

Verslag voor de gebruikersgroep

Januari 2022

Inhoud

1. Algemene stand van zaken.....	2
1.1. Bronnen	2
1.2. Webtoepassingen	4
1.3. Documentatie en het CMS	5
2. Stand van zaken projecten.....	6
2.1. Onderzoeksproject: indicatoren ontwikkelen die gewoonlijk berekend worden a.d.h.v. enquêtegegevens	6
2.1.1. Het beschikbaar inkomen op huishoudniveau	6
2.1.2. Transitie werklozen	6
2.1.3. Telling werklozen en grensarbeiders	6
2.1.4. Indicatoren niveau statistische sector	7
2.1.5. Tewerkstellingsindicatoren	7
2.1.6. NEET-jongeren	7
2.2. Onderzoeksproject: een meet- en opvolgingsinstrument voor regionale tewerkstellingsmaatregelen	8
2.3. Onderzoeksproject: informatie over het scholingsniveau van de bevolking	9
2.4. Onderzoeksproject: indicatoren voor een boordtabel sociale bescherming	10
2.5. Onderzoeksproject: variabelen m.b.t. de socio-economische achtergrond	11
2.5.1. Ontwikkeling van een afgeleide variabele loopbaanhistoriek	11
2.5.2. De groep 'andere' in de nomenclatuur	11
2.5.3. Voorbereiden en integreren beroepencode	11
2.5.4. Indicatoren effectiviteit arbeidsmarktmaatregelen	12
2.5.5. Informatie herkomst, immigratiegeneratie en cohorte nieuwkomers	12
2.5.6. Indicator socio-economische achtergrond	12

Inleiding

In dit verslag geven we een **algemene stand van zaken** met betrekking tot het DWH AM&SB (hierna 'het Datawarehouse') en behandelen we de **voortgang van de onderzoeksprojecten** van het wetenschappelijk project voor het Datawarehouse, die werden bepaald in de stuurgroep met de financiers. Deze onderzoeksprojecten worden uitgevoerd door het consortium CeSO, METICES, Steunpunt Werk. Het betreft volgende onderzoeksprojecten:

- Taak 1. De ontwikkeling van indicatoren met betrekking tot de arbeidsmarkt en sociale bescherming op basis van de administratieve gegevens van het Datawarehouse, die de indicatoren indien mogelijk reproduceren die gewoonlijk berekend worden op basis van enquêtegegevens
- Taak 2. De ontwikkeling van een meet- en opvolgingsinstrument voor de regionale tewerkstellingsmaatregelen
- Taak 3. De verbetering van de informatie over het scholingsniveau van de bevolking
- Taak 4. De verbetering van indicatoren voor een bordtabel sociale bescherming
- Taak 5. Ontwikkeling van variabelen met betrekking tot de socio-economische achtergrond

We bespreken eerst de algemene stand van zaken (deel 1) en overlopen nadien achtereenvolgens bovenstaande onderzoeksprojecten (deel 2).

1. Algemene stand van zaken

1.1. Bronnen

Stand van zaken van het inladen van de data in het Datawarehouse:

- Het inladen van de data 2019 is nagenoeg afgerond. FWB-ARES (SATURN) en RSVZ-Eclipz moeten nog ingeladen worden.
- Het inladen van 2020 wordt verdergezet.
- Het inladen van de voorlopig bestanden in het kader van de COVID-19-crisis wordt ook verdergezet.

FOD Buitenlandse Zaken: er werd aan de FOD gevraagd een nieuw bestand aan te leveren. Hier is echter tot nu toe geen reactie op gekomen. Het dossier is geblokkeerd.

LIMOSA: dit betreft werknemers die tijdelijk of gedeeltelijk in België werken voor een buitenlandse werkgever en buitenlandse zelfstandigen die tijdelijk of gedeeltelijk in België opdrachten uitvoeren. Beiden vallen niet onder de Belgische sociale zekerheid. De periode waarvoor de data beschikbaar zijn werd uitgebreid. De data zijn nu beschikbaar voor de periode 2010-2020.

WABRO: dit betreft werknemers die tijdelijk of gedeeltelijk in het buitenland werken voor Belgische werkgevers. De data werden ingeladen voor de periode 2007-2020 maar na testen bleek dat een nieuwe inlading nodig is. De data moeten ook nog gedocumenteerd worden.

KBO: de data werden gecheckt. Omdat het bestand bijzonder groot is zal het worden opgesplitst in deelbestanden. De data moeten ook nog gedocumenteerd worden.

FWB - verplicht onderwijs, de FWB-ARES (SATURN) en het Vlaams Departement Onderwijs: deze instellingen vragen een protocol af te sluiten met de KSZ. De KSZ is echter eerder geen voorstander van het afsluiten van protocollen en vindt het wenselijker dat deze instellingen zou toetreden tot het netwerk van de KSZ. Over dat laatste heeft er in de komende weken overleg plaats tussen de KSZ en de regionale overheden. Als die overheden effectief toetreden dan hoeft er geen protocol meer afgesloten te worden en is het Informatieveiligheidscomité bevoegd.

STATBEL: de gegevens over het opleidingsniveau werd ingeladen voor de jaren 2012-2017. Voor 2016 zijn er ook gegevens over schoollopenden.

Er werden bestanden ingeladen met geo-coördinaten, zowel op het niveau van de adressen als van de statistische sectoren.

Er werden drie nieuwe bestanden gecreëerd voor de POD Maatschappelijke Integratie, op trimestriële basis. Het eerste bestand bevat een classificatie, die hiërarchisch is opgebouwd, die de aard van de OCMW-steun weergeeft. Het tweede bestand bevat informatie over het Geïndividualiseerd Programma Maatschappelijke Integratie (GPMI). Het derde bestand bevat de personen die louter een installatiepremie hebben gekregen om uit de dakloosheid te treden (wetgeving 6). Deze populatie wordt niet opgenomen in de nomenclatuur.

Het inladen van tijdelijke bestanden naar aanleiding van de COVID-crisis om data-aanvragen hierover te kunnen beantwoorden, wordt verdergezet. Momenteel zijn er bestanden beschikbaar van de RSZ, het RSVZ, het RIZIV, het NIC en de RVA. Wanneer de structurele bestanden ingeladen zijn, zullen deze tijdelijke bestanden niet langer bewaard worden. Op die manier werden al bestanden verwijderd van de FPD, de POD Maatschappelijke Integratie, de kinderbijslaggegevens (FAMIFED-ORINT), de RVA. en VDAB/FOREM/ACTIRIS/ADG. Sommige van die tijdelijke bestanden bevatten echter informatie die niet aanwezig is in de structurele bestanden. In dat geval zal die informatie wel permanent bewaard blijven.

Er zullen nieuwe bestanden worden opgenomen m.b.t. de arbeidsongevallen. Er zijn nieuwe bestanden over de arbeidsongevallen en de tijdelijke en permanente arbeidsongeschiktheid. Er is ook een bestand over de rechthebbenden (nabestaanden). De gegevens hebben zowel op de private- als de publieke sector betrekking. De gegevens over de publieke sector starten echter maar vanaf 2014. Later moeten er nog gegevens opgenomen worden over de betalingen.

FPD - pensioenkadaster: de variabelen “soort pensioen” en “soort pensioen volledig” werden aangepast vanaf 2019. Dit heeft een nagenoeg verwaarloosbare impact op de nomenclatuur van de socio-economische positie. Het heeft wel een grotere impact op de bijkomende socio-economische posities m.b.t. de pensioentrekkenden.

Regionale gegevens over de kinderbijslagen naar aanleiding van de 6de staatshervorming: de gegevens van de Vlaamse- en Duitstalige Gemeenschap werden al opgenomen voor 2019. Voor Wallonië en Brussel zal er gestart worden met de gegevens van 2020. De kadastragegevens werden door ORINT nog bijgehouden voor de jaren 2019 en 2020 voor Brussel, Wallonië en de Duitstalige Gemeenschap. Deze werden opgenomen in het Datawarehouse.

Gegevens over personen met een handicap: langs Vlaamse zijde werden de gegevens over het Zorgbudget voor Ouderen opgenomen voor de jaren 2017-2020, als opvolger voor THAB. Er zal met de andere regio's contact opgenomen worden.

De nomenclatuur moet nog uitgebreid worden met een aantal populaties. Deze worden dan uit de groep n4 gehaald.

Er wordt gewerkt aan een systeem van remote access. In eerste instantie voor het model BELMOD, nadien zal bekeken worden of dit systeem ook kan gebruikt worden voor andere data-aanvragen. Er worden ook testen uitgevoerd m.b.t. scrambling van data.

1.2. Webtoepassingen

De webtoepassingen werden volledig geactualiseerd naar 2019. De webtoepassingen globale en lokale cijfer werden ook gedeeltelijk geactualiseerd naar 2020. Met een volledige actualisatie wordt bedoeld dat alle socio-economische posities aanwezig zijn.

Er werden 2 uitbreidingen voorzien in de webtoepassing lokale cijfers:

- De werkplek werd toegevoegd vanaf 2007, op het niveau van de gemeente.
- Een nieuwe classificatie (IBEFE) voor de woonplaats werd ingevoerd in Brussel en Wallonië vanaf 2015.

In de webtoepassing globale cijfers werd de populatie THAB weer toegevoegd, vanaf het jaar 2017. De opvolger van THAB in Vlaanderen, het Zorgbudget voor Ouderen, is eveneens consulteerbaar, vanaf 2017.

De socio-economische positie m.b.t. OCMW-hulp werd hernoemd. Vroeger werd "leefloon/financiële hulp" gebruikt als label voor de nomenclatuurcodes n331 en n332, nu wordt dat "OCMW-steun - Recht op Maatschappelijke Integratie" en "OCMW-steun - Recht op Maatschappelijke Hulp". Dit is algemener, maar ook correcter. Het is echter geen trendbreuk: de berekeningswijze en de cijfers zijn nog steeds dezelfde. De documentatie zal worden aangepast.

De definitie van interimarbeid werd aangepast. Dit heeft vooral een impact op de cijfers van de laatste jaren. Niettemin werden de basistabellen van de webtoepassing globale cijfers opnieuw aangemaakt vanaf het eerste beschikbare trimester (het eerste trimester van 2003). Het opstellen van nieuwe basistabellen brengt mee dat sommige cijfers wat kunnen veranderen. Dit is voornamelijk het geval in de nomenclatuurcode n4 en voor personen met een woonplaats buitenland of onbekend. Ook de cijfers over de pensioentrekkenden met een buitenlandse woonplaats kunnen wat wijzigen.

De anonimiseringsregel is verstrengd: vroeger werd "1-3" vermeld voor aantallen kleiner dan 4 wanneer de statistische sector werd bevraagd. Nu is de regel verstrengd naar "1-9" en hij wordt ook toegepast wanneer de herkomst en de werkplaats worden opgevraagd.

1.3. Documentatie en het CMS

Er werden een aantal aanpassingen aangebracht aan de website: voortaan zijn er tooltips bij de instellingen en de bronnen, die een beknopte beschrijving bevatten. De vernieuwing van de themapagina moet nog gebeuren.

2. Stand van zaken projecten

2.1. Onderzoeksproject: indicatoren ontwikkelen die gewoonlijk berekend worden a.d.h.v. enquêtegegevens

2.1.1. Het beschikbaar inkomen op huishoudniveau

Het doel van deze eerste opdracht is om afgeleide variabelen te ontwikkelen over *het beschikbaar inkomen op huishoudniveau* op basis van administratieve data uit het Datawarehouse en de IPCAL-databank. Met behulp van een administratief inkomensconcept kunnen vervolgens allerhande indicatoren worden ontwikkeld die gewoonlijk worden berekend op basis van enquêtegegevens (zoals bijvoorbeeld de 'AROP-indicator'). We gebruiken de *HY020*-variabele van de SILC-enquête als referentiepunt voor de ontwikkeling van de afgeleide variabelen. Concreet trachten we volgende variabelen te ontwikkelen:

- Basisvariabele: het beschikbaar inkomen op huishoudniveau
- Aanvullende variabelen: afzonderlijke inkomensvariabelen op huishoudniveau zoals bijvoorbeeld het beschikbaar inkomen uit ouderdomsuitkeringen per huishouden en het beschikbaar inkomen uit loondienst per huishouden.

In de voorbije maanden hebben we een voorlopige versie van de basisvariabele opgesteld op basis van een nieuwe gegevensaanvraag en hebben we de achterliggende methodologie verder verfijnd. Daarnaast werd ook een eerste analyse uitgevoerd over de verschillen tussen het ontwikkelde inkomensconcept en de *HY020*-variabele. Hieruit blijkt dat de resultaten van de basisvariabele afwijken van de *HY020*-variabele maar dat de verschillen wellicht deels te wijten zijn aan het gebruik van administratieve gegevens. Zo kunnen de verschillen het gevolg zijn van ontbrekende inkomens of beperkt geregistreerde inkomens in de gebruikte databanken. Verder onderzoek is nodig om de geobserveerde verschillen tussen de basisvariabele en de *HY020*-variabele verder te verklaren. Eind september werd het project echter *on-hold* geplaatst omdat Statbel ook werkt aan de ontwikkeling van een gelijkaardig inkomensconcept. Om dubbel werk te vermijden werd bijgevolg beslist dat er eerst overleg moet plaatsvinden met Statbel. Deze gesprekken zullen kortelings worden opgestart.

2.1.2. Transitie werklozen

Voor deze opdracht werd de vraag gesteld variabelen te ontwikkelen die de mobiliteit van werkzoekenden op de arbeidsmarkt in kaart brengen. Op dit ogenblik is er nog geen onderzoeksplan opgesteld. Hierover zal binnenkort een overleg plaats hebben met de opdrachtgever(s).

2.1.3. Telling werklozen en grensarbeiders

In deze taak vragen de opdrachtgevers om werklozen en uitgaande grensarbeiders af te bakenen. Op dit ogenblik is er nog geen onderzoeksplan opgesteld. Hierover zal binnenkort een overleg plaats hebben met de opdrachtgever(s).

2.1.4. Indicatoren niveau statistische sector

Sinds kort is er een bestand met geo-coördinaten aanwezig in het Datawarehouse. Deze coördinaten kunnen ook gelinkt worden aan de statistische sectoren. In het kader van deze taak vond een eerste verkennend overleg plaats tussen de financiers en de onderzoekers. In een volgende fase zullen de financiers intern nagaan welke ideeën zij hebben voor zulke indicatoren en zullen ze deze ideeën overmaken aan het CeSO. Het CeSO zal zelf ook nagaan wat mogelijk is met deze geo-coördinaten en zal een meeting plannen om het project verder met de financiers te bespreken.

2.1.5. Tewerkstellingsindicatoren

In het voorbije jaar werd de opdracht voor de ontwikkeling van tewerkstellingsindicatoren in het Datawarehouse afgerond. We leverden een begeleidende nota bij de tewerkstellingsindicatoren op waarin we de constructie en enkele resultaten van de indicatoren bespreken. De variabelen om de tewerkstellingsgraden in het Datawarehouse te berekenen, zijn gedocumenteerd op de website van het Datawarehouse. Deze variabelen zijn nu beschikbaar voor de gebruikers van het Datawarehouse en zitten vervat in de bron DWH_BCSS_Emploi. Ook de begeleidende nota kan teruggevonden worden bij de documentatie op de website.

2.1.6. NEET-jongeren

Om NEET-jongeren af te bakenen op basis van de gegevens uit het Datawarehouse is het aangewezen om nieuwe databronnen over onderwijs- en opleidingsdeelname in het Datawarehouse te integreren. In het voorbije jaar werd geen vooruitgang geboekt in de integratie van deze databronnen in het Datawarehouse. We geven hieronder een stand van zaken mee over dit integratieproces.

Eerder ontving het Datawarehouse al gegevens van ARES (Franstalig niet-universitair hoger onderwijs) en de Cref (Franstalig universitair onderwijs) over onderwijsdeelname. De gegevens van ARES zijn beschikbaar voor de jaren 2017 en 2018. De data van de Cref gaan van 2018 terug tot in de jaren 1980. Vanwege ARES, de Fédération Wallonie-Bruxelles (Franstalig leerplichtonderwijs) en het Departement Onderwijs en Vorming (onderwijs van de Vlaamse Gemeenschap) werden in het afgelopen jaar geen (nieuwe) data ontvangen.

Wel werden in november 2021 data afkomstig van Statbel over het onderwijsniveau geïntegreerd in het Datawarehouse en dit voor de periode 2012-2017. In deze data is voor het bestand van 2016 ook een variabele opgenomen die aanduidt wie leerling of student is. Het onderzoeksteam zal samen met de medewerkers van de KSZ de Statbel-gegevens testen op hun bruikbaarheid voor het bepalen van de NEET-jongeren, in combinatie met de huidige gegevens die al beschikbaar waren in het Datawarehouse.

2.2. Onderzoeksproject: een meet- en opvolgingsinstrument voor regionale tewerkstellingsmaatregelen

In eerste instantie moeten de doelstellingen van de vorige financiersperiode worden voortgezet en afgerond: met name, de ontwikkeling van een inventaris van de regionale tewerkstellingsmaatregelen, de analyse daarvan en de integratie van deze maatregelen in het Datawarehouse.

Daarnaast moet volgende opdracht worden uitgevoerd:

Een typologie definiëren (of een bestaande typologie verfijnen) die het mogelijk maakt om de verschillende maatregelen te classificeren, waarbij in het bijzonder rekening wordt gehouden met het budgettaire gewicht dat aan de verschillende soorten maatregelen (stimulansen, verplichtingen, enz.) is verbonden. Een dergelijke typologie zou vergelijkingen in de tijd mogelijk maken (zelfs als de maatregelen veranderen) binnen dezelfde regio en tussen regio's.

Wat betreft de eerste opdracht werd in een overleg op 7 juni 2021 besloten dat de tijdslijn van de tewerkstellingsmaatregelen in België moet afgewerkt worden, omdat deze noodzakelijk is voor het verdere verloop van het project. Verder werd er besloten dat de nieuwe taak over het budget dat aan de verschillende tewerkstellingsmaatregelen is toegewezen in de tijdslijn kan worden opgenomen, op voorwaarde dat dit de verwezenlijking van de tijdslijn niet belemmert. De informatie rond de financiering kan ook in een afzonderlijk document worden opgenomen.

Er werd in de periode mei-december verder gewerkt aan de tijdslijn van tewerkstellingsmaatregelen.

Momenteel wordt deze tijdslijn van tewerkstellingsmaatregelen en de informatie die daarin wordt opgenomen nagekeken en aangevuld door de betrokken instellingen.

Wat betreft de tweede opdracht: het Centre METICES wacht op een uitgebreide toelichting van de opdrachtgever(s) vooraleer ze start met deze opdracht en een planning opmaakt. Er zal echter reeds een nota worden opgesteld waarin wordt nagegaan of het mogelijk is de budgetten die zijn toegewezen aan de werkgelegenheidsmaatregelen die in de inventaris van het consortium zijn opgenomen, te onderzoeken.

De discussie over de tijdslijn is nog niet afgerond. Er wordt nog gewacht op opmerkingen van verschillende instellingen. Daarna kan er overgegaan worden tot de volgende stap. Er zal opnieuw contact opgenomen worden met de instellingen om hun opmerkingen op de tijdslijn te noteren.

Het Centre METICES zal verifiëren of de elementen van financiering al werden geïntegreerd in de analyse.

2.3. Onderzoeksproject: informatie over het scholingsniveau van de bevolking

De financiers van het wetenschappelijk project voor het Datawarehouse formuleerden de vraag om na te gaan hoe de informatie over het scholingsniveau van de bevolking in het Datawarehouse kan worden verbeterd. Het initiële opzet voor de verbetering van de informatie over het scholingsniveau was om het scholingsniveau zoals bepaald door de FOD WASO & UNIA (2019) over te nemen in het Datawarehouse en te verfijnen waar mogelijk.¹ In de loop van het project werd evenwel duidelijk dat er sterke gelijkenissen zijn tussen deze opdracht van het wetenschappelijk project voor het Datawarehouse en het project van Statbel over de creatie van een onderwijsniveau op basis van administratieve data. Er werd besloten om de data met het onderwijsniveau bepaald volgens Statbel te integreren in het Datawarehouse. De Statbel-data werden in november 2021 ingeladen in het Datawarehouse voor de jaren 2012 tot en met 2017. De gegevens over het onderwijsniveau zullen weldra gedocumenteerd worden op de website van het Datawarehouse en beschikbaar gesteld worden voor de gebruikers.

¹ FOD WASO & UNIA (2019). *Socio-economische monitoring. Arbeidsmarkt en origine 2019*. Brussel: FOD WASO & UNIA.

2.4. Onderzoeksproject: indicatoren voor een boordtabel sociale bescherming

De Universiteit Antwerpen heeft een boordtabel ontwikkeld waarin te ontwikkelen indicatoren zijn opgesteld die nodig zijn om de kwaliteit van het Belgische sociale beleid te kunnen monitoren. Eén van die indicatoren gaat over de uitkeringsontvangers. De bedoeling is om indicatoren te ontwikkelen die het aantal uitkeringsontvangers meet en de gemiddelde en mediane hoogte van die uitkeringen weergeeft. Daarnaast wordt er ook gewerkt aan een indicator over personen met een handicap.

In 2021 is het project van start gegaan. Na het verkennen van de data werd een eerste indicator opgesteld. Bij het opstellen van die indicator kwamen een aantal methodologische kwesties naar boven over de manier waarop de indicatoren geconstrueerd worden en hoe ze gepresenteerd worden. In overleg met de financiers hebben we beslist om één indicator volledig op te stellen en op basis van die voorbeeldcase te bepalen welke methodologie en presentatie we zullen gebruiken. Deze voorbeeldcase is afgewerkt. De resultaten hiervan worden begin januari met de financiers besproken. Voor de indicator over personen met een handicap hebben we onderzocht welke data nog in het Datawarehouse zouden moeten worden ingeladen.

Het plan voor 2022 is om, na een volgend overleg met de financiers, de indicatoren over het aantal uitkeringsontvangers en de indicatoren over de hoogte van de uitkeringen op te stellen. We nemen ook contact op met verschillende instanties om de inlading van nieuwe gegevens over personen met een handicap verder te plannen.

2.5. Onderzoeksproject: variabelen m.b.t. de socio-economische achtergrond

2.5.1. Ontwikkeling van een afgeleide variabele loopbaanhistoriek

Het zal in de toekomst mogelijk worden om op basis van het Datawarehouse de volledige loopbaan als werknemer, zelfstandige en ambtenaar in kaart te brengen met behulp van de aanwezige Eclipz-gegevens en de nieuwe Sigedis-gegevens (die in de loop van dit jaar beschikbaar zullen zijn voor de gebruikers; zie ook het verslag van de gebruikersgroep van december 2020). Het inzichtelijk voorstellen van de volledige loopbaanhistoriek is een complex gegeven. We willen daarom een aantal afgeleide variabelen – gelijkaardig aan diegenen die tot en met 2010 aanwezig waren in de bron DWH_SIGEDIS_VarDer – over de loopbaanhistoriek integreren in het Datawarehouse om dit gegeven op een gebruiksvriendelijke manier ter beschikking te stellen aan het brede publiek. Het ontwikkelen van deze afgeleide variabelen vereist evenwel een uitgebreide methodologie. Deze methodologie wordt op dit moment uitgewerkt voor werknemers. Tijdens de volgende periode dient hetzelfde te gebeuren voor zelfstandigen en ambtenaren. Een alternatieve piste is om reeds bestaande afgeleide variabelen in te laden in het Datawarehouse. Afgeleide variabelen die toelaten de loopbaanhistoriek te reconstrueren zijn immers aanwezig in de databanken die aan de basis liggen van mypension.be en mycareer.be. Tijdens de volgende periode zullen we bijgevolg eveneens onderzoeken of het mogelijk én aangewezen is om dergelijke afgeleide variabelen rechtstreeks in het Datawarehouse in te laden.

2.5.2. De groep ‘andere’ in de nomenclatuur

De positie ‘andere’ in de nomenclatuur bevat een erg grote groep. Voor individuen in deze groep is het niet duidelijk of het werkenden, werkzoekenden of niet-beroepsactieven zijn. We zullen nagaan hoe deze groep verder kan worden verkleind door ze toe te wijzen aan één van de 16 andere statuten. Momenteel onderzoeken we of een deel van deze groep kan worden toegewezen aan het statuut ‘pensioengerechtigden’. We vermoeden immers dat veel individuen uit deze groep huwelijkspartners zijn van de ontvangers van een gezinspensioen.

2.5.3. Voorbereiden en integreren beroepencode

Een belangrijk hiaat in het Datawarehouse zijn de gegevens over het beroep van personen. In enkele van de databronnen van het Datawarehouse is informatie opgenomen over het beroep, maar voor het overgrote deel van de werkenden is het beroep niet gekend. In het kader van het wetenschappelijk project voor het Datawarehouse werd de vraag gesteld om dit hiaat te verhelpen en gegevens over het beroep te integreren in het Datawarehouse. Als eerste stap van deze opdracht hebben we een nota opgesteld met een inventaris van de beschikbaarheid van beroepsgegevens in administratieve databronnen. Een korte blik op het datalandschap in België leert ons dat het hiaat in beroepsgegevens niet beperkt is tot het Datawarehouse. De administratieve databronnen van de verschillende instellingen bevatten weinig of geen informatie over het beroep of de informatie is van onvoldoende kwaliteit. Voor de volgende editie van de Census werkt ook Statbel aan een statistiek over de beroepen in België op basis van administratieve data. Hierbij stuiten ze op dezelfde problemen als deze die we hierboven vermelden. Omwille van het beperkte aanbod van administratieve beroepsgegevens geeft Statbel aan om voor slechts een vierde van de beroepsbevolking administratieve data te gebruiken.

In een volgende stap stellen we ons tot doel om de kennis over administratieve beroepsgegevens verder uit te diepen en te verbreden. Via inzicht in het hoe en waarom van de (niet-)registratie van beroepsgegevens in binnenlandse- en buitenlandse voorbeelden willen we de noden van de registratie van beroepsgegevens en de mogelijkheden dit biedt voor onderzoek en beleid verder mee op de kaart zetten.

2.5.4. Indicatoren effectiviteit arbeidsmarktmaatregelen

In deze opdracht worden indicatoren ontwikkeld die de effectiviteit van arbeidsmarktmaatregelen meten. Indien mogelijk worden deze indicatoren geïntegreerd in de webtool rond socio-economische mobiliteit, zodat doorstroom naar (niet-gesubsidieerde) tewerkstelling in kaart wordt gebracht.

In de komende periode wordt in overleg met de opdrachtgevers een planning en onderzoeksproject uitgewerkt.

2.5.5. Informatie herkomst, immigratiegeneratie en cohorte nieuwkomers

Voor deze taak vragen de opdrachtgevers een opvolgingssysteem voor nieuwkomers te ontwikkelen in het Datawarehouse. We zullen in een eerste stap nagaan welke gegevens voorhanden zijn in het Datawarehouse voor de ontwikkeling van een dergelijk opvolgingssysteem. Hierbij kijken we zowel naar demografische gegevens als naar arbeidsmarktgegevens. In een volgende stap zullen we een methodologie ontwikkelen om de populatie af te bakenen en de bijhorende socio-economische gegevens te definiëren die relevant zijn om de integratie van de populatie in kaart te brengen.

2.5.6. Indicator socio-economische achtergrond

Het Centre METICES werkt, in samenwerking met de FOD WASO, aan de ontwikkeling van indicatoren met betrekking tot diversiteit op de arbeidsmarkt (socio-economische monitoring). Deze opdracht bestaat erin één of meerdere standaardvariabelen met betrekking tot de “socio-economische achtergrond” van personen te ontwikkelen op basis van administratieve gegevens van het Datawarehouse, en deze vervolgens te koppelen aan hun arbeidsmarktpositie.

De volgende stappen worden voortgezet:

1. Het uitvoeren van verkennende analyses.
2. Het koppelen van de gegevens van de Census en de gegevens van de KSZ. In een eerste stap, wordt er voor de periode tussen 1998 en 2013 voor elk jaar een afzonderlijk databestand opgesteld waarin alle informatie wordt opgenomen. In een tweede stap, worden er op basis daarvan vijf bestanden aangemaakt waarin telkens informatie over de geselecteerde individuen wordt opgenomen van toen zij (a) 8 jaar, (b) 10 jaar, (c) 12 jaar, (d) 14 jaar en (e) 16 jaar oud waren.
3. Het selecteren van één van deze vijf ‘jaar’-bestanden voor de opstelling van de indicator. Hier zal de werkgroep verdere analyses voor moeten verrichten, op basis waarvan de keuze zal worden gemaakt.

Stap 1 werd reeds uitgevoerd voor de jaren 2003 en 2005. Stap 2 is momenteel aan de gang.

In de periode tussen mei en december 2021 werd verder gewerkt aan de koppeling van de gegevens van de Census en van de KSZ. Voor elk jaar tussen 1998 en 2013 afzonderlijk werd een dataset ontwikkeld waarin deze gegevens worden opgenomen. Dit gebeurde aan de hand van een 'standaardsyntax', die voor elk jaar werd uitgeschreven.

Er werd eveneens bij STATBEL nagegaan of er aanvullende gegevens over de inkomens beschikbaar zijn voor de periode vóór 2005. Dit bleek niet het geval te zijn. De jaarbestanden vanaf 2005 bevatten dus de meeste informatie, aangezien daarin zowel de nomenclatuurpositie (beschikbaar vanaf 2003) als het inkomen van de betrokken personen voor het betreffende jaar aanwezig zijn.

De volgende stap in dit project is het ontwikkelen van vijf bestanden per leeftijd (8, 10, 12, 14 en 16 jaar) en het kiezen van de ideale leeftijd op basis waarvan de indicator zal worden ontwikkeld.