**Adviesnotitie verbeteringen t.b.v. bronontsluitingsproces DIM**

1. **Inleiding** 
   1. **Opdracht**

De programmamanager Datafabriek heeft naar aanleiding van diverse vertragingen bij de migratie van de bronnen UZS en Excasso in september 2023 aan de projectleider Datafabriek-bronontsluitingen gevraagd inzicht te geven in de Non Value Adding activiteiten (verspillingen) van het bronontsluitingsproces. Aanvullend is aan de projectleider gevraagd een adviesvoorstel op te stellen met hoe de totale doorlooptijd (Lead Time) voor het ontsluiten van een bron kan worden verkort. Om tot een beantwoording van de adviesvragen te komen zijn de volgende deelvragen geformuleerd:

1. *Hoe ziet het huidig DIM bronontsluitingsproces er uit?*
2. *Wat zijn de hoofdoorzaken van de verspillende activiteiten (Root Cause Analyse)?*
3. *Welke maatregelen en verbeteracties leiden tot de grootste reductie van Lead Time?*
4. *Welk advies voor verkorting leadtime bronontsluiting DIM kan worden opgesteld?*

**1.2 Aanpak**

Om tot een realisatie van de gestelde opdracht te komen, is vanuit de gedachte ‘Getting the whole system in the room’, een werkgroep samengesteld met alle stakeholders. De werkgroep heeft tijdens de Value Stream Mapping (VSM) sessies de volgende aanpak gehanteerd:

* In kaart brengen van huidig bronontsluitingsproces inclusief cycle time per activiteit;
* Identificeren van Non Value Adding activiteiten (verspilling)
* Uitvoeren van Root Cause Analyse (de hoofdoorzaak van een verspillende activiteit achterhalen)
* Bepalen welke hoofdoorzaak leidt tot de grootste reductie van Lead Time
* Bepalen welke maatregelen en verbeteracties leiden tot het gewenste resultaat;
* Prioriteren van de maatregelen en verbeteracties die tot grootste reductie van Lead Time leiden;
* Formuleren van advies op basis van geprioriteerde maatregelen en verbeteracties

Parallel aan bovenstaande aanpak hebben zowel DWH Gegevensdiensten als de projectleider Datafabriek-bronontsluitingen inzichtelijk gemaakt wat de mogelijke oorzaken zijn van de issues met bronleveringen.

Dit document geeft antwoord op de gestelde vragen van de programmamanager Datafabriek. Het gevraagde adviesvoorstel is uitgewerkt.

**2. Hoofdadvies: Welk advies voor verkorting leadtime bronontsluiting DIM kan worden opgesteld?**

Op basis van de verbeteracties en maatregelen die aan de orde zijn geweest in deelvraag 3 (hoofdstuk 3) zijn ter verbetering (lees: verkorting) van de leadtime van het DIM bronontsluitingsproces een zevental aanbevelingen naar voren gekomen. Deze zijn in twee categorieën te verdelen:

1. Oorzaken met een oplossing die op korte termijn (binnen enkele maanden) en voornamelijk met inspanning door teams binnen het programma kan worden gerealiseerd, zonder grote impact op andere geplande activiteiten
2. Oorzaken met een oplossing die een besluit van stuurgroep vragen, en acties door management / divisies buiten het programma

Per aanbeveling is een expert inschatting gegeven van de versnelling op de totale doorlooptijd van het bronontsluitingsproces door de PO van de Datafabriek bronnenteam: Medium (10-15%), Groot (circa 20%) Zeer groot (≥30%).

**2.1 Aanbevelingen - Oplossing binnen invloedssfeer programma Datafabriek**

Aanbeveling 1 – Datakwaliteitsproblemen eerder in het proces signaleren

|  |  |
| --- | --- |
| **Grondoorzaak** | Datakwaliteitsproblemen en afwijkingen t.o.v. specificaties in RLO worden pas erg laat in het proces geconstateerd (nu pas bij het laden van gemaskeerde productiedata op acceptatie omgeving) |
| **Consequentie** | Meerdere langdurige iteraties met rework n.a.v. deze bevindingen laat in het proces die leiden tot vertraging. Deze bevindingen zijn al eerder te constateren en dan vroeger te verhelpen met minder herhaling van zetten. |
| **Advies** | Controles of productiedata aan RLO voldoet op basis van geanonimiseerde subset van data (die buiten harnassen om wordt geïmporteerd in een apart schema voor analyse) direct na definitief maken RLO. |
| **Op te volgen door:** | Datafabriek bronontsluiting team i.s.m. enabling |
| **Impact / versnelling** | Groot |

Aanbeveling 2 – RLO document interne DIM deliverable maken

|  |  |
| --- | --- |
| **Grondoorzaak** | Het RLO, dat wordt ingelezen om m.b.v. harnassen data te kunnen laden in DIM, bevat naast informatie over het gegevensmodel van de bron ook technische specificaties voor verwerking en aanlevering, die alleen relevant zijn voor DIM, en waarvan bron zich geen eigenaar voelt. |
| **Consequentie** | De afdelingen verantwoordelijk voor de bronapplicaties (divisies) hebben onvoldoende kennis van deze datawarehouse specifieke zaken. Het aanleveren van deze informatie in een definitief RLO gaat daardoor moeizaam en met lage kwaliteit. Verkeerd invullen van deze datawarehouse specificaties leiden tot rework later in het proces als de configuratie moet worden aangepast. |
| **Advies** | RLO tot interne deliverable maken van DIM / Datafabriek. Als interne deliverable wordt het RLO dan samengesteld uit:   1. FUGEM-TEGEM extract vanuit de bronafdeling, geautomatiseerd (zie volgende aanbeveling, hier bestaat al een script voor). Bron levert dit extract bij iedere release met impact op DIM aanlevering. 2. Specificatie van technische sleutels en parameters voor verwerking. Opgesteld door informatie analist DIM i.s.m. deskundigen van de bronapplicatie. Bron geeft definitief akkoord op deze (in principe niet wijzigende) specificaties. |
| **Op te volgen door:** | Datafabriek i.s.m. lijn en bronafdelingen |
| **Impact / versnelling** | Middel |

Aanbeveling 3 – Script voor export metadata uit FUGEM/TEGEM optimaliseren

|  |  |
| --- | --- |
| **Grondoorzaak** | Huidige script om de metadata in RLO geautomatiseerd te vullen op basis van FUGEM/TEGEM heeft tekortkomingen, waardoor nog handmatige acties nodig zijn om het RLO volledig en juist te maken. |
| **Consequentie** | Handmatige aanpassingen na genereren m.b.v. script leiden tot versieproblemen en fouten in de verwerking. Informatie benodigd voor DIM verwerking kan niet volledig uit FUGEM/TEGEM worden gehaald. Consequentie is een extra verzoek tot aanpassing RLO, foutgevoelig proces met wachttijd. |
| **Advies** | Verbeteringen voor first-time-right vullen van RLO inventariseren uit recente ervaringen, en script laten aanpassen zodat alle input over gegevensmodel m.b.v. script kan worden aangeleverd, zonder nabewerking. In aanvulling hierop wordt ook een oplossing uitgewerkt om middels een GUI met user/audit om metadata te beheren. |
| **Op te volgen door:** | Data office, bronontsluiting team Datafabriek en solution architect |
| **Impact / versnelling** | Middel |

Aanbeveling 4 – Verkorten van draaitijden harnassen op A en P

|  |  |
| --- | --- |
| **Grondoorzaak** | Lange draaitijden van harnassen op A en P omgeving (bijv. >2 uur per daglevering voor medium size bron WWF, één van de 30+ bronnen) |
| **Consequentie** | Stuwmeer van te verwerken data op P. Onvoorspelbare planning voor in beheer nemen van opgeleverde bronnen. Veel tijdverlies bij het stuklopen van een job (herstarten van harnassen na een stop is risicovol, alles moet dan opnieuw). |
| **Advies** | Optimaliseren van draaitijden op A en P m.b.v. efficiëntere Datastage jobs, database inrichting/partitionering en indexen, en efficiënter gebruik onderliggende hardware. De techniek van het inladen van data (DIMIF) moet opnieuw worden ontworpen om sneller te zijn. Hier zijn de architecten mee bezig. Hier moeten andere disciplines (o.a. Enabling team, INFRA en DBA ) |
| **Op te volgen door:** | Stuurgroep.  Actie is (al maanden) belegd bij architectuur. Meer aandacht van stuurgroep nodig om te sturen op concrete milestones en resultaten van deze actie. |
| **Impact / versnelling** | Groot (>20% versnelling)  Zeer grote impact op slagen van het DIM. |

**2.2 Aanbevelingen - Oplossing buiten invloedssfeer programma Datafabriek**

Aanbeveling 5 – Data huishouding op orde binnen de divisies

|  |  |
| --- | --- |
| **Grondoorzaak** | Data huishouding in divisies niet op orde: Het bijhouden van gegevensmodellen (FUGEM en TEGEM) van bronapplicaties is onvoldoende belegd. Verantwoordelijkheid voor bijhouden van gegevensmodellen is niet belegd. Update FUGEM en TEGEM bij nieuwe releases van de bron is geen onderdeel van release proces. |
| **Consequentie** | DIM is meestal de eerste partij die hogere eisen aan FUGEM/TEGEM stelt, met lange opstarttijd tot gevolg. Divisies hebben het op orde brengen en houden van FUGEM en TEGEM niet op orde, zowel qua verantwoordelijkheid, proces als bemensing.  Ook wordt na oplevering en overdracht van een bron aan de lijn, het structureel up-to-date houden van FUGEM/TEGEM bij releases niet geborgd binnen de divisies. |
| **Advies** | Centrale sturing inrichten op datamanagement binnen de divisies. Checks op FUGEM/TEGEM en datakwaliteit. Bij releases ook de update van de FUGEM en TEGEM als harde eis.  Hygiëne van FUGEM-TEGEM onderdeel maken van een periodieke review, met dashboard over de divisies en consequenties bij niet op orde zijn van gegevensmodellen.  Gegevensmanager verantwoordelijk maken en plan-do-check-act inregelen op actualiteit FUGEM en TEGEM. |
| **Op te volgen door:** | Stuurgroep / management divisies |
| **Impact / versnelling** | Zeer groot (>30% versnelling) |

Aanbeveling 6 – Prioriteit DIM in PI planningen en change kalender divisies

|  |  |
| --- | --- |
| **Grondoorzaak** | Op tactisch niveau (product owners, PI planning, verandertafels) krijgen de activiteiten benodigd voor aansluiten op DIM lage prioriteit binnen de divisies. |
| **Consequentie** | Planning en tijdslijnen van bronontsluiting zijn zeer afhankelijk van opvolging van acties en verzoeken aan deskundigen bij de bron. Dit speelt vanaf het plannen van de kick-off t/m het akkoord op de acceptatie test. Medewerkers willen vaak wel maar moeten het “erbij doen”. Dit leidt tot onvoorspelbare tijdslijnen en uitloop. |
| **Advies** | Handshake op de inzet van deskundigen en geplande tijdslijnen per divisie, met commitment en hulp van management. Hier is een template voor opgesteld door PO Bronontsluiting DIM. |
| **Op te volgen door:** | Stuurgroep |
| **Impact / versnelling** | Middel (10-15% versnelling) |

Aanbeveling 7 – DevOps inrichting voor DIM

|  |  |
| --- | --- |
| **Grondoorzaak** | Scheiding “ontwikkelteam” en “beheerteam” leidt tot langere doorlooptijd en lagere kwaliteit en toekomstbestendigheid |
| **Consequentie** | Beheer voert draaiboek uit maar is niet betrokken bij de analyse en bouwtraject voorafgaand aan in beheer name. Deze kennis is wel nodig om incidenten te kunnen analyseren en vervolgreleases van bronnen te kunnen verwerken.  Omgekeerd doet het ontwikkelteam een hand-over naar beheer, en is niet verantwoordelijk voor het op lange termijn stabiel en goed lopen van techniek en proces.  Deze scheiding in verantwoordelijkheden tussen DEV en OPS leiden tot lage kwaliteit en problemen in beheren van DIM ontsluiting op langere termijn. Ook gezien de disbalans in bemensing (het werk van 10 FTE ontwikkelteam komt in productie nu terecht bij 1 FTE beheerder). |
| **Advies** | Focus op het inrichten van Dev/Ops teams, die naast het bouwen ook verantwoordelijk zijn voor het beheren van hun software. Hiermee wordt ook een goede overgang van DIM naar de lijn organisatie bevorderd. Om DEVops te laten werken moet ook naar het proces van beheer, change management en deployments (CI/CD) worden opgezet. |
| **Op te volgen door:** | Stuurgroep |
| **Impact / versnelling:** | Zeer groot (>30% versnelling) |

**3. Conclusies en aanbevelingen**

In dit hoofdstuk vormen de bevindingen uit de VSM sessies de basis voor de conclusies die worden getrokken en de aanbevelingen die worden gedaan. In paragraaf 3.1 worden de deelvragen beantwoord. Vervolgens worden de bevindingen, maatregelen en verbeteracties vertaald naar aanbevelingen/ adviezen voor het verkorten van de leadtime voor het ontsluiten van bronnen (Hoofdstuk 2).

**3.1 Beantwoording deelvragen**

**Deelvraag 1: *Hoe ziet het huidig DIM bronontsluitingsproces er uit?***

Het bronontsluitingsproces start bij het vastleggen van een bron die ontsloten moet worden op de backlog van programma Datafabriek tot en met het in beheer nemen (accorderen van de overdracht) van een ontsloten bron door DWH Gegevensdiensten.

De werkgroep heeft tijdens de eerste VSM sessie inzichtelijk gemaakt hoe het huidige DIM bronontsluitingsproces eruitziet (Bijlage 1). In onderstaand tabel zie je per activiteit/ deliverable de geschatte bewerkingstijd[[1]](#footnote-1) en Procestijd[[2]](#footnote-2) in de praktijk. Wanneer de Procestijd aan de lange kant is t.o.v. de bewerkingstijd houdt dat in dat in het processtap verspilling plaatsvindt. De *deliverables* zijn cursief gedrukt en de **processtappen met de meeste verspilling** zijn vetgedrukt.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Nr.** | **Processtap/ deliverable** | **Wie** | **Bewerkings- tijd** | **Procestijd** |
| 1 | Epic bronontsluiting op backlog DF | Projectmanager | 1 uur | 1 uur |
| 2 | Afgeven signaal bij bron | PM/PO | 1 uur | 1-2 wk |
| 3 | Kickoff houden met bron | IA/PO | 1-5 dag | 1-2 wk |
|  | *Planningsvoorstel met duidelijke afspraken (deliverable)* |  |  |  |
| 4 | Bronovereenkomst sluiten met bron | IA/PO | 1-2 dag | 1-2 wk |
|  | ***Requirementsdocument v0.1 (deliverable)*** |  | **1 wk** | **5 wk** |
| 5 | **Uitvoeren informatieanalyse** | **IA** | **3 wk** | **6 wk** |
|  | *Requirementsdocument (deliverable)* |  |  |  |
|  | *Acceptatiecriteria(deliverable)* |  |  |  |
|  | *RLO v0.9 (deliverable)* |  |  |  |
| 6 | Verkrijgen testset | Tester | 1 dag | 1 sprint |
| 7 | **Tot stand brengen stuurmetadata o.b.v. RLO op O** | **Configurator** | **1 wk** | **3 wk** |
| 8 | Uitvoeren tests op O | Configurator | 1-3 dag | 5 dagen |
| 9 | Genereren ontkoppelviews | Configurator | 1 dag | 1 dag |
| 10 | Samenstellen en installeren deployment package op O | Configurator | 1-2 dag | 5 dag |
| 11 | **IMAM en vertical lineage** | **IA** | **1-2 dag** | **1-2 wk** |
|  | *Ingerichte O omgeving (deliverable)* |  |  |  |
| 12 | Deployen op T en testdata laden op T | Configurator | 1-5 dag | 5 dag |
| 13 | Testen op T | Tester | 1-2 dag | 2 dag |
| 14 | Inrichten Trudi | Tester | 1-2 dag | 2 dag |
|  | *Ingerichte T omgeving (deliverable)* |  |  |  |
| 15 | Maken draaiboek deployment op A | Configurator | 1-2 dag | 2 dag |
| 16 | **Draaiboek uitvoeren op A** | **Configurator** | **1-5 dag** | **3 wk** |
| 17 | Testrapport maken | Tester | 1-2 dag | 2 dag |
|  | *Ingerichte A omgeving (deliverable)* |  |  |  |
|  | *Testrapport (deliverable)* |  |  |  |
| 18 | Maken draaiboek P | Configurator | 1-2 dag | 2 dag |
| 19 | **Draaiboek uitvoeren op P** | **Team Up!** | **1-5 dag** | **3 wk** |
| 20 | **Maken overdrachtsdocument** | **IA** | **1-3 dag** | **1-2 wk** |
|  | *Ingerichte P omgeving (deliverable)* |  |  |  |
|  | *Overdrachtsdocument (deliverable)* |  |  |  |
| 21 | Overdracht sessie plannen en houden | PO | 1 dag | 3 wk |
|  | *Akkoord overdracht (deliverable)* |  |  |  |
|  |  |  |  |  |

***Deelvraag 2: Wat zijn de hoofdoorzaken van de verspillende activiteiten (Root Cause Analyse). Welke hoofdoorzaak leidt tot de grootste reductie van Lead Time?***

Vervolgens heeft de werkgroep voor de processtappen met de meeste verspilling de oorzaken van verspilling aangegeven met behulp van TIMWOODS[[3]](#footnote-3)

Uit de VSM sessies zijn de volgende terugkerende problemen die tot vertraging leiden naar voren gekomen:

* Verschillende subtaken moeten worden uitgevoerd door verschillende teams, omdat elk team (DBA, DIM Beheer, Enabling, Bronontsluiting team) eigen en kleine verantwoordelijkheid en rechten op omgevingen heeft
* Divisies / bronnen hebben niet de kennis of niet de randvoorwaarden op orde (FUGEM/TEGEM) om in korte tijd benodigde specificaties voor DIM op te leveren
* Handmatige vulling van RLO's of nabewerking van RLO's leidt tot versieproblemen, fouten en rework
* Datakwaliteitsproblemen die niet vooraf bekend waren, leiden laat in het proces tot fouten en rework
* Aanleverende partijen (divisies, DXC, TDA, DWH) hebben andere prioriteiten / moeten het "ernaast" doen

Hierna heeft de werkgroep gebruik gemaakt van het Ishikawa diagram[[4]](#footnote-4) om inzichtelijk te maken wat de grondoorzaken (hoofdoorzaken) zijn van bovenstaande terugkerende problemen.

In onderstaande tabel zijn de belangrijkste grondoorzaken genoemd. De top 3 grondoorzaken (grijs gearceerd) zijn door de werkgroep d.m.v. dot voting bepaald. Het aanpakken van deze grondoorzaken leidt tot de grootste reductie van Lead Time.

| **Grondoorzaak** | **Korte omschrijving** |
| --- | --- |
| Aannames datakwaliteit pas laat in het proces getoetst met productiedata | Datakwaliteitsproblemen en afwijkingen t.o.v. specificaties in RLO worden pas erg laat in het proces geconstateerd (nu pas bij het laden van gemaskeerde productiedata op acceptatie omgeving) |
| RLO is 2 in 1 | Het RLO, dat wordt ingelezen om m.b.v. harnassen data te kunnen laden in DIM, bevat naast informatie over het gegevensmodel van de bron ook technische specificaties voor verwerking en aanlevering, die alleen relevant zijn voor DIM, en waarvan bron zich geen eigenaar voelt. |
| Script gebruikt voor RLO geeft handmatige nabewerking | Huidige script om de metadata in RLO geautomatiseerd te vullen op basis van FUGEM/TEGEM heeft tekortkomingen, waardoor nog handmatige acties nodig zijn om het RLO volledig en juist te maken. |
| Lange draaitijden op A en P | Lange draaitijden van harnassen op A en P omgeving (bijv. >2 uur per dag levering voor medium size bron WWF, één van de 30+ bronnen) |
| Basis niet op orde; FUGEM en TEGEM | Data huishouding in divisies niet op orde: Het bijhouden van gegevensmodellen (FUGEM en TEGEM) van bronapplicaties is onvoldoende belegd.  Verantwoordelijkheid voor bijhouden van gegevensmodellen is niet belegd. Update FUGEM en TEGEM bij nieuwe releases van de bron is geen onderdeel van release proces. |
| Beschikbaarheid ketenpartners | Op tactisch niveau (Product Owners, PI planning, verandertafels) krijgen de activiteiten benodigd voor aansluiten op DIM lage prioriteit binnen de divisies. |
| Vele hand-overs en scheiding ontwikkel en beheer | Scheiding “ontwikkelteam” en “beheerteam” leidt tot langere doorlooptijd en lagere kwaliteit en toekomstbestendigheid |

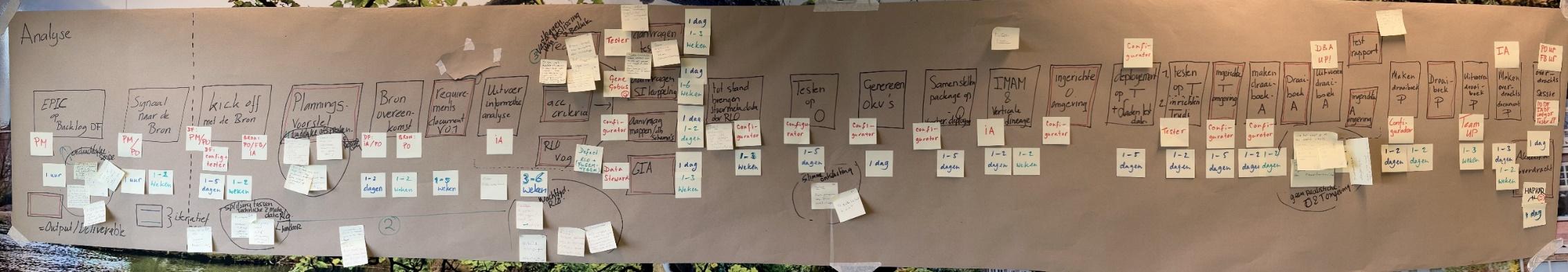
**Deelvraag 3 *Welke maatregelen en verbeteracties leiden tot de grootste reductie van Lead Time?***

Doormiddel van een brainstorm met de werkgroep hebben we voor elk van deze 3 grondoorzaken een aantal verbeteracties bedacht:

| **Grondoorzaak** | **Aangedragen verbeteracties** |
| --- | --- |
| 1. Basis niet op orde; Fugem en Tegem | Data office in de lead zetten en laten vastleggen welke FUGEMS en TEGEMS op orde zijn |
| FUGEM laten automatiseren |
| Toetsing van de benodigde onderdelen |
| Lik op stuk beleid |
| Projectteam FUGEM/TEGEM |
| Data officer per divisie, verantwoordelijk maken voor FUGEM, TEGEM en datakwaliteit |
| Gegevensmanager per divisie aan het werk zetten |
| 1. RLO is 2 in 1 | RLO uitkleden, minimale verwerking |
| RLO splitsen; metadata uit Fugem mbv script, parameters voor verwerking in ander bestand |
| Bron van tevoren goed uitleggen over RLO |
| Slim werken met RLO |
| Kritisch kijken naar maskeringsklasse, administratieve tijdslijnen, etc. Weet welke impact dit heeft. |
| 1. Beschikbaarheid ketenpartners | DF representatie bij PI-events ketenpartners |
| Prioriteitsstelling op juiste niveau en naar niveau vertaald |
| Resultaatgericht samenwerken |
| Eigenaarschap bevorderen |
| Projectteams opstellen over divisies heen |
| Dedicated mensen vrij laten maken per bron |
| Beter verkopen van DF/DIM bij het UWV |
| Commitment en naleving op tijdslijn bij de bron |
| Meer duidelijkheid over deadlines voor leveringen bij bron |
| Dataleveringen op tijd aanvragen |

Deze verbeteracties zijn vervolgens uitgewerkt in groepjes. De adviezen van deze groepjes zijn gepresenteerd aan de gehele groep en vervolgens verder uitgediept. In kleiner verband zijn de verbeteracties en adviezen concreter gemaakt en vertaald naar de zeven aanbevelingen, gekoppeld aan grondoorzaak en consequentie, die zijn beschreven in hoofdstuk 2.

**Bijlage 1 – Procesplaat bronontsluiting**



1. Bewerkingstijd is de werktijd die nodig is voor elke processtap. [↑](#footnote-ref-1)
2. De Procestijd is gemiddelde tijd die het duurt om een processtap te doorlopen. [↑](#footnote-ref-2)
3. TIMWOODS, is een ezelsbruggetje voor de 8 vormen van verspilling: Transport, Inventory, Motion, Wachten, Overproductie, Overprocessing, Defecten(Rework), Skills. [↑](#footnote-ref-3)
4. Ishikawa diagram is gebruikt om alle mogelijke oorzaken (problemen) visueel te rangschikken en een gestructureerde brainstormsessie te houden met de werkgroep. Met als doel komen tot de belangrijkste grondoorzaken. [↑](#footnote-ref-4)