatie enigszins kan variëren, en dit eindresultaat verspreiden volgens de gebruikelijke kanalen (infoshops, internet...).

1. Op te leveren eindproduct:

a. door de onderzoeksploeg

Voor de onderzoeksploeg zal het eindproduct driedelig zijn:

- een *rapport* dat volgens de formule van action-research het proces in kaart brengt, bespreekt en alternatieven en oplossingen voorstelt. Dit zal voor de openbare instelling van nut zijn als logboek voor de activiteiten rond de longitudinale statistische databanken en hun optimale uitbouw en ontwikkeling.

- een *statistische databank* die zo getrouw mogelijk het formaat volgt van de klassieke volkstellingen als foto van de socio-economische realiteit. De mate waarin deze eerste «administratieve foto» ook voor verdere distributie en wetenschappelijke exploitatie geschikt is zal nader bepaald moeten worden, maar met de mogelijkheid ervan wordt in elk geval van bij de startrekening gehouden. Deze databank is uiteraard volledig gelokaliseerd en in beheer van de indienende overheidsinstelling. Vanzelfsprekend gelden voor de onderzoeksteams binnen en buiten de instelling de klassieke beperkingen en clausules rond het garanderen van de privacy.
- Een *gepubliceerd* wetenschappelijk eindproduct dat naargelang de bereikte kwaliteit varieert tussen een *mini-censusmonografie* en een *kritisch overzicht* van de actuele staat van socio-economische indicatoren zoals deze uit diverse bronnen naar voren komt. Daarin zullen ook de andere steekproefenquêtes van de Directie Statistiek betrokken worden.

b. voor de openbare instelling

De openbare instelling verbindt er zich toe als eindproduct een exploiteerbare databank uit te bouwen en bij te houden rond de gegevens die in het kader van dit project verzameld worden, en daarbij dezelfde regels en gebruiken te respecteren die klassiek van toepassing zijn voor censusinformatie. Er wordt wel met de mogelijkheid rekening gehouden bijkomende beperkingen in acht te nemen voor de verspreiding van resultaten van onzekere kwaliteit.

Voor de FOD Economie, K.M.O., Middenstand en Energie (Algemene Directie Statistiek en Economische Informatie) vormt het eindresultaat in de vorm van een databank en gepubliceerde resultaten een belangrijke meerwaarde. Er moet echter de nadruk op gelegd worden dat het minstens even belangrijk is bij de diepgaande omwentelingen die zich aandienen rond de verzamelwijze van de socio-economische informatie tijdig te kunnen anticiperen op de technische en inhoudelijke uitdagingen. Het zal immers gaan om informatie die moet samengebouwd worden uit elementen die vanuit verschillende tijdsdimensies, met een verschillende invalshoek en met onderscheiden bedoelingen werden verzameld. Bij een overgang van deze grootteorde en complexiteit is het essentieel een beroep te kunnen doen op een gekwalificeerde onderzoeksploeg, die vanuit methodologisch oogpunt competent is om de nieuwe uitdagingen aan te pakken.

Valorisatie van het eindproduct:

a. voor de openbare instelling: integratie van het eindproduct in haar eigen werking

Artikel 21 van de wet van 2 januari 2001 maakt een einde aan de verplichte tienjaarlijkse algemene tellingen en voorziet in plaats daarvan in de uitbouw en het updaten door de Directie Statistiek van statistische databanken die vanuit administratieve bronnen gevoed dienen te worden. De minister van Economie vraagt de Algemene Directie Statistiek en Economische informatie de werkzaamheden hierrond verder uit te bouwen.

Dit onderzoeksvorstel is er precies op gericht methodologische en inhoudelijke vooruitgang te boeken in diverse dossiers die hieraan verbonden zijn. Bovendien wordt er al sinds een tweetal jaren voorbereidend onderzoek in deze materie verricht. De lange traditie betreffende Belgische sociaal-economische statistiek op basis van de volkstellingen, de inspanningen die nu al op dit vlak gebeuren en het belang dat de bevoegde minister hecht aan deze opdracht, garanderen de integratie van het eindproduct in de werking van de FOD-Economie.

Bovendien verwacht de Directie Statistiek van dit project een positieve wisselwerking met de grote steekproefenquêtes die door de instelling worden georganiseerd, en past dit voorstel in een bredere visie rond de uitbouw van het Belgisch statistisch systeem. Dit systeem is erop gericht zowel de Belgische sociaal-economische statistiek als de rapportering t.a.v. de Europese statistiekinstellingen longitudinaal te verzekeren en de kwaliteit ervan te bewaken.

b. voor het “publiek”: toegangsvoorwaarden en modaliteiten voor wetenschappelijke gebruikers en voor het grote publiek

De valorisatie van het eindresultaat voor het publiek - zowel wetenschappelijke gebruikers als het bredere publiek - op langere termijn staat vast; de toekomst van de volkstelling (of Socio-Economische Enquête) heeft immers baat bij dit onderzoeksvoorstel. Ook nu de resultaten van de Enquête 2001 geleidelijk beschikbaar komen voor verspreiding en voor verder onderzoek blijkt weer hoe groot de belangstelling van de media en de onderzoekswereld is voor deze unieke bron van informatie.

Zoals hoger aangegeven is het op dit ogenblik minder zeker of ook de resultaten van de hier voorgestelde eerste proefronde volgens de administratieve formule al van dezelfde waarde zullen zijn. Daarop wordt echter ingespeeld door de formule van het gepubliceerde eindresultaat te laten afhangen van de kwaliteit van de gegevens, zodat in elk geval een waardevol en wetenschappelijk verantwoord eindresultaat wordt bereikt.

Inzake de toegangsvoorwaarden en modaliteiten voor diverse groepen van gebruikers van censusgegevens heeft de Directie Statistiek een beproefde werkwijze. Daarin komt uiteraard de bescherming van de persoonlijke levenssfeer van de individuele burger op de eerste plaats. Dat betekent concreet dat gegevens die rechtstreeks of door kruising tot individuele herkenning zouden kunnen leiden onder geen beding worden doorgegeven, en dat de verspreiding van informatie gebeurt op het hoogste aggregatieniveau waarop de informatie nog nuttig is voor de gebruiker. Bovendien wordt er systematisch met steekproeven gewerkt i.p.v. met exhaustieve populaties. Slechts de Commissie voor de bescherming van de persoonlijke levenssfeer (“privacycommissie”) is bevoegd afwijkingen hierop toe te staan.

Timing, planning van uit te voeren onderzoekstaken en toegekend budget

a. timing van activiteiten:

Zie verder “Tijdsbalk deeltaken”. Er moet benadrukt worden dat deze timing afhangt van het toeleveren van gegevens door derden. De uitvoering van de deeltaken kan bijgevolg enigszins variëren.

b. planning van activiteiten: fasering van taken

1. voor de onderzoeksploeg

Indeling in vier deels overlappende fasen, tentatieve timing gezien de afhankelijkheid van derden voor datatoelevering.

Fase 1: (september 2005 – juni 2006)

- overleg met beheerders administratieve databanken: dit overleg omvat een reeks werkvergaderingen en mikt uiteindelijk op de oprichting van vaste structuren rond datatransactie in het kader van de administratieve tellingen
- de formulering van de precieze data-aanvraag bij de diverse administratieve beheerders van databanken: bepalen van de variabelen en hun specificaties (tijdstip, formaat, trekkingswijze van de steekproef,...), samen met de meta-informatie nodig om deze informatie te exploiteren en op haar validiteit en betrouwbaarheid te controleren

- in kaart brengen en inbouwen van lopend en afgerond onderzoek in deze materie

Fase 2: (maart 2006 – oktober 2006)

- behandelen van proefbestanden op kwaliteitsaspecten: 2 doelstellingen:
- (1) data-aanvraag verder preciseren en zo nodig bijsturen (uit meerdere bronnen de kwalitatief beste selecteren, technische aspecten rond het dataformaat uitklaren en/of harmonisering met het oog op koppeling met andere bronnen optimaliseren)
- (2) voorbereiding voor de uiteindelijke bewerking van de administratieve censusbestanden (fase 3) : voorstudie rond het maken van een statistische dwarsdoorsnede in de administratieve databank en inschatting van de problemen; effecten van de nieuwe observatiewijze
- komen hier o.m. aan bod: alle deelactiviteiten vanaf de transactie van de gegevens in aangeleverde vorm tot en met de productie van statistieken, het koppelen van informatie van verschillende oorsprong, waarborgen van de vertrouwelijkheid, controles en kwaliteitsanalyse (o.a. volledigheid, dubbeltelling, interne consistentie, tijdsindex van opmeting, ...), vastleggen van de procedures voor het opkuisen van de gegevens.

Fase 3: (november 2006 – oktober 2007)

- ontvangst administratieve censusbestanden ref.datum 1/1/2006
- exploitatie van de bestanden in meerdere fasen (met logistieke en technische ondersteuning door de Directie Statistiek en I.C.T.)
- data-cleaning, harmonisering van anonieme identificator die koppeling toelaat zonder herkenning, consistentie en controle binnen de bron
- koppeling, checks op consistentie tussen bronnen, behandeling van missing values, documenteren van variabelen en codes (meta-informatie)
- creatie van afgewerkte, gedocumenteerde en exploiteerbare censusbestanden
- opstellen rapport "Datakwaliteit administratieve censusgegevens"

Fase 4: (juli 2007 – augustus 2008)

- aanmaak en beschrijving van de censusresultaten 2006
- aanmaak van kruisingen, tabellen, kaarten, socio-economische indicatoren
- wetenschappelijk eindresultaat: de Belgische socio-economische situatie in 2006 (precieze formule i.f.v. datakwaliteit)
- procesmatig eindresultaat: naar een administratief systeem van gegevensverzameling en productie van socio-economische statistiek in België.

2. voor de openbare instelling

Fase 1: (september 2005 – juni 2006)

- (vernieuwde) contactname met beheerders administratieve databanken om het overleg rond datatransactie t.b.v. administratieve tellingen op te starten
- aanduiden van verantwoordelijken voor diverse deeltaken en de opvolging van de dossiers rond gegevensuitwisseling
- overlegfase: afvaardigen van verantwoordelijken Directie Statistiek en I.C.T. op de werkvergaderingen, bemannen van vaste structuren rond datatransactie in het kader van de administratieve tellingen
- precisering data-aanvraag bij de diverse administratieve beheerders van databanken: technische specificaties (formaat, ...), en praktische aspecten van de transactie

Fase 2: (maart 2006 – oktober 2006)

- implementatie en uitbouw van de statistische databank, verdere evolutie van het project "DEWOLO": Informatica - infrastructuur, inhoudelijke structuur van de databank, gegevensmodellen, queries, objecten en hun relaties (geografische eenheden, woningen, huishoudens, gezinskernen, ...)
- databank vullen met bestanden (administratieve proefbestanden, censusinformatie, database rond adressen en hun notering) en exploiteren (trekken van steekproeven en maken van koppelingen voor studie door de onderzoeksploeg, procedures voor de actualisering van informatie binnen de databank)

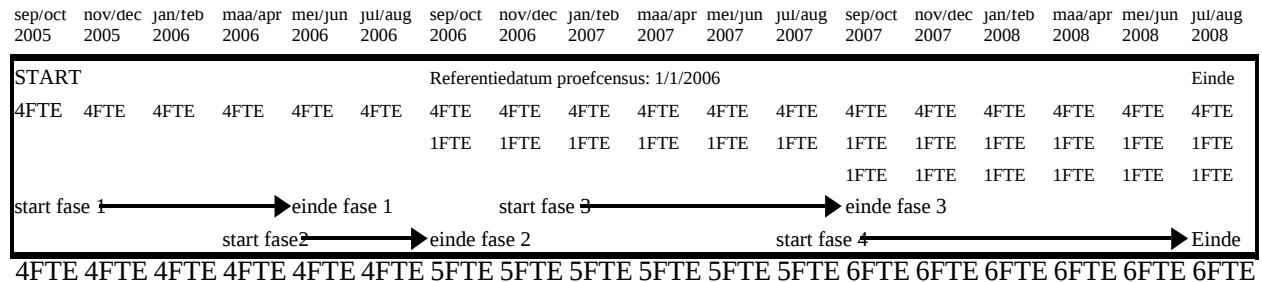
Fase 3: (november 2006 – oktober 2007)

- ontvangst administratieve censusbestanden ref.datum 1/1/2006
- creatie fichier "census 2006" binnen de databank
- taakverdeling datacleaning: onderzoeksploeg ontwerpt procedure datacleaning en test deze uit op steekproeven, Directie Statistiek/I.C.T. implementeert voor de volledige populatie en geeft feedback: datacleaning, harmonisering van anonieme identifier die koppeling toelaat zonder herkenning, consistentie en controle binnen de bron
- Exploitatie van de censusbestanden 2006

Fase 4: (juli 2007 – augustus 2008)

- exploitatie databank: verdere verspreiding geanonimiseerde resultaten (indien kwaliteit gewaarborgd is)
- longitudinale opvolging: vernieuwing van administratieve informatie, introductie nieuwe bronnen
- publicatie en verspreiding wetenschappelijk eindresultaat: "de Belgische socio-economische situatie in 2006" (precieze formule i.f.v. datakwaliteit)
- publicatie procesmatig eindresultaat: "Naar een administratief systeem van gegevensverzameling en productie van socio-economische statistiek in België".

Onderstaande tijdsbalk "Personeelsbezetting" vat deze indeling en timing schematisch samen.



Bijzondere voorwaarden

a. suggesties voor de samenstelling van het begeleidingscomité

- vertegenwoordigers van de Directie Statistiek en van ICT;
- daarenboven is het wenselijk dat vertegenwoordigers van het Comité voor de bescherming van de persoonlijke levenssfeer afgevaardigd worden alsook de verschillende openbare instellingen, beheerders van de betrokken administratieve registers zoals:
 - het Rijksregister van de natuurlijke personen;
 - de Kruispuntbank van de Sociale Zekerheid;
 - de Administratie van het Kadaster;
 - de regionale en gemeenschapsinstellingen bevoegd voor het onderwijs.

b. clausule van vertrouwelijkheid van gegevens

Behalve een voorafgaande herziening van de statistiekwet, moet een contract over de vertrouwelijkheid van de gegevens ondertekend worden door de onderzoeksploeg voor het ter beschikking stellen van anonieme individuele gegevens. De onderzoeksploeg moet zich anderzijds schikken naar de eisen opgesteld door de instellingen die de gegevens leveren of door het begeleidingscomité.

c. aanwezigheid van de onderzoeksploeg in de instelling (indien van toepassing)

De aanwezigheid van de onderzoeksploeg in de instelling is niet vereist. Maar moeten de leden en in het bijzonder de persoon die belast zal worden met de coördinatie van de werkzaamheden en de liaison met de Directie Statistiek en ICT, permanent in contact blijven met de instelling, deelnemen aan overlegvergaderingen waarop ze worden uitgenodigd en antwoorden op de vereisten die door het begeleidingscomité naar voren worden gebracht.

d. Andere

Zoals hoger aangegeven worden bij de Directie Statistiek en ICT samen ongeveer 5 FTE aan gekwalificeerd personeel gepland voor de uitbouw van de databanken ter exploitatie en bijhouding van de censusinformatie. Hiervoor werd een interne projectaanvraag ingediend. Daarnaast wordt voor de onderzoeksploeg het benodigde wetenschappelijke potentieel voor het uitvoeren van de onderzoeksactiviteiten geraamd op 5 FTE aan personeel over 3 jaren, nochtans progressief

opgebouwd volgens de vorderingen van het project. Voor het toewijzen van takenpakketten aan het personeel van de onderzoeksploeg(en) wordt bij wijze van voorbeeld de volgende indeling voorgesteld:

- Een eerste takenpakket behelst de wetenschappelijke en organisatorische coördinatie, de liaison met de Directie Statistiek en ICT, de wetenschappelijke en administratieve opvolging en rapportering van het project als geheel. Dit kan ingevuld worden in 1FTE gedurende de volledige looptijd van het project.
- Een tweede pakket van 1FTE concentreert de aspecten databank, beheer en exploitatie van gegevens, bijhouding, koppeling, integratie van deelinformatie en andere overkoepelende technische, statistische en methodologische aspecten. Ook hier wordt uitgegaan van 1FTE over de volledige looptijd van het project.
- Dan worden in een eerste jaar 2FTE en later telkens 3FTE gepland voor themaspecifieke werkzaamheden die zich concentreren op afzonderlijke bronnen en/of domeinen van socio-economische informatie (demografie en gezin, onderwijs, arbeid, gezondheid, wonen, mobiliteit,...). De onderzoekers moeten van bij de start bij elk deelthema worden betrokken. De onderzoeksploegen zullen echter sterker bezet worden wanneer vanaf het tweede jaar proefbestanden of definitieve steekproeven beschikbaar zijn.
- Ten slotte wordt voor het derde werkjaar een extra FTE gepland. Naast het afronden en rapporteren van het studiewerk moet dan nog de consolidatie van de databank, verspreiding en toelichting van de resultaten en de kennisoverdracht van de onderzoeksploeg aan de Directie Statistiek en ICT plaatshebben.