



> Retouradres Postbus 20011 2500 EA Den Haag

Ministerie van Sociale Zaken en Werkgelegenheid
T.a.v. de minister, de heer drs. W. Koolmees
Postbus 90801
2509 LV Den Haag

Bureau ICT-toetsing

Muzenstraat 95
Den Haag
Postbus 20011
2500 EA Den Haag
www.bureauicttoetsing.nl

Contactpersoon

BIT@rijksoverheid.nl

Kenmerk

2020-0000581495

Uw kenmerk

2018-0000147370

Datum 6 oktober 2020
Betreft Definitief BIT-advies project Datafabriek

Geachte heer Koolmees,

U heeft het Bureau ICT-toetsing (BIT) verzocht een toets uit te voeren op het project Datafabriek. De opdrachtgever is de Directeur Divisie Gegevensdiensten van het Uitvoeringsinstituut Werknemersverzekeringen (UWV). Hieronder vindt u een korte beschrijving van het project. Daarna geven we de conclusie van de toets, onze analyse en adviezen. Wij concentreren ons hierbij op de belangrijkste risico's van het project.

UWV beoogt een betere dienstverlening te bieden door haar toekomstige bedrijfsvoering te baseren op meer datagedreven werken. Het project Datafabriek maakt hiervoor het Data Integratie Magazijn (DIM). Met behulp van dit nieuwe datawarehouse kunnen sturings- en verantwoordingsrapportages (informatieproducten) worden gemaakt die data uit meerdere applicaties combineren. Het DIM gaat gebruikt worden door alle divisies en directies binnen UWV en zal tussen de 1.200 en 1.500 rapportages voorzien van informatie.

Het DIM moet drie bestaande datawarehouses gaan vervangen: DWH 1.0, DWH 2.0 en DWH 3.0. Van deze drie is DWH 2.0 significant de grootste. UWV geeft voor deze vervanging de volgende redenen:

- voorkomen van continuïteitsproblemen, omdat support voor de technologie van DWH 3.0 eind 2021 eindigt, en omdat de kennis en documentatie van DWH 1.0 onvoldoende is;
- kunnen voldoen aan de AVG;
- beperken van het risico op falende IT, door een vereenvoudigd IT-landschap;
- meer flexibiliteit bij het combineren van data en maken van rapporten.

UWV heeft eerder een project uitgevoerd voor het vervangen van DWH 1.0 en DWH 2.0 door DWH 3.0. Dit project duurde veel langer dan verwacht en is in 2011 gestaakt voordat DWH 1.0 of DWH 2.0 konden worden vervangen.

Het project Datafabriek is gestart in 2017. In fase 1 tot en met 3 (tot aan april 2020) is een nieuw data-integratieplatform geselecteerd, te weten InfoSphere van IBM. Ook is een ontwerp gemaakt hoe gegevens in stappen moeten worden ingelezen en verwerkt.

In 2019 is UWV gestart met het programma DIAscoop. Drie activiteiten van dit programma zijn cruciaal voor project Datafabriek: de regie voeren op het in kaart brengen van de informatiebehoefte van de afnemers, het ontsluiten van de brondata en het monitoren van de ontwikkeling van het metadatamodel door de divisie Gegevensdiensten. DIAscoop loopt tot en met 2023.

De resterende fases van het project zijn bedoeld om eerst DWH 3.0 uit te faseren (fase 4 en 5) en daarna DWH 1.0 en DWH 2.0 (fase 6 en 7). Fase 4 is begin april 2020 van start gegaan en loopt door tot eind 2020; de andere fases duren elk één kalenderjaar. Het project is momenteel gebudgetteerd op € 30 miljoen. Tot en met fase 3 is € 8 miljoen uitgegeven. Het programma DIAscoop heeft een eigen begroting van € 5 miljoen¹.

De BIT-toets is uitgevoerd tussen januari en mei 2020. De conclusie ervan luidt als volgt:

Het UWV werkt al geruime tijd aan een nieuw datawarehouse. Desondanks verwachten we dat het nog lang zal duren voordat dat in gebruik wordt genomen en een eerste bestaand datawarehouse kan worden ontmanteld. Dit komt doordat het project geen sturingsmogelijkheden heeft op het aansluiten van bronsystemen en het omzetten van rapportages. Bovendien is er onvoldoende grip op de voortgang en blijft het project te lang hangen in de ontwerpfase.

Wij adviseren daarom het project te voorzien van de benodigde bevoegdheden om te kunnen sturen op afhankelijkheden en daarmee op resultaat. Verander bovendien de ontwikkelaanpak en maak daarbij ontwikkelteams integraal verantwoordelijk voor oplevering van informatieproducten voor afnemers.

Wij lichten onze conclusie hieronder toe.

NOG GEEN ZICHT OP INGEBRUIKNAME NIEUW DATAWAREHOUSE

Het is goed te begrijpen dat UWV bestaande datawarehouses wil vervangen, gezien de continuïteitsrisico's, de AVG-eisen die moeten worden ingevuld en de redundantie in de bestaande datawarehouses. Ook is de keuze om per informatieproduct te migreren verstandig omdat zo stapsgewijs de doelstellingen kunnen worden gerealiseerd.

We verwachten evenwel dat het nog lang gaat duren voordat het UWV het DIM kan gebruiken. Na ruim 2,5 jaar is nog geen enkel nieuw informatieproduct aan een afnemer ter beschikking gesteld. Ook zien we nauwelijks maatregelen waarmee het project wel resultaat oplevert. Het eerder door UWV zelf gesignaleerde risico dat er een vierde datawarehouse bij komt in plaats van dat vereenvoudiging in het systeemlandschap wordt bewerkstelligd, wordt steeds meer realiteit².

¹ Overgenomen uit het programmaplan DIAscoop van 28 augustus 2019.

² CIO-advies UWV, d.d. 15 juli 2019 en 14 januari 2020.

Dat het nog lang gaat duren heeft drie oorzaken:

A. Te weinig sturingsmogelijkheden

Er is te weinig regie door het project, de stuurgroep en het programma DIAscoop op de activiteiten die nodig zijn om het DIM in gebruik te nemen en bestaande datawarehouses uit te zetten. Hiertoe moeten bronsystemen volledig worden aangesloten en bestaande rapportages worden omgezet. Dit aansluiten en omzetten valt echter niet onder de verantwoordelijkheid van het project dat hiervoor afhankelijk is van activiteiten van DIAscoop en van werkzaamheden die plaatsvinden in de lijn. Het project heeft daar te lang in berust en dreigt zonder stevige ingrepen een voortmodderend project te worden waarbij Datafabriek, DIAscoop en de lijnorganisaties naar elkaar kunnen gaan wijzen.

Het gaat om de volgende afhankelijkheden:

- Het is onbekend wanneer welke bron wordt aangesloten en het project heeft daarop geen invloed. In fase 4 en 5 moeten 44 aansluitingen van 13 bronsystemen en in fase 6 en 7 nog eens 121 aansluitingen van de overige 23 bronsystemen worden opgeleverd. Desondanks hebben het project en de onderhoudsteams van de 36 bronsystemen niet afgesproken wanneer deze 165 aansluitingen op het DIM worden gerealiseerd door de onderhoudsteams³.
- Het is onbekend wanneer de bestaande 1.200 rapporten worden aangepast en het project heeft daarop geen invloed. Circa 80 procent van de bestaande rapporten moet door de afnemers worden omgezet naar het DIM. Noch door het project noch door DIAscoop zijn afspraken gemaakt met de afnemers over planning en uitvoering van de hiervoor benodigde werkzaamheden.
- Het project heeft de invulling van diverse AVG-eisen buiten scope geplaatst. Ondanks dat het UWV aangeeft de voorgeschreven procedures te volgen, gaat DIM niet alle gestelde AVG-eisen invullen. Dit betreft:
 - *Het anonimiseren van persoonsgegevens waar nodig.* De functionaris gegevensbescherming eist dat persoonsgegevens geanonimiseerd worden vastgelegd in het DIM. Het project realiseert echter alleen een voorziening om gegevens gepseudonimiseerd op te slaan. Het project heeft nagelaten om met de functionaris gegevensbescherming af te spreken welke extra procedurele maatregelen nodig zijn.
 - *Het kunnen verwijderen van gegevens bij afloop van bewaartermijnen.* Het UWV moet rekening houden met wettelijke bewaartermijnen. Het is mogelijk om persoonsgegevens gepseudonimiseerd langer te bewaren voor statistische verwerking. Maar ook aan die termijn komt een eind. Juist omdat de gegevens mogelijk herleidbaar zijn naar individuen, zal DIM functionaliteit moeten krijgen om deze gegevens te verwijderen (lang nadat ze al uit het bronsysteem zijn verwijderd). Deze functionaliteit is geen onderdeel van het huidige ontwerp.

³ Zie Faseplan Vernieuwing IV middelen en inrichting Datafabriek – Fase IV van 24 april 2020, versie concept.

B. Er is onvoldoende grip op de voortgang

De voortgang kan onvoldoende worden bewaakt. Dat komt door het ontbreken van een realistische planning met voldoende meetmomenten. Zo blijven opgelopen vertragingen lang onbesproken en wordt er niet tijdig bijgestuurd. Dat de planning tekort schiet leiden wij af uit het volgende:

- *Het ontwerp en de technische realisatie van het DIM zijn aanzienlijk vertraagd.* Fase 3 heeft niet opgeleverd wat was gepland en dit is pas laat signaleerd. Alle activiteiten betreffende het ontwerp en de technische realisatie van het DIM zijn doorgeschoven naar fase 4.
- *De OTAP⁴-straat is nog niet beschikbaar en de realisatie is vertraagd.* De OTAP-straat is nodig om informatieproducten in gebruik te kunnen nemen en bronnen aan te kunnen sluiten en is nog niet gereed. Er zijn nog geen afspraken gemaakt met de leverancier over de planning.
- *Inzicht in de hoeveelheid werk is onvoldoende om een realistische planning op te baseren.* Er zijn veel kengetallen verzameld maar er is nog geen goede raming van de benodigde inspanning opgesteld. De inspanning voor het maken van scripts om data in te lezen, te vertalen en te pseudonimiseren is niet ingeschat. Bovendien zijn de doorgeschoven activiteiten uit fase 3 nog niet verwerkt in de planning. Wij verwachten dat er nog zeker zes jaar nodig is om de bestaande datawarehouses uit te faseren.
- *De planning gaat zonder onderbouwing uit van een toename van de ontwikkelsnelheid.* Wij zien de ontwikkelsnelheid afnemen bij vier van de vijf teams. Dat komt mede doordat de projectaanpak is gebaseerd op losse teams die wachten op elkaars halffabrikaten. Deze versnipperde aanpak met veel overdrachtsmomenten leidt tot inefficiënt werken.
- *De planning houdt geen rekening met de datacentrummigratie.* De verhuizing van het DIM van het datacentrum van IBM naar het nieuwe van DXC moet uiterlijk in 2024 zijn afgerond. Met de huidige ontwikkelsnelheid voorzien we dat die verhuizing plaats moet vinden gedurende het project. Wij verwachten dat dit nieuwe vertragingen oplevert.

C. Door verkeerde focus in ontwerpfase moet lang gewacht worden op bruikbare resultaten

Het realiseren van één nieuw datawarehouse voor heel UWV is per definitie een omvangrijke operatie: het gaat over veel gegevens, gegevensleveringen, rapporten en een uitdagend uniformeringsvraagstuk. We begrijpen dat UWV kiest voor een indeling van het DIM in zones omdat dat een gangbare indeling is. Maar de aanpak van het project is gericht op het op voorhand gedetailleerd ontwerpen en nadien realiseren van volledige generieke technische deeloplossingen. Hierdoor duurt het lang voordat het project resultaat voor eindgebruikers levert. Bovendien lukt het ook niet goed. Dit blijkt uit het volgende:

- Een voldoende uitgewerkt globaal ontwerp van de zone-indeling ontbreekt, evenals de technische uitgangspunten voor schaalbaarheid en foutafhandeling waaraan de zones en de te realiseren bouwblokken moeten voldoen.

⁴ Een OTAP-straat omvat alle omgevingen die nodig zijn om een systeem te kunnen ontwikkelen, testen, accepteren en uiteindelijk in productie te nemen.

- Slechts het ontwerp van één van de vier zones, de bronzone, is beschikbaar. Voor de ontsluitingszone zijn wel een aantal technische zaken gebouwd, maar een ontwerp daarvoor is niet aanwezig.
- Er wordt veel tijd besteed om bouwblokken in detail uit te werken, waarbij maximale herbruikbaarheid (genericiteit) wordt nagestreefd zonder dat vaststaat dat de extra inspanning die daarvoor nodig is de moeite waard is. Het is nog onzeker in welke mate en wanneer van deze bouwblokken gebruik gemaakt gaat worden. Ook wordt niet tijdig gevalideerd of deze bouwblokken inzetbaar zijn zoals gepland.
- Er wordt een gegevensmodel voor het DIM nagestreefd en uitgewerkt dat is voorbereid op alle mogelijke situaties en uitzonderingen. Wij vragen ons af of dat de inspanning op dit moment waard is.

ADVIES: HERSTRUCTUREER HET PROJECT

Wij adviseren om alle activiteiten die essentieel zijn voor het projectsucces onder aansturing te brengen van het project Datafabriek, zodat UWV meer grip krijgt op de voortgang. Bovendien adviseren wij de ontwikkelaanpak aan te passen en te richten op het frequent opleveren van bruikbare informatieproducten, zodat het project sneller waarde gaat leveren. Onze aanbevelingen lichten we hieronder verder toe.

1. Zorg dat gestuurd kan worden op essentiële afhankelijkheden en daarmee op resultaat

Tref maatregelen waarmee het project daadwerkelijk de verantwoordelijkheid voor het uitfasen van de bestaande datawarehouses kan waarmaken. Zorg dat het project bevoegdheden krijgt waarmee het tijdig aansluiten van bronsystemen en het tijdig omzetten van rapportages voor eindgebruikers kan worden afgedwongen. Bereid zo nodig de stuurgroep uit met leidinggevendenden uit de divisies die op dat moment betrokken zijn bij het aansluiten van bronsystemen of het omzetten van rapportages.

Bepaal samen met de functionaris gegevensbescherming welke mitigerende maatregelen nodig zijn om aan de AVG te kunnen voldoen. Maak afspraken over wie daarvoor verantwoordelijk is.

2. Verander de ontwikkelaanpak, en pas planning en sturing daarop aan

Richt in dat snelle, kort-cyclische oplevering van de informatieproducten op basis van 'goed genoeg' bouwblokken kan worden gerealiseerd. Pas de ontwikkelaanpak aan. Stop met het eerst volledig generiek uitwerken van bouwblokken. Wacht niet met het valideren van de werking. Op basis van deze aanpak kan een planning worden gemaakt waarop beter gestuurd kan worden. Doe daartoe het volgende. Stel de volgende richtlijnen vast:

- Stop met het werken in teams die elkaar halffabrikaten doorgeven. Organiseer de teams naar (groepen van) informatieproducten. Elk team dient eenduidige verantwoordelijkheid te krijgen voor het opleveren van

informatieproducten op het DIM en het buiten gebruik stellen van informatieproducten op bestaande datawarehouses.

- Ontwerp en ontwikkel alleen de bouwstenen zoals ze op dat moment nodig zijn. De ontwikkelde bouwstenen kunnen nadien worden aangepast voor het gebruik bij andere informatieproducten.

Wij bevelen de volgende acties aan:

- Laat zien dat project Datafabriek voor einde van het jaar wel degelijk resultaten kan leveren door zorg te dragen dat:
 - de OTAP-straat op de kortst mogelijke termijn beschikbaar is;
 - de uitgangspunten voor schaalbaarheid en foutafhandeling op de kortst mogelijke termijn beschikbaar zijn;
 - de aansluitingen op de bestaande datawarehouses zonder aanpassingen worden overgezet naar het DIM, tenzij dit echt niet anders kan;
 - de eerste rapportage uiterlijk eind 2020 is geïmplementeerd.Lukt dat niet dan moet UWV zich afvragen of het project niet beter kan worden beëindigd.
- Maak één planning voor het uitfasen van DWH 3.0 en 1.0. Neem ook het migreren naar het datacentrum van DXC op in de planning.
- Laat het project periodiek voortgang rapporteren op percentage afgerond werk afgezet tegen bestede kosten. Grijp in indien substantieel wordt afgeweken van de planning.
- Beoordeel na het uitzetten van DWH 3.0 of er nog een zakelijke rechtvaardiging bestaat voor vervanging van het veel omvangrijkere DWH 2.0. Zo ja, doe dat dan als volgt:
 - Sta in beginsel geen wijzigingen meer toe voor DWH 2.0.
 - Geef prioriteit aan migratie van de informatieproducten uit DWH 2.0 die frequent worden aangepast.
 - Maak een planning voor de ontmanteling van DWH 2.0.
 - Voorzie in additionele financiering en houdt rekening met eventuele tegenvallers. Het uitfasen van DWH 2.0 is immers eerder onderschat.

* * *

Tot slot danken wij alle geïnterviewden voor hun medewerking en openheid bij deze toets. We hopen u met dit advies aanknopingspunten te hebben gegeven voor het vervolg.

Met de meeste hoogachting,
namens het Bureau ICT-toetsing,

Sander van Amerongen
wnd. Hoofd BIT