

Datacoalitie Datagedreven onderzoek mbo

Activiteitenplan voor verbeterde data over werving, instroom en doorstroom

Inhoudsopgave

Inleiding.....	3
Hoofdstuk 1: Doelstelling, aanpak en bevindingen.....	3
Bevindingen:	4
Werving & Instroom	4
Doorstroom binnen het MBO (intern of tussen instellingen)	5
Doorstroom naar HBO	5
Doorstroom naar Arbeid.....	5
Hoofdstuk 2. - De uitgangspunten.....	6
Uitgangspunt 1: Open data tenzij.....	6
Uitgangspunt 2: Een dashboard is aanvullend en vervangt de databehoefte niet	7
Uitgangspunt 3: we houden rekening met volwassenheidsverschillen.....	7
Uitgangspunt 4: wat we doen is voor de gehele sector	8
Hoofdstuk 3. - Vervolgactiviteiten.....	8
Stap 1. Vertrekpunt vaststellen en toetsen	8
Stap 2. Dataleveringen op orde brengen.....	8
Stap 3. Bruikbaarheid en bekendheid van de databronnen verhogen.....	9
Hoofdstuk 4. - Integratie met lopende initiatieven.....	11
Hoofdstuk 5. - Organiseren van het werk	12
Bijlage 1. Gewenste basisset van gegevens per informatiebehoefte.....	13

Inleiding

Datagedreven onderzoek en werken staat of valt met de beschikbaarheid van de benodigde data.

Binnen werkgroep 2 van de datacoalitie datagedreven onderzoek in het mbo is gekeken welke informatiebehoefte instellingen hebben ten aanzien van werving, instroom en doorstroom. Vervolgens is gekeken welke data (of bestanden) gewenst is, in hoeverre die data nu van voldoende kwaliteit beschikbaar is én zo niet, wat er moet gebeuren om deze kwaliteit beschikbaar te krijgen

In dit activiteitenplan is een overzicht opgenomen van het doorlopen proces en de belangrijkste bevindingen die het heeft opgeleverd (hoofdstuk 1). In hoofdstuk 2 & 3 worden een aantal uitgangspunten geformuleerd en activiteiten om de behoefte en het aanbod beter op elkaar aan te laten sluiten. In hoofdstuk 4 & 5 wordt uiteengezet hoe de verbeterde aansluiting in de praktijk geoperationaliseerd kan worden, hoe het werk kan worden verdeeld en tot welk resultaat dit zou moeten leiden.

Aangezien de mbo-sector in zijn geheel belang heeft bij een passend informatieaanbod zijn de bevindingen, uitgangspunten en voorstellen ook getoetst bij de MBO Raad. De MBO Raad herkende de bevindingen en wensen en denkt mee in het vervolg hoe de verbetervoorstellen kunnen worden gerealiseerd.

Goed om te weten:

In de werkgroep is nauw samengewerkt met DUO. De geformuleerde wensen zijn opgesteld in samenspraak met DUO, maar kunnen ook voor andere “aanbieders” relevant zijn. Hoe beter de databestanden qua detailniveau en definities op elkaar aansluiten hoe effectiever ze door de sector te gebruiken zijn.

Hoofdstuk 1: Doelstelling, aanpak en bevindingen

Vanuit de datacoalitie datagedreven onderzoek in het MBO zijn er 3 werkgroepen gevormd die zich elk bezig houden met een ander onderzoeksonderwerp. Het beoogd doel van werkgroep 2 is het in kaart brengen en verbeteren waar nodig van relevante data rondom werving, instroom, doorstroom binnen het mbo en doorstroom naar hbo of arbeid.

Binnen deze opzet is met de instellingen gekeken welke behoefte er is vanuit de onderwijsinstellingen, welke aanbod van data er is vanuit DUO en waar die behoefte en het aanbod niet op elkaar aansluiten. Waar behoefte en aanbod niet aansluiten, is onderzocht of daar verbeteringen mogelijk zijn en welke verbeteringen dat zijn. In dit hoofdstuk zijn de belangrijkste bevindingen opgenomen.

Bevindingen:

Er is behoefte aan eenduidige, heldere databestanden

De instellingen binnen deze werkgroep zijn voorlopers op het gebied van werken met data. Ze gebruiken data uit de eigen systemen, maar maken ook gebruik van dataleveringen vanuit partners of dashboards (bijvoorbeeld het dashboard <https://www.onderwijsinderegioincijfers.nl/>). Ze geven aan dat ze data vanuit verschillende bronnen willen (kunnen) combineren. Hiervoor ontvangen ze het liefst (platte) databestanden. Die zijn makkelijker te combineren dan data uit bijvoorbeeld dashboards.

Wat ontvangen ze het liefst?

- Platte databestanden (csv of dergelijk)
- Bestanden met daarin een vastgestelde set stamgegevens (zoals BRINNR, CREBO). Hierdoor is het te koppelen met andere bronnen
- Indien mogelijk als open-data bestand, maar als het detailniveau dat niet toelaat dan achter autorisatie

Dashboards zijn bonus

Een dashboard kan helpen om inzicht te verschaffen in complexe datavraagstukken. Het is een goede manier om gemiddelde gebruikers te helpen inzicht te vinden. Er is echter een groot nadeel dat met name voor gevorderde gebruikers als beperkend ervaren wordt. Een dashboard visualiseert data, maar het is vaak niet mogelijk om de onderliggende data ook uit het dashboard te halen en te hergebruiken/combineren met andere bronnen. Het heeft dan ook de voorkeur om dashboards als complementair te zien en niet als supplementair op de dataleveringen.

Verschillende informatiebehoefte vragen om verschillend detailniveau

Het detailniveau van data is afhankelijk van de behoefte. Als je een beeld wilt vormen van een individuele student dan vraagt dat om informatie op een laag detailniveau (granulariteit), maar voor een macro-economisch doel is data op een hoger detailniveau afdoende. Hoe hoger het detailniveau hoe beter de behoefte past bij open data.

Gedetailleerder dan BRIN

De dataleveringen zijn nu veelal gericht op instellingen (BRIN). Het gewenste niveau, voor vrijwel alle dataleveringen, is echter nauwgezetter. Het gaat dan om een combinatie van Brin, onderwijsaanbieder & onderwijslocatie. Daarbinnen wordt de student dan nog toegewezen aan een specifieke opleiding (gekoppelde CREBO). De locaties kunnen sterk van elkaar verschillen en daarmee ook de aandachtspunten. Het heeft de voorkeur om waar mogelijk ook de onderwijslocatie (op basis van RIO) mee te nemen in de dataleveringen.

Werving & Instroom

Instellingen hebben behoefte aan een volledig beeld van de (toekomstige) student. DUO heeft data over iedereen die is ingeschreven bij een onderwijsinstelling (PO, VO, MBO, HO). In het

kader van instroom is informatie over de eerdere opleiding van groot belang. De instellingen hebben op het gebied van werving en instroom 2 specifieke wensen:

- Inzicht in studenten die niet kiezen voor de eigen instelling (via bijvoorbeeld centraal aanmelden). Specifieker: de instellingen hebben graag zicht op de aantallen studenten die voor een andere instelling kiezen terwijl ze wonen in het voedingsgebied van de instelling.
- Inzicht in de toekomstige student middels informatie over de vooropleiding en aanleverende school

Doorstroom binnen het MBO (intern of tussen instellingen)

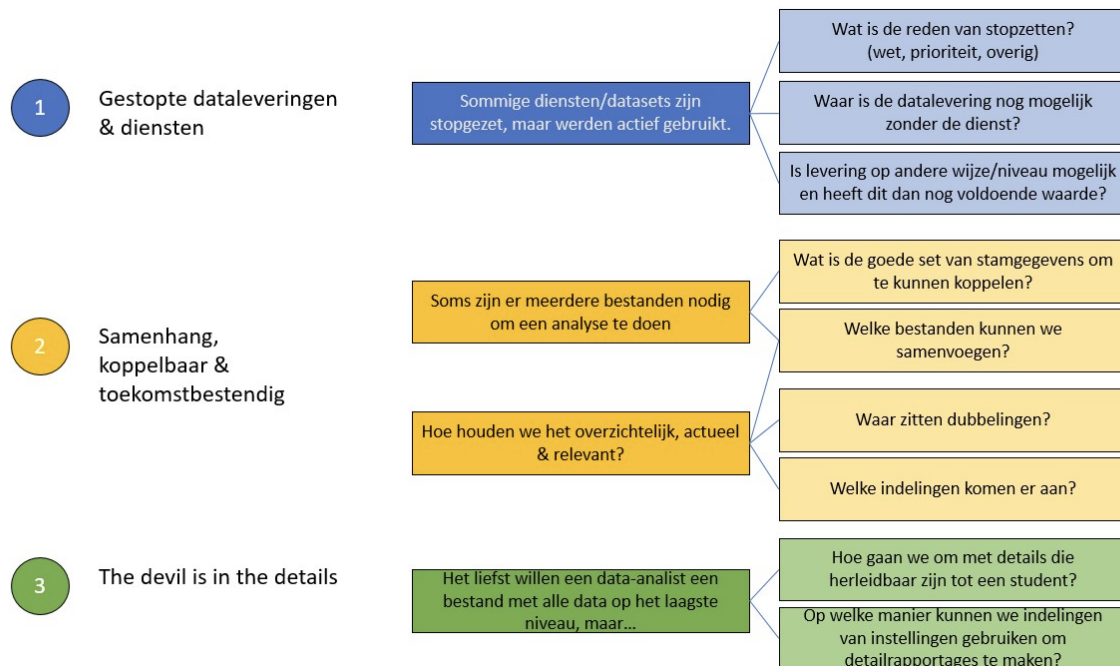
Via de eigen administratiesystemen hebben scholen een duidelijk beeld van de studenten die intern doorstromen. Inzicht in de doorstroom naar (en succes op) andere mbo-instellingen is in het kader van kwaliteitsonderzoek en -borging en voor alumni-beleid wel wenselijk.

Doorstroom naar HBO

Eerder hadden instellingen via de voorziening “Inzien” een goed beeld van doorstroom naar (en succes op) HBO-opleidingen. De ondersteuning van dit dashboard is echter stopgezet. De werkgroep heeft aangegeven dat de behoefte aan deze inzichten niet is afgenomen. Ze geven aan dat ze graag de onderliggende data (en KPI’s) ontvangen die is gebruikt ten behoeve van “Inzien”.

Doorstroom naar Arbeid

Vanuit de doorstroom naar arbeid zijn de wensen minder uitgewerkt. Wel is er interesse om een samenwerking met SBB of CBS te verkennen om te kijken naar databestanden die daar beschikbaar zijn zoals “Kans op werk”.



Figuur 1.) schematische weergave van de bevindingen

Hoofdstuk 2. - De uitgangspunten

Voordat we aan de slag kunnen met de bevindingen is het nodig om een aantal uitgangspunten te formuleren waar een oplossing aan moet voldoen. Dat helpt bij het creëren van duurzame oplossingen. De werkgroep heeft hiervoor 4 uitgangspunten geformuleerd.

Uitgangspunt 1: Open data tenzij...

In de mbo-sector wordt veel gebruik gemaakt van open data, data die vrij toegankelijk is voor alle gebruikers. Open data bestanden hebben een aantal voordelen. Deze bestanden kunnen breed gebruikt worden door de mbo-sector en ze zijn efficiënt door bijvoorbeeld DUO samen te stellen en beschikbaar te stellen. Daarnaast heeft DUO het beleid om geen ad-hoc leveringen voor 1 instelling te faciliteren, maar dataleveringen ten behoeve van de gehele sector te realiseren. Deze databestanden zijn ook goed toe te passen op de informatiesystemen van de instellingen.

Open data bestanden hebben dus de voorkeur, maar zijn niet altijd geschikt gezien het doel of de gevoeligheid van de data. Naast open data bestaat de mogelijkheid om data wel beschikbaar te stellen voor de hele sector, maar in een beveiligde omgeving middels autorisatie.

We hebben een aantal voorwaarden geformuleerd wanneer een open data bestand de geschikte optie is en wanneer gekozen dient te worden voor een levering achter autorisatie.

Wanneer is een open data bestand wel geschikt?

- Wanneer het bestand gegevens bevat van alle instellingen binnen de sector en
- Wanneer het bedoeld is voor een brede weergave van trends & aantallen in de mbo-sector en
- Wanneer het gaat om studentaantallen (in een enkele rij/kolom) van meer dan 5 studenten *.

Wanneer is data in een beveiligde omgeving beter geschikt?

- Wanneer alleen de gegevens van de eigen instelling worden geleverd of
- Wanneer het gaat om studentaantallen (in een enkele rij/kolom) van minder dan 5 studenten of
- Wanneer er sprake is van “verrijkte data”.
Bijvoorbeeld berekende gegevens of aangevulde gegevens over de vervolgopleiding en dergelijke. **

* NB: Er zijn verschillende methodes om databestanden aan te passen waardoor levering als open data alsnog tot de mogelijkheden behoort. Bijvoorbeeld wanneer slechts enkele velden of rijen kleine aantal studenten representeren. In samenspraak met sector kunnen de mogelijkheden besproken worden en kan gekozen worden voor een open data levering met geanonimiseerde records.

** Er wordt binnen OCW gekeken naar de mogelijkheid van verrijkte teruglevering op individueel niveau en de benodigde grondslag daartoe. Als dit mogelijk is en er een passende grondslag wordt geformuleerd dan biedt dit de mogelijkheid om meer detailinformatie terug te leveren aan de instellingen.

Uitgangspunt 2: Een dashboard is aanvullend en vervangt de databehoefte niet

Data heeft waarde als het leidt tot inzichten. Een dashboard kan helpen om inzicht te verschaffen in complexe datavraagstukken. Het is een goede manier om gemiddelde gebruikers te helpen inzicht te vinden. Vaak is het niet mogelijk om de onderliggende data ook uit het dashboard te halen en te hergebruiken/combineren met andere bronnen. Het heeft dan ook de voorkeur om dashboards als complementair te zien en niet als supplementair op de dataleveringen.

Uitgangspunt 3: we houden rekening met volwassenheidsverschillen

Niet iedere instelling is op een gelijke manier bezig met datagedreven onderzoek en -werken. Daarnaast zijn er ook binnen instellingen grote verschillen tussen domeinen, locaties of teams. Houd hier rekening mee in de keuze van de dataleveringen. Bij een beperkte volwassenheid is actie nodig vanuit zowel de “aanbieder” als de “afnemer”. Vanuit DUO vraagt dit om meer interpretatie en voorbereidend werk. Dit kan bijvoorbeeld door te kijken naar een goede uitleg, heldere bestandsindelingen en structuren, maar ook door voorbeelden of dashboards te bieden

op de databestanden. Vanuit de instellingen vraagt dit om aandacht om databestanden te doorgronden en door het kennis- en vaardigheidsniveau van haar medewerkers te verbeteren.

Uitgangspunt 4: wat we doen is voor de gehele sector

De insteek van de datacoalitie datagedreven onderzoek is om met en voor de mbo-sector te kijken hoe we datagedreven onderzoek en -werken verder kunnen brengen. Dit doen we voor de gehele sector wat betekent dat de resultaten relevant (kunnen) zijn voor de hele sector en de resultaten ook beschikbaar worden gesteld voor de sector. Als het resultaat is dat een datalevering wordt gestart (open of achter de pin) dan wordt deze datalevering dus ook beschikbaar voor de gehele sector.

Hoofdstuk 3. - Vervolgactiviteiten

Op basis van de bevindingen en uitgangspunten heeft de werkgroep een activiteitenplan opgesteld. Dit plan is opgedeeld in 4 stappen met daarin een serie uit te voeren activiteiten.

De werkgroep heeft in sessies de informatiebehoefte van data geïnventariseerd. Hierbij is de prioriteit aangegeven 1) instroom in het mbo, 2) doorstroom naar hbo, 3.) doorstroom binnen mbo. Op basis van deze prioriteiten zullen diverse (open) databestanden worden gemaakt. Ten aanzien van prioriteit 1 en 2 gaat het om stromenbestanden. Voor alle 3 de prioriteiten worden onderstaande stappen doorlopen.

Stap 1. Vertrekpunt vaststellen en toetsen

- Geïnterpreteerde informatiebehoefte toetsen op sectorniveau. Nalopen en vastleggen van definities voor nieuw te leveren gegevens
 - o Sectorbreed inventariseren: welke indeling benadert een team?
- Toetsen voorstel voor een passende, uniforme set dimensies om databestanden te kunnen koppelen. Het voorstel van de werkgroep is opgenomen onder bijlage 1.
 - o Onderzoeken of de opgestelde “minimale” set dimensies overeenkomt met de sectorbrede wens (MBO Raad)
 - o Nagaan via de MBO Raad of het wenselijk is om de dimensie-structuur redundant te leveren (dwz het laagste en alle bovenliggende niveaus)
 - o De mogelijkheden van RIO onderzoeken om data te kunnen leveren op niveau van de locatie & onderwijsaanbieder
 - o Definitieve set van gegevens vaststellen

Stap 2. Dataleveringen op orde brengen

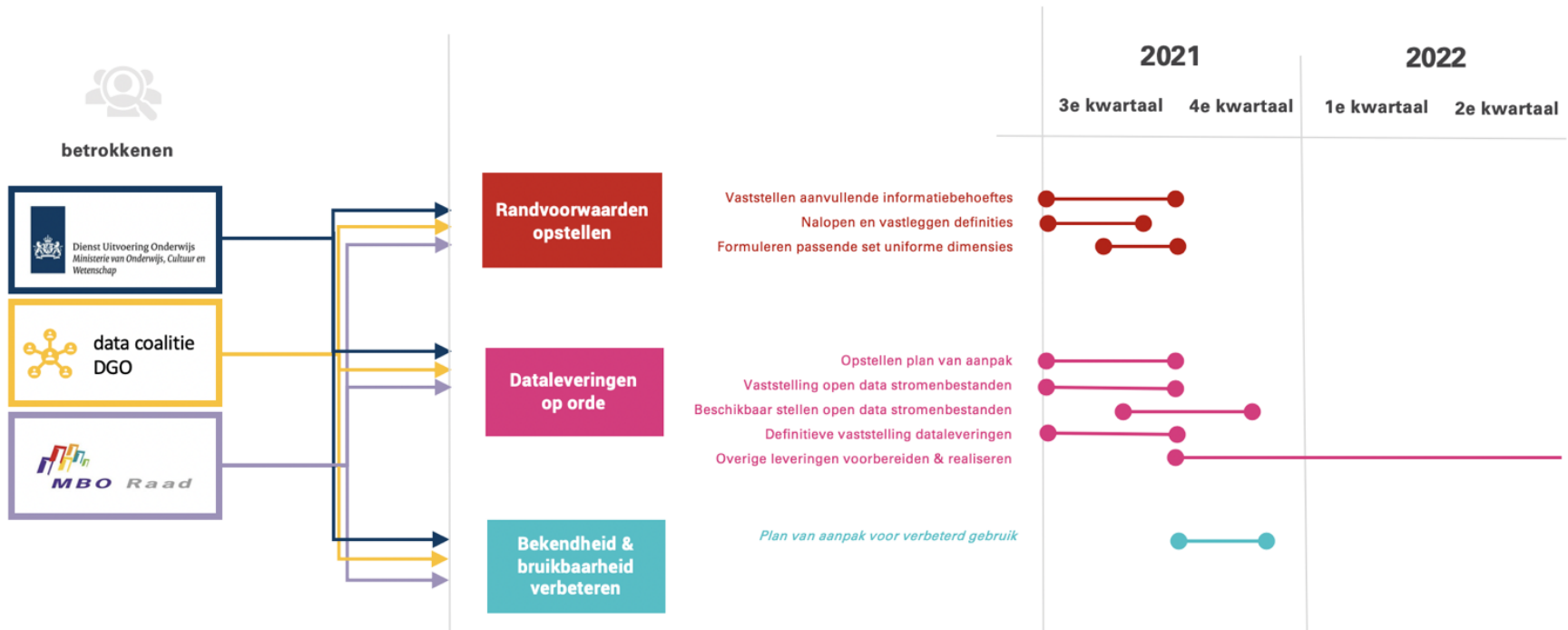
- Opstellen van plan van aanpak om de vastgestelde wensen om te zetten in (aanpassing van) dataleveringen. Daarbij passen onder andere de vragen:
 - o Past de geformuleerde datalevering bij de randvoorwaarden mbt open data?

- Kunnen de huidige / bestaande databestanden aangepast worden naar de gevraagde stamgegevens? Volstaat dat niet dan dienen er wellicht aparte databestanden voor worden ontwikkeld. De voor- en nadelen van vervanging dienen te worden afgewogen tegen de introductie van nieuwe bestanden.
- Op welke wijze kan de wens om de differentiëren naar onderwijslocatie & aanbieder vorm krijgen in deze datalevering?
- Definitieve vaststelling van de dataleveringen
 - Definities
 - Frequentie en planning
 - Formaat
 - AVG-voorwaarden / anonimiseren
- Dataleveringen plannen, voorbereiden en realiseren

Stap 3. Bruikbaarheid en bekendheid van de databronnen verhogen

- Passende planning opstellen: wanneer hebben de instellingen de databestanden nodig?
- Opstellen plan van aanpak om databekendheid te verbeteren. Bijvoorbeeld middels het realiseren van een leeswijzer voor de databronnen of (online) bijeenkomsten:
 - Wat zijn relevante bronnen om data/informatie te vinden?
 - Hoe verhouden de bronnen zich ten opzichte van elkaar?
 - Wat zijn de onderliggende definities?
 - Hoe kan ik de bron gebruiken?

Conceptplanning dataleveringen



Figuur 2.) Voorlopige planning en betrokken partijen per stap.

Hoofdstuk 4. - Integratie met lopende initiatieven

Er is een aantal bestaande initiatieven die kunnen bijdragen aan het realiseren van bovenstaande doelen. De werkgroep stelt voor om hier actief de aansluiting bij te zoeken. De samenhang dient verder te worden onderzocht met het oog op de toekomst.

- **Introductie & invoer RIO**

Het Register Instellingen en Opleidingen (RIO) is sinds anderhalf jaar in gebruik in de mbo-sector. Er zijn diverse nieuwe en gewijzigde velden in het register. De belangrijkste toevoegingen in RIO in deze context zijn de onderwijslocatie en onderwijsaanbieder.

- Onderwijslocaties zijn gekoppeld aan de Basisregistratie Adressen en Gebouwen, de BAG. Zo kunnen er nooit verkeerde of niet-bestaande adressen in RIO staan. De onderwijslocatie is een code van een instelling, bestaand uit 7 karakters, die koppelbaar is aan adres, postcode en plaats (voorbeeld van Albeda: onderwijslocatie 100X749, met daarbij adres Europa 1, 2672ZX, Naaldwijk).
- De onderwijsaanbieder is een code, bestaand uit 7 karakters, en is te koppelen aan een herkenbare naam van de onderwijsaanbieder (voorbeeld van Albeda: onderwijsaanbieder 101A442, met daarbij de naam van de onderwijsaanbieder Albeda Zorgcollege). Doordat studenten in het mbo worden ingeschreven op de instelling en onderwijslocatie (en onderwijsaanbieder) is het mogelijk om inschrijvingen binnen een instelling te onderscheiden naar locatie.

Met de introductie van RIO komt de wens van instellingen in zicht om data beschikbaar te stellen op het niveau van onderwijslocatie en onderwijsaanbieder in plaats van het bovenliggende BRIN-niveau. Voorwaarde daarbij is dat de administratieve inschrijving van de student overeenkomt met de realiteit van de onderwijssituatie.

- **Centraal aanmelden (CAMBO) biedt nieuwe data**

Met de introductie van centraal aanmelden komt er, mogelijk, een nieuwe set gegevens beschikbaar voor onderwijsinstellingen rondom aanmeldingen. Vanuit de werkgroep is er interesse om te onderzoeken welke data CAMBO genereert en welke data, er op welke wijze, beschikbaar kan worden gemaakt voor de onderwijsinstelling.

- Binnen DUO is er een project herstructurering open duo data gestart dat kijkt waar er een verbeterslag te maken is (betere vindbaarheid van bestanden, AVG-proof maken, beter structureren) met betrekking tot de bestaande open databestanden. Dit initiatief is in een opstartende fase en is sectoroverstijgend, breder dan enkel de mbo-sector. Vanuit de werkgroep is het voorstel om namens de mbo-instellingen de gewenste gegevensset te bespreken, input te geven op de herziening van de open databestanden en feedback te leveren waar nodig. DUO is daarbij bezig om alle bestanden die er nu

beschikbaar te zijn te toetsen aan de AVG. Dit moet leiden tot een passende set qua zowel informatiebehoefte als qua privacy-voorwaarden. Met DUO wordt besproken dat de realisatie van dit activiteitenplan in samenspraak gebeurt met het project herstructurering open duo data, waar het de mbo- en stromenbestanden betreft.

- In de datacoalitie wordt samengewerkt met DUO, maar ook SBB en CBS hebben interessante datasets rondom met name uitstroom naar arbeid. Voorstel van de werkgroep is om in een vervolg te onderzoeken hoe deze datasets te gebruiken zijn voor de onderwijsinstellingen.

Hoofdstuk 5. - Organiseren van het werk

Nu een beeld is opgebouwd over de uit te voeren activiteiten kan worden bekeken op welke wijze de betrokken partijen de uitvoering samen kunnen organiseren.

Daarvoor dient een aantal stappen te worden doorlopen:

1. Bepalen definitieve scope op basis van prioriteit
 - a. Wat is de prioritering (1. Instroom, 2. Doorstroom HBO, 3. doorstroom binnen het mbo)
 - b. Inventariseren in hoeverre de aangepaste dataleveringen vanuit team stromen prio 1 & 2 al dekken
 - c. Vaststellen van de scope
2. Opstellen plan van aanpak met begroting en planning op basis van prioritering
3. Afstemmen organisatie van het werk; wie doet wat?
 - a. MBO Raad nodig bij
 - i. Randvoorwaarden opstellen (indelingen, definities et cetera)
 - ii. Bekendheid verbeteren van de databestanden
 - b. DUO nodig bij
 - i. Randvoorwaarden opstellen (indelingen, definities et cetera)
 - ii. Plan van aanpak
 - iii. Opschonen & nieuwe leveringen open data
 - iv. Nieuwe leveringen achter de pin
 - v. Leveren van stromenbestanden
 - c. OCW nodig voor
 - i. Randvoorwaarden opstellen (indelingen, definities et cetera)
 - ii. Budget voor alle activiteiten van DUO
 - iii. Wettelijke basis van berekende data + terugleveringen
 - d. Data-coalitie nodig bij
 - i. Randvoorwaarden opstellen (indelingen, definities et cetera)
 - ii. Feedback / input op dataleveringen
 - iii. Bruikbaarheid & bekendheid: planning & leeswijzer opstellen. Optioneel kunnen we ook kennissessies organiseren om het veld verder te bedienen.
4. Vaststellen plan van aanpak, planning en financiering met opdrachtgever
5. Uitvoeren plan van aanpak conform taakverdeling

Bijlage 1. Gewenste basisset van gegevens per informatiebehoefte

Dimensie	Velden	Macro doelmatigheid	Instroom in het mbo	Doorstroom binnen de mbo-sector	Doorstroom HBO	Uitstroom Arbeid	Student juiste plek
Organisatie	Brinnr Instelling	x	x	x	x	x	x
	Instellingsnaam	x	x	x	x	x	x
	Locatie Instelling (RIO)	x	x	x	x	x	x
Opleiding	mbo Domein	x	x	x	x	x	x
	mbo Subgroep	x	x	x	x	x	x
	BC-code	x	x	x	x	x	
	BC-naam	x	x	x	x	x	
	mbo-niveau	x	x	x	x	x	x
	Crebo-code	x	x	x	x	x	x
	Crebo-naam	x	x	x	x	x	x
	Leertraject (BOL, BBL, OVO, ODT, EX)		x	x	x		x
	Diploma MBO			x	x	x	x
Student	Geslacht		x	x	x	x	x
	Leeftijdscategorie		x	x	x	x	x
Locatie / Geo	Provincie student	x	x				
	Arbeidsmarktregio	x					
	Gemeentenaam student	x	x				
	Gemeentecode student	x	x				
Voor-opleiding	Type vooropleiding (VO, MBO, overig)		x				
	Brinnummer leverende school		x				
	Opleiding op het VO		x	x			x
	Leerweg op het VO		x	x			x
	Laatste jaar op het VO			x			x
	Sector op het VO		x	x			x
	Laatste opleiding MBO			x			x
	Laatste bekende MBO niveau			x			x
	Laatste jaar op het mbo			x			x
	Hoogste vooropleiding vo						
Hoogste vooropleiding mbo							

Vervolg- opleiding	Brinnummer HBO				x		
	Brinvolgnummer HBO				x		
	Onderdeel				x		
	Subonderdeel				x		
	Croho-code				x		
	Croho-naam				x		