



# Bijlage B - Architectuurkaders

*Europese aanbesteding IDP-software*

22 juli 2026

# Inleiding

In dit document worden de architectuorkaders benoemd voor de nieuwe Oplossing.

De Architectuureisen waaronder maar niet alleen de vertaling van architectuurprincipes en te gebruiken open standaarden zijn integraal opgenomen in het Programma van Eisen (PvE).

## Inhoud

|                                                      |           |
|------------------------------------------------------|-----------|
| <b>Inleiding.....</b>                                | <b>2</b>  |
| <b>1 Definitie en afbakening voorziening .....</b>   | <b>4</b>  |
| 1.1 Beschrijving van de scan- en herkenvoorziening   | 4         |
| 1.2 Afbakening van de voorziening                    | 4         |
| 1.3 Applicatiefunctionaliteit                        | 4         |
| <b>2 Implementatiewijze van de voorziening .....</b> | <b>6</b>  |
| 2.1 Dienstverleningsmodel en hosting                 | 6         |
| 2.2 Lokale software                                  | 6         |
| <b>3 Applicatielandschap .....</b>                   | <b>7</b>  |
| 3.1 Overzicht na as-is vervanging                    | 7         |
| 3.2 Koppelingen in de projectscope                   | 9         |
| 3.3 Potentiële koppelingen op termijn                | 10        |
| <b>4 AI.....</b>                                     | <b>14</b> |
| 4.1 AI versus traditionele technologieën             | 14        |
| 4.2 Integratie van generatieve AI functionaliteit    | 16        |

# 1 Definitie en afbakening voorziening

## 1.1 Beschrijving van de scan- en herkenvoorziening

De nieuwe voorziening biedt centraal de volgende functionaliteiten:

- het digitaliseren van een fysiek document inclusief beeldverwerking en -verbetering en karakterherkenning,
- het ontvangen van een digitaal document ter verwerking,
- het metadateren van een document,
- het extraheren van gegevens in een document (optioneel). De geëxtraheerde gegevens worden in een separaat object (bestand) bewaard in een daartoe aangewezen opslagvoorziening. Voor de documenten aangaande DSV en DZW wordt dit gegevensobject bewaard in het DMS Primair.
- Routeren van ten eerste een document met metadata en ten tweede optioneel de geëxtraheerde gegevens naar de correcte opslagvoorziening,
- Notificeren van een bedrijfssysteem dat een nieuw document verwerkt is en opgehaald kan worden uit de opslagvoorziening.

In de visie convergeren alle scanoplossingen naar deze centrale voorziening; te beginnen met de harmonisatie van de Kofax scanstraat oplossing en de Kofax scanoplossing voor F&C. Op termijn is de verwachting dat de PGB 2.0 scanoplossing vervangen wordt door deze nieuwe voorziening.

Vanwege de uitbreiding van de functionaliteit is niet meer sprake van een scanvoorziening maar van een scan- en herkenvoorziening. Deze voorziening wordt gerealiseerd met een Intelligent Document Processing (IDP) oplossing. Dit type technologie biedt meer functionaliteit, maar voor de scan- en herkenvoorziening mag alleen bovenstaande functionaliteit in gebruik worden genomen. Naast traditionele scan- en herkentechnieken bieden IDP oplossingen steeds meer AI technologieën. Inzet van de combinatie van beide technologieën behoort tot de mogelijkheden afhankelijk van de geschiktheid voor een bepaalde taak en de daaraan gestelde kwaliteitseisen. Zie de duiding van de AI en traditionele technologieën in het hoofdstuk AI.

## 1.2 Afbakening van de voorziening

Inlezen van gestructureerde gegevens, zoals elektronische XML- of JSON-berichten, vallen buiten de reikwijdte van deze PSA, voor wat betreft de verwerking door scansoftware. Alleen semi- of ongestructureerde documenten worden ingescand, mits hier een bedrijfsmatige noodzaak voor is (lees: er is een taak vanuit een bedrijfsproces die dit vereist). Indien een document reeds volledig van metadata is voorzien en data niet geëxtraheerd hoeft te worden, wordt dit document niet verwerkt door de scan- en herkenoplossing.

## 1.3 Applicatiefunctie

De scansoftware moet de volgende functionaliteit ondersteunen.

| Nr. | Applicatiefunctie       | Opmerking                                                                                        |
|-----|-------------------------|--------------------------------------------------------------------------------------------------|
| 1   | Integratie met scanners | De softwarematige integratie met scanners (multifunctional printers (MFP's), bulkscanners, e.d.) |

|    |                                                                                                 |                                                                                                                                                                                                                                                                                                                  |
|----|-------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2  | Inlezen van 'RAW', beeld- en documentformaten                                                   | Het inlezen van authentieke 'RAW' <sup>1</sup> scanoutput, vereiste bestandsformaten voor afbeeldingen (bijvoorbeeld TIFF, PNG, JPG of BMP) en voor documenten (bijvoorbeeld PDF of Word).                                                                                                                       |
| 3  | Optische tekenherkenning                                                                        | Omzetten van afbeeldingen in bewerkbare en doorzoekbare bestanden, zoals PDF- of Word-documenten.                                                                                                                                                                                                                |
| 4  | Automatische data-extractie (scan-en-herken)                                                    | Scan-en-herken software zet papieren en digitale documenten (zoals bonnetjes en facturen) automatisch om naar gestructureerde data; waaronder de herkenning van bar- en QR-codes zoals meegegeven vanuit output management.                                                                                      |
| 5  | Beeldbewerking & verbetering                                                                    | Automatisch corrigeren van documenten zoals de afmetingen, verwijderen van schaduwen en draaien van scheve pagina's voor optimale leesbaarheid.                                                                                                                                                                  |
| 6  | Ondersteuning afhandeling uitval met handmatige taken                                           | Uitval binnen het scan-en-herkenproces wordt afgehandeld door Facilities en uitvoerende afdelingen, vanuit de scansoftware.                                                                                                                                                                                      |
| 7  | Notificeren van gekoppelde applicaties.                                                         | De scansoftware is in staat gekoppelde applicaties te notificeren (events) betreffende de nieuw opgeslagen documenten.<br><br><i>De exacte conversatiepatronen voor applicatie integratie worden door solution architectuur bepaald.</i>                                                                         |
| 8  | Aanleveren voor beschikbaarstelling van de uitvoer van het scannen aan een DMS/RMS              | Het aanleveren van documenten, metadata en optioneel eventueel hun gegevensbestanden aan een DMS (combinatie van document management systeem (DMS) en record management systeem (RMS)). Voor operationele doeleinden de verbeterde versie (PDF/A) en voor archiveringsdoeleinden het originele gescande bestand. |
| 9  | Aanleveren voor beschikbaarstelling van de uitvoer van het scannen aan administratieve systemen | Het aanleveren van documenten, metadata en optionele gegevens(bestanden) aan specifieke administratieve systemen, zoals financieel systeem voor verwerking van facturen.                                                                                                                                         |
| 10 | Aanleveren van metadata over verwerkte documenten specifiek voor archivering aan een RMS        | Het aanleveren van metadatabestanden over verwerkte documenten aan een record management systeem (RMS). Dit is geen onderdeel van de gewenste situatie maar moet in de toekomst mogelijk zijn.                                                                                                                   |

<sup>1</sup> De fysieke scanners zullen de gedigitaliseerde documenten zoveel als mogelijk in 'raw' formaat gaan aanbieden aan de IDP software, waardoor de software of medewerker alsnog aanpassingen kan doen nadat de eerste scanning van het papieren document heeft plaatsgevonden. 'raw' formaat betekent in deze context de authentieke scanoutput met geen of minimale beeldverwerking. Want als de fysieke scanner zelfstandig (via scanner-embedded functionaliteit) reeds bepaalde beeldbewerkingen foutief heeft uitgevoerd, kan een medewerker of de IDP software, deze mogelijk nooit meer herstellen.

## 2 Implementatiewijze van de voorziening

### 2.1 Dienstverleningsmodel en hosting

De scan- en herkenoplossing moet volgens het dienstverlening- en hostingmodel Software as a Service (SaaS) – Virtual Private Cloud (VPC) voorziening worden gerealiseerd. In deze paragraaf wordt nader ingegaan op de keuze en consequenties voor SaaS. De eisen hieromtrent zijn opgenomen in het Programma van Eisen (PvE).

De keuze voor SaaS is ingegeven door een tweetal factoren. Ten eerste is hier sprake van standaard software die door commerciële partijen wordt aangeboden. Daarmee is dit in lijn met de cloud richtlijnen van de SVB. Ten tweede heeft de SVB gekozen voor ontzorging zowel ten aanzien van software als ten aanzien van infrastructuur. Het laatste betekent dat de SVB wegbeweegt van on premises infrastructuur ten faveure van cloud infrastructuur.

Consequentie is dat ten aanzien van beheer, alleen het functioneel beheer door de SVB gedaan wordt en het technisch applicatiebeheer en het infrastructuurbeheer door de dienstverlener moet worden gedaan. Hiermee is de dienstverlener verantwoordelijk om zorg te dragen dat de oplossing functioneert volgens de afgesproken kwaliteitscriteria en hiertoe alle benodigde activiteiten verricht waaronder bijvoorbeeld monitoring, versie releasing en patching en incident- en probleemoplossing.

### 2.2 Lokale software

Voor de aansturing van scanner-hardware heeft een cloudgebaseerde inrichting de voorkeur. Indien lokale software aantoonbaar noodzakelijk is voor hardware-aansturing, moet dit met rationale worden onderbouwd en vooraf door SVB worden goedgekeurd. Lokale software mag geen verplichte on-premises runtime, klassiek domein, separate beheerstraat of structurele afhankelijkheid binnen het SVB-netwerk introduceren. De component moet passen binnen het SVB-beheer-, security-, identity-, monitoring- en lifecyclemodel en mag uitsluitend worden ingezet voor het doel waarvoor deze noodzakelijk is.

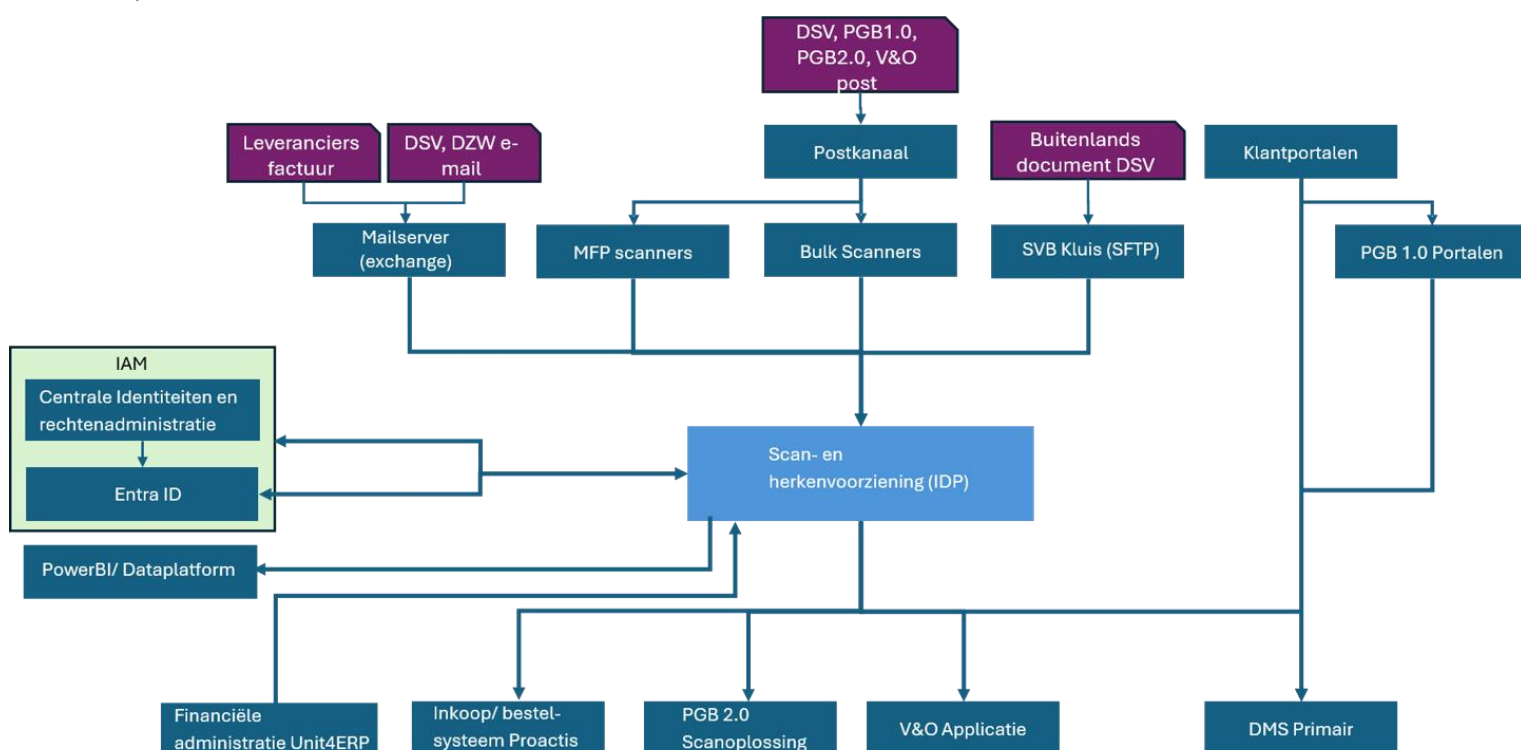
## 3 Applicatielandschap

In dit hoofdstuk wordt een toelichting gegeven op het applicatielandschap na implementatie van de gekozen oplossing inclusief koppelingen door het project en wordt een overzicht gegeven van mogelijke potentiële koppelingen in de toekomst.

De volgende tabellen tonen de verwachte toekomstige koppelingen tussen de systemen in de doelsituatie. Bij de implementatie van de koppelvlakken moet vanuit solution architectuur bekeken worden of aanvullende vereisten ingevuld moeten worden. Hierbij dient tevens de Doelarchitectuur Netwerkbeveiliging betrokken te worden.

### 3.1 Overzicht na as-is vervanging

De volgende figuur toont het applicatielandschap na de as-is vervanging van de huidige software voor de scanstraat en voor de F&C scanoplossing. Na de figuur volgt een korte beschrijving van de gekoppelde systemen.



| Nr. | Naam                  | Omschrijving                                                                                                            |
|-----|-----------------------|-------------------------------------------------------------------------------------------------------------------------|
| 1   | Mailserver (exchange) | E-mail voorziening voor verwerking van inkomende en uitgaande e-mails.                                                  |
| 2   | SVB Kluis (SFTP)      | Managed File Transfer voorziening (EFT Enterprise van Globalscape) om grote bestanden intern en extern uit te wisselen. |

|    |                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|----|--------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 3  | Klantportalen                                          | Gesloten 'mijn'-portalen voor de klanten, burgers/cliënten en werkgevers, van de SVB om activiteiten te verrichten aangaande de regelingen die SVB uitvoert uit hoofde van haar wettelijke taken, bijvoorbeeld de aanvraag voor de uitkering van een bepaalde regeling.                                                                                                                                                            |
| 4  | PGB 1.0 Portaal                                        | Medewerkerportaal voor PGB 1.0 waarlangs onder andere documentuploads in een klantportaal worden gecontroleerd.                                                                                                                                                                                                                                                                                                                    |
| 5  | DMS Primair                                            | Het Document Management Systeem (verder genoemd als DMS Primair) voor het opslaan en raadplegen van alle uitgaande als inkomende correspondentie die de SVB met haar klanten en ketenpartners gevoerd heeft voor de Dienstverlening Sociale Verzekeringen (DSV) en Dienstverlening Zorg en Welzijn (DZW) voor PGB1.0. Aan dit systeem zijn alle administraties gekoppeld van DSV en DZW die hun documenten opslaan in DMS Primair. |
| 6  | V&O applicatie                                         | Op zichzelf staande administratie voor de verwerking van gegevens van Verzetsdeelnemers en Oorlogsgetroffenen (V&O) in het kader van de uitvoering van wetten en regelingen voor deze groepen.                                                                                                                                                                                                                                     |
| 7  | PGB 2.0 scanoplossing                                  | PGB 2.0 scanstraatoplossing voor metadatering en dataextractie voor de PGB 2.0 systeeminstallatie. Op termijn kan deze specifieke oplossing mogelijk vervangen worden door de centrale scan- en herkenningsvoorziening.                                                                                                                                                                                                            |
| 8  | Inkoop- en bestelsysteem Proactis                      | Financieel systeem dat het inkoop-/bestelproces ondersteunt inclusief verwerking van facturen van leveranciers.                                                                                                                                                                                                                                                                                                                    |
| 9  | Financieel Administratiesysteem Unit4ERP               | Het systeem voor de financiële huishouding van de SVB dat bijvoorbeeld de functionaliteit omvat op het gebied van financiële administratie, debiteuren- en crediteurenbeheer en urenregistratie. De door Proactis goedgekeurde facturen worden vanuit Unit4ERP betaalbaar gesteld.                                                                                                                                                 |
| 10 | Centrale Identiteiten- en toegangsrechtenadministratie | Systeem voor het centraal beheren van eindgebruiker accounts, rollen en rechten.                                                                                                                                                                                                                                                                                                                                                   |

|    |                      |                                                                                                                                 |
|----|----------------------|---------------------------------------------------------------------------------------------------------------------------------|
| 11 | Entra ID             | Identificatie, authenticatie en autorisatie inclusief single sign on (SSO) voor cloud applicaties.                              |
| 12 | PowerBI/Dataplatform | Dashboard/rapportage tooling. Indien nodig via het Dataplatform van de SVB voor formele rapportages op gecontroleerde datasets. |

### 3.2 Koppelingen in de projectscope

De documentstromen die in de huidige situatie verwerkt worden, door de scanstraatoplossing en door de F&C scanoplossing, zijn as-is in scope van het project.

| Nr. | Interface                                                                        | Omschrijving                                                                                                                                              | Van                                                 | Naar                  | Verwachte integratie via                                                                                                      |
|-----|----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|-----------------------|-------------------------------------------------------------------------------------------------------------------------------|
| 1   | Ingescande post                                                                  | Formulieren, brieven, bijlagen en gespreksverslagen aangeboden via de kantoorbalies en de postkamers voor DSV (AxW, TSB, VV) en DZW (PGB1.0, PGB2.0, V&O) | Nieuwe bulk scanner, Multi-functional printer (MFP) | Nieuwe IDP oplossing  | Zie paragraaf 'Lokale software', output in raw format TIFF                                                                    |
| 2   | Email (DSV, DZW)                                                                 | Formulieren, brieven, bijlagen, gespreksverslagen aangeboden via email voor DSV en DZW                                                                    | Microsoft Exchange                                  | Nieuwe IDP oplossing  | Mail import connector (MS Graph API)                                                                                          |
| 3   | Facturen F&C (email)                                                             | Facturen voor F&C aangeboden via email (in PDF formaat)                                                                                                   | Microsoft Exchange                                  | Nieuwe IDP oplossing  | Mail import connector (MS Graph API)                                                                                          |
| 4   | Documenten buitenlandse kantoren (DSV)                                           | Formulieren, brieven, bijlagen (DSV; SVB Buitenlandse kantoren Rabat en Ankara)                                                                           | SVB Kluis (EFT Enterprise)                          | Nieuwe IDP oplossing  | SFTP                                                                                                                          |
| 5   | Rapportage gegevens vanuit de Nieuwe IDP oplossing                               | Levering metrics IDP services                                                                                                                             | Nieuwe IDP oplossing                                | PowerBI/ Dataplatform | Rest-API (heeft voorkeur voor Dataplatform), of PowerBI-connector of als laatste als CSV bestand via SVB Kluis (managed SFTP) |
| 6   | Gescande en ingelezen formulieren, brieven, bijlagen, gespreksverslagen voor V&O | Documenten vanuit de Nieuwe IDP oplossing met metadata en eventueel geëxtraheerde inhoudelijke data (V&O)                                                 | Nieuwe IDP oplossing                                | V&O WMS               | SFTP, PDF/A                                                                                                                   |
| 7   | Ingescande fysieke documenten voor PGB 2.0                                       | Handhaving van de huidige situatie, waarin de PGB2.0 scanstraat nog de ingescande documenten verwerkt                                                     | Nieuwe IDP oplossing                                | PGB Core (PGB 2.0)    | SFTP, TIFF formaat en/of PDF/A                                                                                                |

|    |                                          |                                                                                                                                                                |                                                               |                             |                                                                                                                                                                                                 |
|----|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------|-----------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 8  | Brondata/<br>standaardwaarden            | Leveranciersgegevens                                                                                                                                           | Unit4 ERP                                                     | Nieuwe IDP<br>oplossing     | Rest-API                                                                                                                                                                                        |
| 9  | Nieuwe factuur met<br>geëxtraheerde data | Facturen F&C (Geëxtraheerde<br>data incl. factuur ID + PDF)                                                                                                    | Nieuwe IDP<br>oplossing                                       | Proactis                    | Rest-API,<br>UBL+PDF/A                                                                                                                                                                          |
| 10 | Opslaan document                         | Opgeslagen document met<br>metadata                                                                                                                            | Nieuwe IDP<br>oplossing                                       | DMS<br>primair;<br>OpenText | Dossier service<br>(DMS Primair), TIFF<br>en PDF/A,                                                                                                                                             |
| 11 | Authenticatie en<br>autorisatie          | Verschaffen van toegang (op<br>basis van authenticatie van de<br>identiteit en<br>rechtenverlening) tot de<br>voorziening op basis van<br>Single Sign On (SSO) | Nieuwe IDP<br>oplossing                                       | Entra ID                    | Authenticatie:<br>Open ID Connect<br>(OIDC), SAML 2.0<br>Autorisatie: OAuth<br>2.0                                                                                                              |
| 12 | Centraal<br>gebruikersbeheer             | Beheer van gebruikers, rollen<br>en toegangsrechten                                                                                                            | IBIS (Centrale<br>identiteiten en<br>rechtenadminist<br>ratie | Nieuwe IDP<br>oplossing     | Via een User<br>provisioning<br>standaard. Voor<br>cloudapplicaties<br>verloopt dit via<br>Entra ID, bij<br>voorkeur via het<br>SCIM protocol.<br>Alternatieven zijn<br>SAML JIT en/of<br>OIDC. |

### 3.3 Potentiële koppelingen op termijn

Voor documentstromen waarvan de metadata incompleet is en/of dataextractie moet plaatsvinden zijn de volgende koppelingen potentiële kandidaten, Deze vallen buiten de projectscope en zullen hierom via vervolgonderzoek nut/noodzaak beoordeeld en na goedkeuring via nieuwe veranderinitiatieven gerealiseerd moeten worden (zie ook 'Afbakening van de voorziening').

Op termijn kan de scan- en herkennvoorziening daarnaast van het DMS Primair systeem de verwerking van de barcodegegevens (metadata over een document) en het notificeren van uitvoering-/bedrijfssystemen over gescande documenten overnemen.

| Nr. | Interface                                                   | Omschrijving                                                             | Van                                | Naar                    | Verwachte<br>integratie via |
|-----|-------------------------------------------------------------|--------------------------------------------------------------------------|------------------------------------|-------------------------|-----------------------------|
| 1   | Gescande documenten van<br>Zorgkantoren<br>(PGB2.0)         | Documenten Zorgkantoren<br>(PGB2.0)                                      | SVB kluis (EFT<br>Enterprise)      | Nieuwe IDP<br>oplossing | SFTP                        |
| 2   | Brondata/<br>standaardwaarden                               | Opvraagde<br>brondata/standaardwaard<br>en uit primaire systemen<br>(VV) | VV (DSV<br>uitvoeringssyst<br>eem) | Nieuwe IDP<br>oplossing | Rest-API                    |
| 3   | Notificatie van nieuw<br>document (ter<br>ondersteuning van | Trigger bericht                                                          | Nieuwe IDP<br>oplossing            | VV                      | Rest-API                    |

|    |                                               |                                                                                                 |                                          |                                             |          |
|----|-----------------------------------------------|-------------------------------------------------------------------------------------------------|------------------------------------------|---------------------------------------------|----------|
|    | conversatie-<br>patronen)                     |                                                                                                 |                                          |                                             |          |
| 4  | Barcode en gekoppelde klant en procesgegevens | Barcodes (wetscluster AA)                                                                       | AA (DSV uitvoeringssysteem)              | Nieuwe IDP oplossing                        | Rest-API |
| 5  | Brondata/standaardwaarden                     | Opvraagde brondata/standaardwaarden uit primaire systemen (TSB)                                 | TSB (Agnitio) (DSV uitvoeringssysteem)   | Nieuwe IDP oplossing                        | Rest-API |
| 6  | Notificatie van nieuw document                | Trigger bericht                                                                                 | Nieuwe IDP oplossing                     | TSB (Agnitio)                               | Rest-API |
| 7  | Brondata/standaardwaarden                     | Opvraagde referentiegegevens.                                                                   | AA-systeem                               | Nieuwe IDP oplossing                        | Rest-API |
| 8  | Notificatie van nieuw document                | Trigger bericht                                                                                 | Nieuwe IDP oplossing                     | AA-systeem                                  | Rest-API |
| 9  | Brondata/standaardwaarden                     | Opvraagde brondata/standaardwaarden uit primaire systemen (Treks)                               | Treks (PGB 1.0) (DZW uitvoeringssysteem) | Nieuwe IDP oplossing                        | Rest-API |
| 10 | Notificatie van nieuw document                | Trigger bericht                                                                                 | Nieuwe IDP oplossing                     | Treks (PGB 1.0)                             | Rest-API |
| 11 | Brondata/standaardwaarden                     | Lookup gegevens (opvraagde brondata/standaardwaarden uit primaire systemen (PGB2.0)             | Hebbes DB                                | Nieuwe IDP oplossing                        | Rest-API |
| 12 | Notificatie van nieuw document                | Trigger bericht (nieuwe inrichting die de huidige bestandsaanlevering (19) moet gaan vervangen) | Nieuwe IDP oplossing                     | PGB Core (PGB 2.0) (DZW uitvoeringssysteem) | Rest-API |
| 13 | Documenten uploads vanaf Twinternet           | Formulieren, brieven + bijlagen (DSV-twinternet; uploads)                                       | Twinternet (Klantportaal)                | Nieuwe IDP oplossing                        | Rest-API |
| 14 | Documenten uploads vanaf MijnSVB              | Formulieren, brieven + Bijlagen (DSV-MijnSVB; uploads)                                          | MijnSVB (Klantportaal)                   | Nieuwe IDP oplossing                        | Rest-API |
| 15 | Documenten uploads vanaf MijnPGB              | Formulieren, brieven + Bijlagen (DZW-MijnPGB; uploads)                                          | MijnPGB (Klantportaal)                   | Nieuwe IDP oplossing                        | Rest-API |
| 16 | Documenten uploads vanaf MijnZK               | Formulieren, brieven + Bijlagen (DZW-MijnZK; uploads)                                           | MijnZK (Klantportaal)                    | Nieuwe IDP oplossing                        | Rest-API |

|    |                                                                                     |                                                                                                                                                                                        |                             |                                              |                 |
|----|-------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|----------------------------------------------|-----------------|
| 17 | Documenten uploads vanaf MijnWMO                                                    | Formulieren, brieven + Bijlagen (DZW-MijnWMO; uploads)                                                                                                                                 | MijnWMO (Klantportaal)      | Nieuwe IDP oplossing                         | Rest-API        |
| 18 | Documenten uploads vanaf PGB Web /Portaal                                           | Formulieren, brieven + Bijlagen (DZW-PGB2.0; uploads)                                                                                                                                  | Portal API                  | Nieuwe IDP oplossing                         | Rest-API        |
| 19 | Gescande en ingelezen formulieren, brieven, bijlagen, gespreksverslagen voor PGB2.0 | Documenten vanuit de Nieuwe IDP oplossing met metadata en eventueel geëxtraheerde inhoudelijke data* (DZW-PGB2.0)                                                                      | Nieuwe IDP oplossing        | Document services (PGB 2.0)                  | Rest-API, PDF/A |
| 20 | Brondata/ standaardwaarden                                                          | Lookup gegevens (opvraagde brondata/ standaardwaarden uit primaire systemen (PGB2.0)                                                                                                   | Hebbes DB (PGB 2.0)         | Nieuwe IDP oplossing                         | Rest-API        |
| 21 | Notificatie van nieuw document                                                      | Trigger bericht met Document ID voor PGB Core en eventueel geëxtraheerde inhoudelijke data*                                                                                            | Nieuwe IDP oplossing        | PGB Core (PGB 2.0) (DZW uitvoerings systeem) | Rest-API        |
| 22 | Opvragen document                                                                   | Opgevraagd dossier en document met metadata en eventueel geëxtraheerde inhoudelijke data (DZW-PGB2.0). Bestaande koppeling die wellicht aangepast moet worden tav. geëxtraheerde data. | Document services (PGB 2.0) | PGB Core (PGB 2.0)                           | Rest-API        |
| 23 | Afwijsbrief per email vanuit IDP oplossing                                          | Afwijsbrief per email                                                                                                                                                                  | Nieuwe IDP oplossing        | Microsoft Exchange                           | SMTP            |

### Toekomstige veranderingen in DMS Primair koppelingen

Het opslaan en opvragen van een document zijn bestaande DMS Primair services. Het opslaan en opvragen van een optioneel bestand met de geëxtraheerde gegevens uit een document is voor DMS Primair een nieuw te verwerken object. Hoe dit zal plaats vinden, gecombineerd of via aparte service-aanroepen is een ontwerpkeuze voor solution architectuur van DMS Primair.

| Nr. | Interface        | Omschrijving                                                                                                  | Van                  | Naar                  | Verwachte integratie via                      |
|-----|------------------|---------------------------------------------------------------------------------------------------------------|----------------------|-----------------------|-----------------------------------------------|
| 1   | Opslaan document | Opgeslagen document met metadata in dossier, optioneel opgeslagen gegevensobject in dossier (DSV, DZW-PGB1.0) | Nieuwe IDP oplossing | DMS primair; OpenText | Dossier service (DMS Primair), TIFF en PDF/A, |

|   |                   |                                                            |                           |                                        |          |
|---|-------------------|------------------------------------------------------------|---------------------------|----------------------------------------|----------|
| 2 | Opvragen document | Opgevraagd document met metadata, optioneel gegevensobject | DMS primair; OpenText     | VV                                     | Rest-API |
| 3 | Opvragen document | Opgevraagd document met metadata, optioneel gegevensobject | DMS primair; OpenText     | MAF (AxW medewerker front-end systeem) | Rest-API |
| 4 | Opvragen document | Opgevraagd document met metadata, optioneel gegevensobject | DMS primair; OpenText     | TSB (Agnitio)                          | Rest-API |
| 5 | Opvragen document | Opgevraagd document met metadata, optioneel gegevensobject | DMS primair; OpenText     | Treks (PGB 1.0)                        | Rest-API |
| 6 | Opvragen document | Opgevraagd document met metadata, optioneel gegevensobject | PGB 2.0 Document services | PGB Core (Werkvoorraad)                | Rest-API |

## 4 AI

### 4.1 AI versus traditionele technologieën

Intelligent Document Processing (IDP) oplossingen combineren (generatieve) AI-componenten en traditionele (deterministische) technieken in één geïntegreerde pipeline. Het is juist die hybride aanpak die ze robuust en betrouwbaar maakt in een scanstraat: (Generatieve) AI zorgt voor flexibiliteit, werkt bij variatie, vereist minder configuratie en onderhoud, waarbij traditionele technieken zorgen voor controle en voorspelbaarheid met hoge snelheid en tegen lage kosten.

Deze hybride combinatie wordt ingezet tijdens het trainen/configureren van de IDP oplossing voor de verschillende documenttypen, als ook in productie wanneer de documenten worden ingelezen en verwerkt. Elke individuele IDP-leverancier maakt zijn eigen afweging in hoeverre hij daarvoor generatieve AI inzet, maar de beweging in de markt is wel duidelijk dat generatieve AI steeds meer een standaardonderdeel van moderne IDP oplossingen wordt.

Zie in onderstaande tabel AI technologieën versus de traditionele technieken voor een bepaalde scan en herken toepassing:

| Scan en herken toepassing                                                                      | (generatieve) AI technologie                                                                                                                                                                                                                                                                                                     | Traditionele techniek                                                         |
|------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| Tekstherkenning – door conversie van een niet-doorzoekbaar document (foto) naar digitale tekst | OCR op basis van Machine/Deep Learning modellen, voor gedrukte tekst<br>Handwritten Text Recognition (HTR) <sup>2</sup> op basis van deep learning modellen voor handgeschreven tekst<br>Intelligent Character Recognition (ICR), voor handgeschreven karakters                                                                  | Template matching – vaste layouts, herkenning op basis van documentmodel      |
| Documentclassificatie (herkenning van het documenttype)                                        | Machine Learning (ML) icm. Supervised learning en Natural Language Processing (NLP),<br>NLP, classificeert op basis van betekenis van woorden<br>Multimodale documentclassificatie met Deep Learning en Transformer architecturen bijvoorbeeld door de combinatie met Computer Vision voor lay-outherkenning naast tekst-analyse | Regels & beslisbomen – keywords, barcodes, documentcodes, heuristieken        |
| Informatie- en gevoeligheidsclassificatie                                                      | Named Entity Recognition (NER), een NLP technologie voor entiteitherkenning, zoals bijvoorbeeld BSN, ziekte, fraude<br>NLP en transformer modellen om tekst en context te begrijpen<br>Trainable classifiers – Machine Learning icm. Supervised learning                                                                         | Regels & beslisbomen – keywords, barcodes, documentcodes, heuristieken        |
| Metadata- en gegevensextractie (bijvoorbeeld key-value)                                        | Named Entity Recognition (NER), een NLP technologie voor                                                                                                                                                                                                                                                                         | Regex / vaste patronen – vaste tekststructuren, maskers, reguliere expressies |

<sup>2</sup> Synoniem is Handwriting Recognition (HWR)

|                                                                                                         |                                                                                                                                                                          |                                                                                      |
|---------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| extraction, semantische extractie)                                                                      | entiteitherkenning, zoals bijvoorbeeld bedrag, adres en persoon, NLP, LLM, Domain-Specific Language Model (DSLML), voor beter begrip van organisatie-specifieke context; |                                                                                      |
| Documentstructuur herkenning: layout-analyse, tabel- en sectiedetectie, formulierindeling, handtekening | Computer Vision (kennis van de documentstructuur geeft context aan gevonden gegevens)                                                                                    | Zone-based extractie – coördinaatgebaseerde extractie, vaste veldlocaties            |
| Regels voor bepalen van nauwkeurigheid en compleetheid                                                  | Confidence scoring – scoremodellen, confidence thresholds, human-in-the-loop routing                                                                                     | Validatie & business rules – domeinregels, controletabellen, plausibiliteitscontrole |

Vanuit SVB gezien hebben (generatieve) AI technologieën de voorkeur als het leervermogen van een AI model zorgt voor het sneller kunnen verwerken van een (nieuw) documenttype en/of als traditionele technieken geen oplossing bieden en/of niet adaptief zijn ten aanzien van de voorkomende variatie op een documenttype. OCR op basis van deep learning kan bv. beter overweg met een slechte scan van een document. Het statische karakter van traditionele technieken leidt ook tot een starre en minder robuuste scantechnologie. Een kleine lay-out verandering in een formulier kan bij een traditionele techniek al leiden tot foutsituatie.

Binnen SVB worden vele documentsoorten verwerkt, met een veelvoud aan verschillende context en betekenis. Dit is iets waar generatieve AI bv. in de vorm van een LLM een toegevoegde waarde kan hebben in de verwerking ervan. Echter het trainen van bv. een LLM op alle individuele documenttypen kan wel veel tijd vergen. Daarom moet altijd de afweging gemaakt worden of de inzet opweegt tegen de opbrengst.

Voor de volgende scan-en-herken toepassingen ziet SVB de toegevoegde waarde van de volgende AI technologieën:

1. Tekstherkenning: OCR, LLM
2. Documentstructuurherkenning: ML, LLM
3. Metadata en gegevensextractie: ML, LLM

Voor classificeren en categoriseren is op voorhand in het algemeen geen voorkeur uit te spreken. Bijvoorbeeld via barcodes en QR-codes en daaraan gekoppelde gegevens kunnen stukken traditioneel geclassificeerd en geïndexeerd worden. Daarom is het gewenst dat de oplossing zowel AI-gebaseerde als traditionele technologieën ondersteunt. Daarmee is er flexibiliteit om per documenttype de best passende keuze te kunnen maken in te gebruiken technologie.

Statistische AI modellen, zoals OCR, werken op basis van waarschijnlijkheid en zijn daarmee behoorlijk deterministisch. Generatieve AI werkt veel meer op een aangeleerd patroon in combinatie met overeenkomsten, en zijn daardoor veel minder voorspelbaar/nauwkeurig. Voor SVB moet de nauwkeurigheid en compleetheid hoog zijn. Generatieve AI modellen zullen daarom vaak aanvullend getraind moeten worden op de SVB documenten om de specifieke context van de organisatie beter te begrijpen. Verwachting is dat het trainen van modellen sneller gaat daar waar het varianten op documenttypen betreft. Bedrijfsmatig zullen besluiten moeten worden genomen over de confidence scoring en confidence thresholds, wat zich vertaalt naar wanneer een medewerker de gegevensverwerking moet overnemen.

## 4.2 Integratie van generatieve AI functionaliteit

Er zijn verschillende manieren hoe generatieve AI functionaliteit binnen de nieuwe scanoplossing kan worden aangeboden. Elk van deze hebben voor- en nadelen. Het is voor de SVB relevant om te weten hoe generatieve AI functionaliteit in de nieuwe scanoplossing is geïntegreerd omdat daarmee bepaald kan worden of deze kan voldoen aan het SVB beleid inzake generatieve AI, en of deze conform de generatieve AI referentiearchitectuur is.

1. **Third-party plugin;** de nieuwe scanoplossing levert generatieve AI functionaliteit via een derde partij, waarbij deze partij verantwoordelijk is voor het gebruikte AI model, de training ervan, en evt. finetuning voor specifieke scanfunctionaliteit.  
Voordelen:
  - relatief lage investering voor scansoftwareleverancier in kennis en functionaliteitsontwikkeling,
  - biedt de modernste en nieuwste versie van de grote commerciële modellen aan klanten,
 Nadelen:
  - scansoftwareleverancier heeft weinig tot geen controle over generatieve AI functionaliteit,
  - afstemmen AI model op specifieke klant vereist een eigen workspace per klant die door de AI functionaliteitsleverancier moet worden geboden aan de scansoftwareleverancier
  - weinig tot geen afdwingbare controle door klant over zaken als soevereiniteit en (gegevens)privacy
2. **Zelfgehost model;** de nieuwe scanoplossing levert generatieve AI functionaliteit via een eigen beheerd model van de scansoftwareleverancier, waarbij deze verantwoordelijk is voor het gebruikte model, de training ervan, en evt. finetuning voor specifieke scanfunctionaliteit.  
Voordelen:
  - scansoftwareleverancier heeft volledige controle over generatieve AI functionaliteit,
  - afstemmen AI model op specifieke klant kan door scansoftwareleverancier zelf worden ingericht
  - veel meer afdwingbare controle door klant over zaken als soevereiniteit, (gegevens)privacy
 Nadelen:
  - vraagt meer investering van de scansoftwareleverancier voor het leveren van generatieve AI functionaliteit
  - kan nog steeds uitdagingen geven op het gebied van soevereiniteit en (gegevens)privacy
3. **Bring-Your-Own-AI model;** de nieuwe scanoplossing biedt de mogelijkheid aan klanten om hun eigen AI modellen te gebruiken voor het leveren van generatieve AI functionaliteit. De scansoftwareleverancier biedt plug-in functionaliteit voor externe AI modellen, naast de eigen AI modellen die de scansoftware biedt.  
Voordelen:
  - maximale flexibiliteit in het gebruik van generatieve AI functionaliteit voor klanten,
  - maximale afdwingbare controle voor klanten op het gebied van soevereiniteit en (gegevens)privacy
 Nadelen:
  - scansoftwareleverancier moet extra investeringen doen om eigen AI functionaliteit te kunnen integreren in de IDP oplossing,
  - klanten zijn zelf verantwoordelijk voor implementeren van een eigen AI model, en dit te trainen en finetunen voor gebruik binnen de nieuwe scanoplossing

Vanuit de SVB gezien biedt optie 3 de meeste controle over de eigen gegevens. Optie 2 is ook mogelijk indien deze voorzien is van voldoende afdekkende contractuele afspraken inzake soevereiniteit en (gegevens)privacybescherming. Optie 1 is enkel mogelijk als de derde partij garanties kan geven over het bieden van een eigen private workspace; anders is deze optie niet mogelijk gegeven het beleid van de SVB.

