Project Start Architectuur

Voor project:

Datalaag mboRijnland

Opdrachtgever:	Alexander Dortland	Datum:	15-07-2022
Auteur(s):	Edzo Botjes & Oscar Zonneveld & Alexander Dortland & Carlo Poli	Versie:	20210122 v1.1
Documentnaam:	PSA Datalaag - mboRijnland v1.1.docx		

mborijnland

Inhoudsopgave

Backlog & Changelog

Status	Datum	Door wie uitgevoerd	Omschrijving
Done	20190926	Edzo, Alexander, Mark	Vier overkoepelende thema's 1.1.1 toegevoegd incl omschrijving
Doing		Edzo Botjes	Impact / toegevoegde waarde vier overkoepelende thema's toevoegen. Direct/indirect is toegevoegd. Uitleg is nog niet compleet.
Done	20190925	Edzo Botjes	Nav feedback vincent, mutlifactor en dataregister requirements aan beveiligings architectuur hoofdstuk toegevoegd.
Done		Edzo Botjes	Feedback alexander vd Braak verwerkt (AD Azure is nog niet 100% ingericht, keuze voor OpenID connect is nog niet definitief)
Open			Het is sterk als in het begin expliciet komt te staan dat elke applicatie binnen het domein van mboRIjnalnd alleen met de datalaag mag koppelen en niet met andere data leveranciers / afnemers.
Open	20201008	Oscar Zonneveld	Herstructurering PSA, toevoegen principes
Done	20210119	Carlo Poli	Omzetten structuur PSA voor leesbaar maken en consistentie aanbrengen in termen.
Done	20210119	Alexander Dortland	Opmerkingen IBP verwerkt. Na akkoord van de Architectuur Review Board zijn de wijzigingen geaccepteerd en de discussies verwijderd.
Done	20210122	Alexander Dortland	Opmerkingen Datateam verwerkt en laatste kleine verbeteringen
Done	20220714	Alexander Dortland	Nieuwe versie, update naar versie 1.1. De PSA is aangepast aan de hand van nieuwe inzichten die zijn opgedaan gedurende het project datalaag-fase1

1 Inleiding

In dit hoofdstuk wordt ingegaan op de doelstelling en het gebruik van de PSA. Daarnaast wordt de aanleiding van het project/PSA beschreven.

1.1. Aanleiding Project/PSA

mboRijnland is een meerjarig traject gestart om de informatievoorziening goed in control te krijgen en toekomstbestendig op te bouwen. Aanleiding hiervoor waren knelpunten in de huidige werkwijze[1] en de ambitie om te groeien naar een hoogwaardige regie-organisatie.

[1] Zie bijlage voor een overzicht van knelpunten zoals geconstateerd in het 2e kwartaal van 2019

Beschikbare documenten:

1. Basisprincipes Architectuur[1]
2. Kwaliteitsagenda mboRijnland 2019-2022[2] (online)

1.2. Doelstelling en gebruik PSA

Deze PSA beschrijft een set van kaders en richtlijnen waarbinnen dit project zal worden uitgevoerd. Deze kaders en richtlijn worden ook vertaald/geconcretiseerd voor dit project. Daarnaast zal deze PSA waar nodig nieuwe referentie architecturen, standaarden, principes en andere kaders introduceren.

1.2.1. Tijdshorizon voor principes en uitspraken in deze PSA

Deze PSA beschrijft de kaders voor het project datalaag - dat is uitgevoerd in de periode 2021-2022. Binnen het project zijn voorzieningen geïmplementeerd die voldoen aan de kaders die zijn gesteld in deze PSA. In 2022 is het project datalaag (fase1) afgerond. De datalaag is echter nog niet uitontwikkeld. In deze PSA zijn uitspraken en principes opgenomen die binnen het project nog niet aan de orde zijn gekomen. Principes die door mboRijnland worden nagestreefd, maar waarvan de implementatie vooralsnog te veel/hoge kosten met zich mee zouden brengen, zijn in deze PSA gemarkeerd met een horizon-symbool 🌈. Hiermee wordt aangegeven dat de implementatie van betreffende uitspraak of principe zal worden opgepakt wanneer er in de markt betaalbare* voorzieningen zijn, die het betreffende principe realiseren.
* Met betaalbare voorzieningen wordt bedoeld: voorzieningen die in de afweging tussen kosten en meerwaarde voor mboRijnland als doelmatig worden beoordeeld.

[1] https://mborijnland.sharepoint.com/sites/Architectuurportaal/SiteAssets/SitePages/Basisprincipes-architectuur/basisprincipes architectuur mboRijnland 1.0.pdf

[2] https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2019/06/27/kwaliteitsagendas-2019-2022-middelbaar-beroepsonderwijs-deel-2-k---n/Kwaliteitsagenda+2019-2022+MBO+Rijnland.pdf

2 Projectinformatie

In dit hoofdstuk wordt ingegaan op het doel van de datalaag en van het project (datalaag). We beschrijven de project context, de scope en de afbakening van het project, alsmede de strategische doelen van mboRijnland. Als laatste wordt de bijdrage van de datalaag aan de strategische doelen van mboRijnland beschreven.

2.1 Doel van de datalaag

Het doel van de datalaag is om ondersteuning te bieden aan het streven van mboRijnland om de informatievoorziening goed in control te krijgen en toekomstbestendig op te bouwen. De datalaag vormt de kern van de informatievoorziening middels een verzameling van digitale diensten die door mboRijnland worden gebruikt om:

- Applicaties en Digitale diensten te integreren (Enterprise Integratie[3])
- Informatie te beheren en beslissingen te ondersteunen (Beslissingsondersteuning[4])

De datalaag zorgt voor juiste samenhang tussen deze twee hoofdfuncties.

- Het doel van Systeem Integratie (SI) is om mboRijnland te helpen om regie te voeren op (informatiestromen in) haar Informatievoorziening. De samenhangende verzameling van alle applicaties en cloudservices die door mboRijnland worden gebruikt, wordt ook wel het IV-landschap genoemd (IV: Informatievoorziening). Dit IV-landschap is complex en continu in verandering door digitalisering in de maatschappij en door nieuwe wensen van de gebruikers. Ook de ambitie van mboRijnland om een knooppunt te vormen in het Lerend Regionaal Netwerk leidt tot nieuwe wensen. Veranderingen in het IV-landschap leiden momenteel te vaak tot verstoringen. Regievoering beoogt het aantal verstoringen te verminderen (ondanks de blijvende veranderingen).

- Het doel van Business Intelligence (BI) is om mboRijnland te helpen om te sturen op haar doelstellingen (zoals onder meer verwoord in de contourennota en de kwaliteitsagenda). Business Intelligence is een bepalende factor die inzicht verschaft. Dit inzicht betreft:
  - inzicht voor directeuren/teamleiders bij het nemen van strategische/tactische besluiten
  - inzicht voor informatiemanagers bij het voeren van regie over de informatievoorziening

  Het business Intelligence gedeelte van de datalaag is gerealiseerd in een component die door mboRijnland de Canonical Data Store wordt genoemd (CDS) en bestaat uit twee delen (in totaal 9 lagen).
  - Het eerste deel 'Data Integratie' (3 lagen) bestaande uit:
    -1- voorzieningen om gegevens uit verschillende bron-systemen in de bron-laag te brengen. Deze voorzieningen noemen we 'Extract Transform Load' (ETL). Een opmerking hierbij dient wel te zijn dat de technische werking van deze voorzieningen divers is, en dat hiermee de term ETL niet helemaal juist is.
    -2- de Staginglaag en de interfacelaag (die beide door de Leverancier worden onderhouden)
  - Het tweede deel zijn de mboRijnland lagen (waaronder het canonieke model CDM). Dit canonieke model bevat een door het mboRijnland ontwikkeld model met daarbij eigen gedefinieerde entiteiten. De mboRijnland lagen bestaan uit 6 lagen: Interfacelaag, Bronspecifieke laag, Integratielaag, CDM, Legolaag, Datamartlaag (presentatielaag)

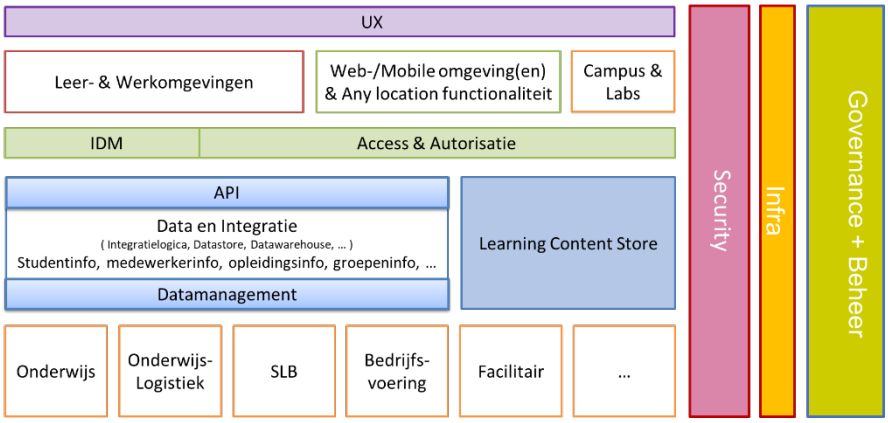
In de gewenste situatie heeft mboRijnland de informatievoorzieningsarchitectuur en de hiermee samenhangende datalaag goed onder controle, en is zij in staat om snel en adequaat in te spelen op snelle veranderingen.

[3] Door verschillende leveranciers in de markt worden de termen Enterprise Integratie en Systeem Integratie van leveranciers-specifieke definities voorzien. In dit document gebruiken we de termen als synoniemen (afgekort met SI)
[4] Beslissingsondersteuning en Business Intelligence (BI) worden in dit document als synoniemen gebruikt

Regievoering op de datalaag vormt een complexe uitdaging. De datalaag vraagt een geïntegreerd organisatorisch kader voor besturing en beheer van de informatiehuishouding. Deze PSA benoemt de uitgangspunten en implicaties hiervoor.

2.2 Project context

mboRijnland is een fusie organisatie van ID College en ROC Leiden in 2017[5].
mboRijnland telt (in januari 2021) ongeveer 18.000 studenten en 1.800 medewerkers. Van de 1.800 medewerkers zijn er ongeveer 1000 docenten en 800 medewerkers in een andere (leidinggevende of ondersteunende) rol. De studenten zijn in te delen in 2 hoofdgroepen. 13.000 studenten volgen een Beroeps Opleidende Leerweg (afgekort BOL: dit is een opleiding die grotendeels op locatie van mboRijnland plaatsvindt, met een kleiner gedeelte stage). Daarnaast volgen ongeveer 4.000 studenten een Beroeps Begeleide Leerweg (BBL-traject waarbij de student een arbeidsovereenkomst heeft met een werkgever en daarnaast vaak een dag in de week les volgt bij mboRijnland). Tenslotte heeft mboRijnland ongeveer 1.000 studenten die VAVO of contractonderwijs volgen (VAVO is voortgezet algemeen volwassenenonderwijs. Op het VAVO kunnen volwassenen een vmbo-tl-, havo- of vwo-diploma of deelcertificaten voor bepaalde schoolvakken halen. Contractonderwijs is ook gericht op volwassenen, maar niet op (middelbare) schoolvakken, maar op additionele beroepsgerichte diploma's of certificaten).



Figuur 1 – Datalaag in de domeinen van mboRijnland

De datalaag is een belangrijk onderdeel van de Architectuur van mboRijnland. Naast de datalaag kent mboRijnland een aantal andere architectuurdomeinen. Onderstaande afbeelding geeft een globaal overzicht van deze architectuurdomeinen.

Het proces- en applicatie-land-schap is reeds gerationaliseerd en geharmoniseerd na de fusie. Ten behoeve van rapportages is er op dit moment één omgeving voor beslissings-ondersteunende informatie op strategisch en tactisch niveau: Power BI. Op dit moment is de keuze gemaakt om de bronsystemen zorg te laten dragen voor de operationele rapportages, en Power BI alleen voor geaggregeerde rapportages in te zetten. Uitzondering hierop zijn operationele rapportages die buiten de eigen afdeling of team worden verspreid

Voor de toekomst heeft mboRijnland Azure als doelarchitectuur. De visie is dan ook dat de komende jaren meer applicatie van on-premise naar PaaS/IaaS gaan, en uiteindelijk als SaaS van derden worden afgenomen.

Het is de visie om de komende 5 jaar binnen mboRijnland te standaardiseren naar het gebruik van de Open Onderwijs API standaard en het gebruik van de mboRijnland API-definitie voor diensten die niet binnen de Open Onderwijs API specificatie vallen.

2.3 Scope, afbakening

2.3.1 Binnen scope Datalaag

Binnen scope van de datalaag vallen onder meer de volgende onderdelen:

[5] https://nl.wikipedia.org/wiki/MboRijnland

Het distribueren van data van bronapplicaties naar doelapplicaties

- Datacontracten (zie Bijlage C : Applicatie architectuur componenten)
- API, API-Management, en overige (aanbiedende en afnemende) koppelvlakken met de datalaag
  - Gebruik standaarden, bijv. OOAPI (Open Onderwijs API specificatie implementatie).
- Technische Architectuur

Het laden van data uit bronapplicaties in de datalaag

- Datacontracten (zie Bijlage C : Applicatie architectuur componenten)
- Technische Architectuur

Het vastleggen, uniformeren en verrijken van data

- Gegevensmodel & Datasets
- Database, Datawarehouse, Datalake
- AVG requirements in ontwerp, realisatie en implementatie
- Logging van gegevens / transacties vanuit administratieve bronsystemen en de datalaag zelf
  - Inclusief Record Management Applicatie (RMA). Noot: RMA is geen bronsysteem omdat – binnen RMA – geen nieuwe feiten worden gecreëerd.

Het aggregeren, transformeren van data tbv het leveren van informatie

- Datasets voor rapportage
- Beveiligingscomponenten (voor gegevensbeveiliging - AVG)

Ondersteuning voor de doorontwikkeling van de datalaag

- Procesbeschrijving van de Applicatieontwikkeling en het deployment proces m.b.v. Azure DevOps (CI/CD), inclusief beschrijving van de OTAP fases (pipeline stages).
- Proces beschrijving rondom versiebeheer (GitFlow / Azure DevOps)
- Procesbeschrijving inclusief sizing rondom transitie van PoC naar Productie
- Tooling om het gegevensmodel te onderhouden
- Test-infra, test-processen & Test-uitvoering (user, functional, system, integration, security/penetration).
- Anonimisering van persoonsgegevens in niet-productie omgevingen.

2.3.2 Buiten scope Datalaag

Buiten scope van de datalaag vallen onder meer:

- Bronnen:
  - Bijvoorbeeld SIS, HR, etc.
  - Bijvoorbeeld document store
  - Onderzoeksgegevens (Practoraten)
  - Autorisatiematrix
  - Referentietabellen
- Logging van operationele/infrastructurele systemen
- Identity Management (proces)
- Rapportage-tooling
- Microsoft/Office 365
- Gegevens in software van eindgebruikers die niet als bron voor een bedrijfsproces dienen, zoals bijvoorbeeld browsergeschiedenis
- Learning Content store

2.4 Strategische doelen van mboRijnland

De datalaag is onderdeel van het IV-landschap (Informatie Voorziening Landschap) van mboRijnland - en draagt bij aan het bereiken van de strategische doelen van mboRijnland, zoals verwoord in de kwaliteitsagenda[6] en Kompas.

De onderstaande tabel geeft enkele voorbeelden van deze bijdragen:

Lerend Regionaal Netwerk	De datalaag maakt het mogelijk om gegevens uit te wisselen met partners in het Lerend Regionaal Netwerk en om een integraal student-beeld (rapportage) te ontsluiten waarin gegevens vanuit verschillende interne en externe bronnen gecombineerd zijn.
Gepersonaliseerd leren	De datalaag ondersteunt bij het modulair maken van het onderwijs. De integratie draagt bij aan het gepersonaliseerd leren en maakt het mogelijk om de student van gepersonaliseerd informatie te voorzien.
Hybride leren	Via de datalaag wordt data van verschillende bronnen geïntegreerd, zodat er inzicht komt in de verschillende trajecten. Daarnaast biedt de datalaag de mogelijkheid om te integreren met applicaties of diensten van partner-bedrijven waar praktijk-onderwijs wordt gevolgd.
Professionalisering in het kader van (de bovenstaande doelstellingen uit) de kwaliteitsagenda.	De datalaag moet Lerend regionaal netwerk, gepersonaliseerd leren, hybride leren ondersteunen middels informatievoorziening. Bijvoorbeeld: KPI's voor professionaliteit van het onderwijzend personeel.
Studieloopbaanbegeleiding	Via de datalaag kan een integraal beeld van de student worden opgebouwd waarmee de kwaliteit van studieloopbaanbegeleiding wordt verbeterd.
Leven Lang Ontwikkelen	De datalaag ondersteunt uitwisseling van gegevens van/naar het persoonlijk dossier - op basis waarvan mogelijkheden/suggesties voor (Leven Lang) Persoonlijk Ontwikkelen ontsloten kunnen worden. Dit zowel voor koppelingen met eigen systemen als ook met externe systemen.
Examinering	De datalaag integreert verschillende applicaties waardoor processen rondom examinering juist en efficiënt kunnen verlopen. Tevens schept de datalaag mogelijkheden voor meer gepersonaliseerde (vormen van) examinering.
Ondersteuning door ondersteunende centra	De datalaag draagt bij aan de integratie van applicaties en informatie, waarmee efficiënte ondersteuning door ondersteunende centra wordt gefaciliteerd

Een nadere toelichting op de bijdragen van de datalaag aan de strategische doelen van mboRijnland is te vinden in 2.4 - Strategische doelen van mboRijnland.

[6] https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2019/06/27/kwaliteitsagendas-2019-2022-middelbaar-beroepsonderwijs-deel-2-k---n/Kwaliteitsagenda+2019-2022+MBO+Rijnland.pdf

2.5 Architectuur drijfveren

De relevante mboRijnland inrichtingsprincipes/uitgangspunten worden ondersteund met de onderstaande datalaag doelstellingen (Architectuur Drijfveren).

Deze zeven drijfveren zijn een samenvatting. Elke implementatie keuze rondom de datalaag zal moeten voldoen aan de toets aan deze zeven drijfveren.

1. **Complexiteit reductie**
   Als een applicatie vervangen wordt, dan hoeft er maar 1 koppeling aangepast te worden. Dit scheelt in maak- en beheer kosten. Minder complexiteit maakt hogere kwaliteit mogelijk.
2. **Stabiliteitsverbetering**
   Als een applicatie een verstoring heeft dan moet de impact op de rest van de organisatie (processen en systemen) minimaal zijn.
3. **Heldere definities & een integraal beeld**
   Gebruikers zien in hun rapportages 1 integraal beeld.
   1 integraal beeld betekent dat het beeld is opgesteld met informatie uit meerdere bedrijfsprocessen en bronsystemen. Het betekent ook dat hieruit een eenduidig en onderling samenhangend beeld wordt gevormd.
4. **Geen tegenstrijdigheden**
   Gebruikers zien in de verschillende rapportages en applicaties geen tegenstrijdige informatie.
5. **Regie ligt bij de bron**
   De bron processen en bron systemen zijn verantwoordelijk voor het opvoeren van gegevens. Hieronder valt het bewaken van de kwaliteitscriteria.
6. **Vertrouwelijkheid**
   De datalaag stelt de organisatie in staat om de herkomst en de toegang tot data te controleren. Dit komt de vertrouwelijkheid ten goede.
   - Systemen dienen daarom gebruik te maken van de integratievoorziening
7. **Personalisatie**
   De gegevens die zijn opgenomen in de datalaag (oa. integraal beeld van de student en de medewerker) stellen mboRijnland in staat om de processen in te richten rondom de persoon.

2.6 Definities

Begrip	Definitie
Bronspecifieke laag	De bronspecifieke laag is vormgegeven in de vorm van views in het schema mboR_Bron. Views die binnen de bronspecifieke laag gecreëerd worden mogen enkel gebruik maken van database objecten in de mboRijnand interface view laag. De Bron laag dient dus om alleen die leverancier database objecten te ontsluiten die voor het CDM relevant zijn. Alleen de relevante velden worden ontsloten. Binnen de bron views moeten, indien nodig maatregelen worden genomen om er voor te zorgen dat ieder record uniek geïdentificeerd kan worden dmv een sleutelveld. Dit betekent dat er soms records met dubbel voorkomende worden weggefilterd. Onderdeel van het CDS.

CDM (conceptueel)	Het Canonieke Data Model CDM is een model dat beschrijft welke gegevens worden gebruikt binnen mboRijnland en hoe deze met elkaar samenhangen. Het CDM heeft de modelvorm die in de theorie ook bekend is als Entity-Relationship-Diagram (ERD). Het CDM wordt in de Datalaag gerealiseerd door een verzameling database objecten (views en tables) die gevoed worden met data afkomstig van de verschillende informatiesystemen die binnen mboRijnland in gebruik zijn. Het CDM heeft een genormaliseerde structuur (derde normaalvorm) en is o.a. gebaseerd op entiteiten die onderkend worden in de standaard datamodellen MORA en RIO. Het CDM is qua structuur dus brononafhankelijk. Het kan organisatie breed gebruikt worden als basis voor rapportage doeleinden en systeem integratie. Noot: een volledig uitgeschreven CDM is momenteel niet beschikbaar. De toegevoegde waarde van één compleet overzicht is beperkt (dit zou eruit zien als een muur-vullend spinnenweb). Er is wel een mboRijnland informatiemodel beschikbaar (dit ziet eruit zoals het MORA informatiemodel). Tevens is er een metamodel beschikbaar (automatisch gegenereerd vanuit de CDM-laag).
CDM entiteit object	Een database object (table of view) in het CDM dat een Entiteit representeert. Een rij in een CDM entiteit object is altijd uniek te identificeren met een natuurlijke sleutel (business key).
CDM laag	Ieder CDM entiteit object zal uiteindelijk gevoed worden vanuit een enkel database object in de integratielaag De CDM entiteiten objecten bevinden zich in de CDM laag in het schema mboR_CDM. De CDM Entiteit objecten zijn normaal gesproken allemaal views omdat in de CDM laag geen datatransformaties gedaan worden. Onderdeel van het CDS.
CDS (fysiek)	Staat voor "Canonieke Data Store", is de fysieke uitwerking van het CDM. De CDS bestaat uit een SQL database en bijbehorende Azure services. Deze worden gebruikt voor informatievoorziening: de productie van informatie t.b.v. mboRijnland en externe stakeholders. Hierin zitten 9 lagen (staginglaag, interface van leverancier, interface (mboRijnland), integratie, CDM, Lego, Datamart, Referentie)
Data Integratie	Data Integratie levert functionaliteit waarin de data uit het applicatielandschap (bronsystemen) beschikbaar gemaakt wordt aan de Canonieke Data Store. Dit kan bestaan uit koppelingen met bronsystemen, maar ook gevoed worden vanuit de Systeem Integratie. Data Integratie bestaat binnen de datalaag uit 3 lagen ETL, Staginglaag en Interfacelaag (leverancier). Data Integratie wordt door mboRijnland als dienst afgenomen (geleverd door een externe leverancier).
Database object	Een table, view, stored procedure, function of index. Een database object is de fysieke realisatie van een Entiteit of Informatie-Object, in de datalaag van mboRijnland.
Datalaag	De datalaag is in principe bedoeld als enige bron van informatie t.b.v. mboRijnland die teamoverstijgend is. De datalaag bestaat uit: 1. CDS: SQL database en bijbehorende Azure services. Deze worden gebruikt voor informatievoorziening: de productie van informatie t.b.v. mboRijnland en externe stakeholders. 2. Systeemintegratie: Azure integration services. Deze worden gebruikt voor systeemintegratie: het geautomatiseerd uitwisselen van gegevens tussen verschillende informatiesystemen van mboRijnland en/of enkele systemen van externe stakeholders (systeemintegratie), voor zover deze niet via rechtstreekse point-to-point koppelingen (stekker-stopcontact) vanuit beide systemen naar elkaar toe al beschikbaar zijn.
Datamart laag	Gebaseerd op entiteiten in de CDM laag, aangevuld met herbruikbare objecten in de Lego laag kunnen diverse datamarts gebouwd worden. Een datamart kan vervolgens weer de bron vormen van een informatieproduct. Een informatieproduct kan bijvoorbeeld een rapportage zijn of een dataexport. Onderdeel van het CDS.
Entiteit	Een Entiteit is 'iets dat echt bestaat'. In de context van de datalaag is een entiteit 'een gegeven' of een informatie-object. Voorbeelden van entiteiten zijn: student, module, rooster, aanwezigheid, resultaat, etc. Een entiteit in deze context is een abstracte beschrijving van het model (CDM) dat de gegevens van mboRijnland en hun samenhang beschrijft.

ETL	ETL is een term die in de markt wordt gebruikt voor "Extract Transform Load". Deze laag levert functionaliteit waarin de data uit het applicatielandschap beschikbaar gemaakt wordt aan de datalaag. Dit kan bestaan uit koppelingen met bronsystemen, of uit componenten die database tabellen van bronsystemen kopiëren (en indien nodig transformeren naar het juiste gegevensformat).
Informatievoorziening	De informatievoorziening (van mboRijnland) bestaat uit het geheel van gegevens, processen en systemen, gericht op de productie, bewerking en beschikbaarstelling van informatie die is bedoeld om te voorzien in de informatiebehoefte van mboRijnland en/of externe stakeholders, waarbij een onderscheid moet worden gemaakt tussen: 1. De voor een MBO-instelling standaard informatiebehoefte zoals beschreven in de MBO Informatie Encyclopedie, waarbij is vastgesteld dat de daarvoor benodigde gegevens worden geregistreerd binnen mboRijnland. 2. De informatiebehoefte specifiek voor mboRijnland en/of externe stakeholders.
Integratielaag	Omdat CDM entiteit objecten soms gevoed zullen worden vanuit verschillende databronnen zullen de datasets uit deze bronnen moeten worden samengevoegd tot een geheel. Dit gebeurt in de integratielaag. Normaal gesproken worden datasets afkomstig van verschillende database objecten in de bronspecifieke laag bij elkaar gebracht binnen een integratieview die verderop in de stroom weer zal dienen als bron voor het corresponderende CDM entiteit object. Onderdeel van het CDS.
Interface laag (leverancier)	leverancier ontsluit iedere leverancier table of leverancier view via een leverancier interface view. Leverancier interface views bevinden zich in schema "Interface".
Interface laag mboR	Bovenop iedere leverancier interface view is een corresponderende mboRijnland interface view gedefinieerd in het schema mboR_Interface. Onderdeel van het CDS.
interface view	Een view die wordt gebruikt om gegevens uit een onderliggende leverancier table of leverancier view te ontsluiten. Er zijn zowel interface views die onder de verantwoordelijkheid vallen van leverancier als interface views die die vallen onder de verantwoordelijkheid van mboRijnland. De leverancier interface views onsluiten rechtstreeks een leverancier table of leverancier view. De mboRijnland interface views ontsluiten hetzelfde object indirect via de corresponderende leverancier interface view.
Lego laag	In Lego laag (schema mboR_Lego) worden herbruikbare bouwstenen gedefinieerd. Met behulp van deze bouwstenen kan vlot een nieuwe datamart in elkaar worden gezet. Voor zaken die telkens weer terugkomen bij het construeren van een datamart zoals bijvoorbeeld een tijdsdimensie of een organisatie dimensie of een bepaald afgeleid feit kan in de Legolaag een prototype worden gedefinieerd die vervolgens kan worden hergebruikt (al dan niet met wat extra filtering of aggregatie) in de diverse datamarts. Dit voorkomt dat zaken dubbel (of erger: in meerdere varianten) worden geprogrammeerd. Uniformiteit in aanpak wordt zo bevorderd. Onderdeel van het CDS.
Leverancier database object	Ieder database object dat niet in een schema startend met mboR_ staat. Deze vallen onder de verantwoordelijkheid van de leverancier. Specifieke varianten waarin in deze tekst naar verwezen wordt: leverancier table, leverancier view, etc.
mboRijnland database object	Ieder database object dat in een schema startend met mboR_ staat. Deze vallen onder de verantwoordelijkheid van mboRijnland. Specifieke varianten waarin in deze tekst naar verwezen wordt: ***mboRijnland table***, ***mboRijnland view***, etc.
MORA	MORA (Middelbaar Onderwijs Referentie Architectuur) is de referentie architectuur voor de MBO-sector die duidelijk maakt hoe een MBO-school werkt.
Object	Zie entititeit. Het begrip 'Object' wordt gebruikt als synoniem voor het begrip 'Entiteit'. Een ander synoniem met dezelfde betekenis is 'informatie-object'.
Persisteren	In plaats van dat data via een view wordt ontsloten, wordt deze in een tabel geladen met behulp van een stored procedure. Deze stored procedure wordt vervolgens op regelmatige basis aangeroepen om zo de data te verversen. Persisteren is soms noodzakelijk om performance problemen op te lossen.

Referentietabel	1. Door mboRijnland zelf onderhouden referentietabellen worden ondergebracht in het schema mboR_Referentie. 2. Door leverancier onderhouden referentietabellen zijn, net als de overige leverancier objecten te bevragen via de corresponderende Interface view (in het door leverancier onderhouden schema 'Interface').
staging laag (leverancier)	Een 1-1 kopie van de gegevens uit de bronsystemen. Vervolgens worden deze beschikbaar gesteld in de interfacelaag van de leverancier. Onder van het CDS.
Systeemintegratie	Systeemintegratie is het onderdeel van de datalaag dat zorg draagt voor de communicatie (het geautomatiseerd uitwisselen van gegevens) tussen applicaties onderling of tussen de database en de applicaties. (Applicaties kunnen zijn: verschillende informatiesystemen van mboRijnland en/of enkele systemen van externe stakeholders). Een sysnoniem voor Systeem Integratie is Enterprise Integratie: Het koppelen van twee applicaties om data en/of functionaliteit over en weer beschikbaar te maken.

2.7 Kaders en standaarden

De volgende kaders, referentie standaarden en referentie architecturen zijn in scope van de datalaag.

2.7.1 Referentie standaarden

- MORA[7] (Middelbaar Onderwijs Referentie Architectuur),
- RIO[8] (Registratie Instellingen en Opleidingen),
- Data Management DMBoK[9] – DAMA (informatie management)
- ISO25010*:2011*[10] (software quality requirements),
- AVG[11] (privacy),
- ISO27001*:2017*[12] (security),
- Open Onderwijs API[13] (koppelvlak standaard)
- OpenID Connect[14] (identificatie)
- Attribute Based Access Control[15] (ABAC)
- GIT & Gitflow[16] (versiebeheer van broncode)

2.7.2 Nog te beoordelen standaarden

- ROSA[17] (keten-referentiearchitectuur voor het hele onderwijs)
- Route21/raMBO[18] ("binnenkort beschikbaar" referentie architectuur voor het MBO)
- BPMN 2.0[19] (taal voor vastlegging bedrijfsprocessen)

---

7 https://mora.mbodigitaal.nl/index.php/Hoofdpagina

8 https://www.edustandaard.nl/standaard_afspraken/registratie-instellingen-en-opleidingen-rio/registratie-instellingen-en-opleidingen-juni-2019/

9 https://dama.org/content/body-knowledge

10 https://iso25000.com/index.php/en/iso-25000-standards/iso-25010

11 https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/algemene-informatie-avg/algemene-informatie-avg

12 https://nl.wikipedia.org/wiki/ISO/IEC_27001

13 https://openonderwijsapi.nl/en/api/

14 https://openid.net/connect/faq/

15 https://en.wikipedia.org/wiki/Attribute-based_access_control

16 https://nvie.com/posts/a-successful-git-branching-model/

17 https://www.noraonline.nl/wiki/ROSA_(Referentie_Onderwijs_Sector_Architectuur)

18 https://triplea.sambo-ict.nl/

19 https://en.wikipedia.org/wiki/Business_Process_Model_and_Notation

- Archimate (taal voor vastlegging relatie bedrijfsprocessen, informatie en informatie-systemen)
- XX als standaard voor API-ontwerp
- Architectuur principes opgesteld door Hogeschool leiden (information framework, IFW)
- Architectuur principes opgesteld door Hogeschool Arnhem en Nijmegen (HAN Data Store)
- Specifieke NORA beveiligingspatronen[20]
- Edustandaard Edukoppelingen[21] en onderdelen hiervan

.

2.8 Afbakening en relaties met andere projecten

De datalaag heeft een centrale rol in het IV-landschap van mboRijnland, doordat het de data-uitwisseling tussen applicaties faciliteert en alle data samenbrengt voor analyse en rapportage. Hierdoor heeft de datalaag een nauwe relatie met de meeste andere IV-projecten binnen mboRijnland zoals invoering AFAS HR, invoering Xedule en Osiris, Blended Education en UX.

Ook van invloed is de vervanging/vernieuwing van het Management Informatie Portaal (MIP) gedeelte.

De data definitie (entiteits-definitie) is rand-voorwaardelijk aan de bouw van de datalaag. Het betreft hier de opslag in het CDM (Canonieke Data Model) als ook de transformatie van en naar het CDM. Zie hiervoor ook de paragraaf over de informatiearchitectuur.

De data definitie (entiteits-definitie) is afhankelijk van de organisatie inrichting of kader gevend aan de organisatie inrichting (denk hierbij aan informatievastlegging in het proces).

[20] https://www.noraonline.nl/wiki/Beveiligingspatronen

[21] https://www.edustandaard.nl/standaard_werkgroepen/werkgroep-edukoppeling/

3 Principes

mboRijnland kent een aantal generieke principes voor de informatievoorziening. Deze staan weergegeven in het document Basisprincipes Architectuur.

3.1 Principes informatie architectuur

3.1.1 Gegevens hebben een eigenaar

Definitie
Van elk gegeven in de datalaag is duidelijk wie de eigenaar is.

Rationale
Door een heldere eigenaar voor elk gegeven aan te wijzen, ontstaat er geen onduidelijk waar men moet zijn indien er vragen zijn over het gegeven zelf of het gebruik ervan.

3.1.2 Gegevens hebben een BIV classificatie

Definitie
Van elk gegeven is duidelijk wat de BIV-classificatie is.

Rationale
De BIV-classificatie van gegevens maakt het mogelijk om eisen ten aanzien van beschikbaarheid, integriteit of vertrouwelijkheid goed in te vullen en daar in het gebruik rekening mee te houden.

3.1.3 Voor ieder gegeven is (precies) één bron bekend waarin het is vastgelegd

Definitie
De bron van elk gegeven is duidelijk en vastgelegd. Die bron kan over tijd veranderen.

Rationale
Door voor elk gegeven een duidelijke bron aan te wijzen ontstaat er geen onduidelijkheid omdat hetzelfde gegeven vanuit meerdere bronnen in de datalaag terecht kan komen. Ook kan de bron verantwoordelijk gemaakt worden voor de kwaliteit van de gegevens. Door de bron over tijd te kunnen laten veranderen wordt flexibiliteit in het landschap van mboRijnland gegarandeerd.

3.1.4 Gegevens hebben een definitie; deze definitie is vastgelegd in het Canonieke Data Model

Definitie
Elk gegeven is helder gedefinieerd in het Canonieke Data Model (CDM). Gegevens uit meerdere bronnen kunnen op een inzichtelijke en consistente wijze worden gecombineerd in het CDM.

Rationale
Het CDM vormt de verbindende schakel tussen gegevensobjecten in de interne structuur van (bron)systemen en bedrijfsobjecten die voor de organisatie herkenbaar zijn en geeft een heldere definitie aan de gegevens. Door gegevens uit meerdere bronnen helder en inzichtelijk te combineren kan één consistent datamodel over alle brondata ontstaan. Voorbeeld: iedereen weet wat met een student, cursist en deelnemer wordt bedoeld.

3.1.5 Van ieder (relevant) gegeven is bekend in welke processen het wordt gebruikt

Definitie
Het gebruik van gegevens in processen is helder gedefinieerd.

Rationale
Door inzichtelijk te maken in welke processen gegevens worden gebruikt, kan snel inzichtelijk gemaakt worden wat de impact is bij wijzigingen en/of problemen.

3.1.6 Wijzigingen in processen en gegevens verlopen via het change proces

Definitie
Elke wijziging op een proces en/of gegeven verloopt via het change proces.

Rationale
Door het changeproces te volgen bij wijzigingen wordt gezorgd dat de data- en procesdefinitie actueel blijven en in lijn zijn met de visie en strategie op die data- en procesdefinities.

3.1.7 Onderscheid bedrijfsobjecten (conceptueel, herkenbaar voor de organisatie) en gegevensobjecten (technische realisatie van bedrijfsobjecten)

Definitie
Maak onderscheid tussen bedrijfsobjecten en gegevensobjecten.

Rationale
Bedrijfsobjecten zijn een geabstraheerde weergave van informatie die in een bedrijf wordt gegenereerd en/of gebruikt. Gegevensobjecten vormen de wijze waarop die informatie wordt opgeslagen in de datalaag. Gegevensobjecten zijn altijd te relateren aan een bedrijfsobject.

3.1.8 Gegevens worden alleen in het aangewezen bronsysteem gewijzigd, aangemaakt en verwijderd

Definitie
Alle wijzigingen op gegevens vinden plaats in de aangewezen bronapplicatie voor dat gegeven.

Rationale
Door helder vast te leggen waar gegevens gewijzigd worden, onderstaat er geen onduidelijkheid over de laatste status van het gegeven.

3.1.9 De actualiteit van gegevens in de datalaag is bekend

Definitie
Van gegevens binnen de datalaag is bekend (wanneer dit relevant is) wanneer ze voor het laatst met het aangewezen bronsysteem zijn gesynchroniseerd.

Rationale
Op basis van de actualiteit van gegevens kunnen processen gestuurd worden die zorgen voor de vereiste tijdigheid van gegevens binnen de informatievoorziening van mboRijnland. Bijvoorbeeld status-informatie, logging of alerts voor (functioneel) beheerders.

3.1.10 Privacy-by-design

Definitie
Bij de ontwikkeling van de datalaag wordt aandacht besteed aan privacy verhogende maatregelen. Er worden zo min mogelijk persoonsgegevens verwerkt, alleen de gegevens die noodzakelijk zijn voor het doel van de verwerking. Gegevens worden niet afgeschermd voor wie het niet mag zien, maar de gegevens worden ontsloten aan wie wat mag zien. (Dus default wordt niets ontsloten.)

Rationale
mboRijnland gaat vertrouwelijk met gegevens om. Deze rationale is toegelicht in de basisprincipes.
Het idee is om al in een vroeg stadium zowel technisch als organisatorisch een zorgvuldige omgang met persoonsgegevens af te dwingen. Het houdt in dat er al bij de ontwikkeling van de datalaag aandacht moet zijn voor privacy.
Bijvoorbeeld door gegevens zo snel mogelijk te wissen/anonimiseren binnen de datalaag. Maar ook door te borgen dat bewaartermijnen van gegevens binnen de datalaag gelijk zijn aan de applicaties. We willen

bijvoorbeeld ook gegevens anonimiseren, zodat ze niet meer herleidbaar zijn naar een persoon, maar nog wel gebruikt kunnen worden in tellingen.

3.1.11 Studenten, docenten/medewerkers kunnen alle benodigde informatie vanaf één punt vinden

Definitie
Er is één plek waarvandaan informatie gevonden kan worden door stakeholders.

Rationale
Er is één informatieportaal van waaruit informatie wordt ontsloten. Informatie kan, via de datalaag, beschikbaar zijn op het informatieportaal maar kan ook worden ontsloten door vanuit het portaal door te routeren naar een bronsysteem.

3.2 Principes applicatie architectuur

3.2.1 Ontkoppeling van applicaties

Definitie
Applicaties worden maximaal ontkoppeld van elkaar zodat afhankelijkheden minimaal zijn.

Rationale
Het moet mogelijk zijn om een applicatie in het applicatielandschap te vervangen, waarbij het effect op de verbinding met de andere systemen minimaal is. Koppelingen lopen via de datalaag. De verandering in een proces of applicatie heeft minimale impact op de rest van het proces- en applicatielandschap.

3.2.2 Scheiding van databewerking en informatielevering

Definitie
Data wordt beheerd en bewerkt in daartoe aangewezen bronapplicaties die deze data delen met de datalaag. Informatie wordt geleverd aan informatiegebruikers via die datalaag.

Rationale
Het doel van de applicatie architectuur is dat de mogelijkheid van levering van gegevens door de applicaties niet beperkend is in de behoefte van gegevens levering aan de afnemer.
Voorbeeld: als een app met 18.000 gebruikers elke 10 minuten om een combinatie van roosters en docenten informatie vraagt, dan zal deze informatiebehoefte door de datalaag opgevangen worden en belast deze vraag de betrokken bronsystemen niet.
Dit betekent ook dat de beschikbaarheidseisen van de systemen lager kunnen worden ingeschaald door gebruik van de datalaag.

3.2.3 Logging van (relevante) gebeurtenissen

Definitie
Van relevante gebeurtenissen binnen de datalaag wordt een registratie bijgehouden.

Rationale
Registratie van relevante gebeurtenissen helpt bij inzicht in de (status van) verwerking binnen de datalaag. Tevens kan deze logging behulpzaam zijn bij monitoring, alerting en oplossen van incidenten (bijvoorbeeld met betrekking tot informatiebeveiliging en privacy).

3.2.4 Integrale informatievoorziening

Definitie
Alle data uit het applicatielandschap worden uniform en ondubbelzinnig met elkaar in verband gebracht zodat er een overkoepelend inzicht ontstaat over alle data en de daaruit volgende informatie.

Rationale
Het doel van de applicatie architectuur is dat de combinatie van alle gegevens vanuit alle systemen, inclusief de vertaling naar een centraal bedrijfsgegevensmodel, gaat leiden tot een integraal beeld op organisatieniveau.

3.2.5 Integrale regie op data van oorsprong tot gebruik

Definitie
Alle informatie is ondubbelzinnig terug te herleiden naar de bron van die informatie in de vorm van duidelijk gedefinieerde data in bronapplicaties.

Rationale

Het doel van de applicatie architectuur is dat er grip is op de oorsprong van informatie (lineage), en waar deze informatie naartoe gaat. Wanneer de rol van de persoon die de informatie opvraagt bekend is, en bekend is met welk doel de informatie wordt opgevraagd (processtap), dan is het mogelijk om per informatieverzoek te bepalen of informatie vertrekt mag worden.
Ditzelfde geldt voor het verwijderen van informatie. Wanneer in de bron informatie wordt verwijderd is het mogelijk om te rapporteren welke applicaties dit verwijder verzoek via de datalaag gekregen hebben.

3.3 Principes technische architectuur

3.3.1 Cloud tenzij

Definitie
De onderdelen van het applicatielandschap zijn bij voorkeur gebaseerd op Cloud-technologie, waarbinnen er geldt dat SaaS gaat boven Paas boven IaaS.

Rationale
De voorkeur gaat uit naar SaaS daarna in afnemende voorkeur PaaS, IaaS, on-premise
mboRijnland heeft gekozen voor Microsoft Azure als default cloud service provider.

3.3.2 Alles heeft een versie

Definitie
Alle componenten in de architectuur hebben een versie die opgehoogd worden bij veranderingen.

Rationale
Door rigoureus te versioneren kunnen afhankelijkheden tussen componenten duidelijker beschreven worden en kan inzichtelijk gemaakt worden welke configuraties compatibel zijn.

3.3.3 Alle bouwblokken kunnen individueel aangesproken en uitgerold kunnen worden.

Definitie
Alle componenten in de architectuur kunnen afzonderlijk gewijzigd en uitgerold worden waarbij alle interfaces gerespecteerd worden.

Rationale
Door componenten individueel uitrolbaar te maken kunnen wijzigingen sneller in gebruik genomen worden, waardoor sneller waarde geleverd kan worden.

3.4 Principes infrastructuur architectuur

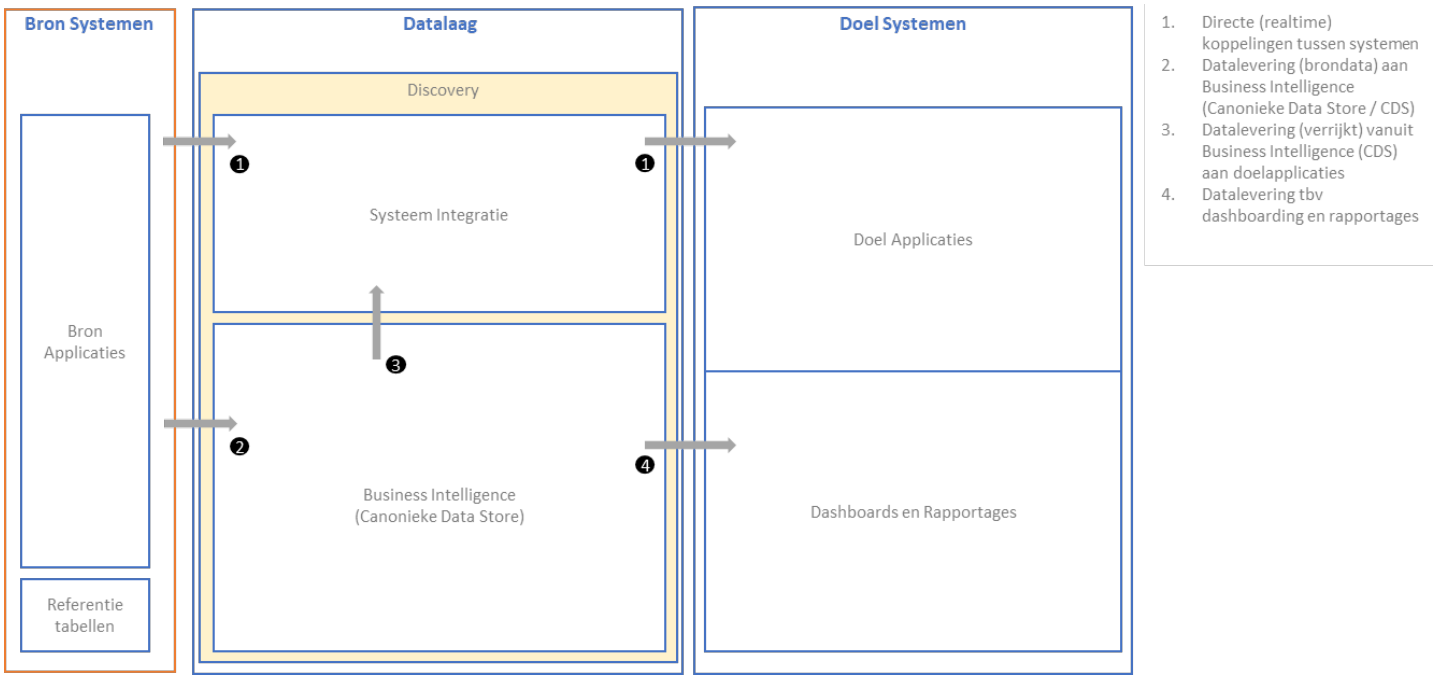
3.4.1 Naamgeving

Definitie
Alle resources worden op basis van een heldere naamgevingsconventie benoemd.

Rationale
Door heldere naamgeving wordt het makkelijker om resources te herkennen, ook als ze niet bekend zijn bij degene die ze ziet, gebruikt en/of aanpast.

4 Architectuur Datalaag

In dit hoofdstuk wordt de beoogde architectuur beschreven voor de mboRijnland datalaag. In onderstaande plaat wordt deze weergegeven.



Figuur 2 – Datalaag functionele architectuur

4.1 Functionele architectuur

Op het hoogste niveau onderscheid mboRijnland binnen de datalaag een tweetal onderdelen:

Business Intelligence
Het doel van het Business Intelligence is mboRijnland te helpen te sturen op haar doelstellingen (zoals onder meer verwoord in de contourennota en de kwaliteitsagenda).
Business Intelligence is een bepalende factor die inzicht verschaft. Dit inzicht betreft:
inzicht voor directeuren/teamleiders bij het nemen van strategische/tactische besluiten
inzicht voor informatiemanagers bij het voeren van regie over de informatievoorziening
Inzicht voor adviseurs als basis voor het geven van betekenisvolle adviezen.

Business Intelligence richt zich op het laden, uniformeren, relateren, verwerken en ontsluiten van data uit de bronsystemen van mboRijnland. Het vormt een bron voor dashboarding en rapportage en voor doelapplicaties die data nodig hebben die niet (direct) uit bronapplicaties leverbaar is, bijvoorbeeld omdat de data geaggregeerd of verrijkt is.

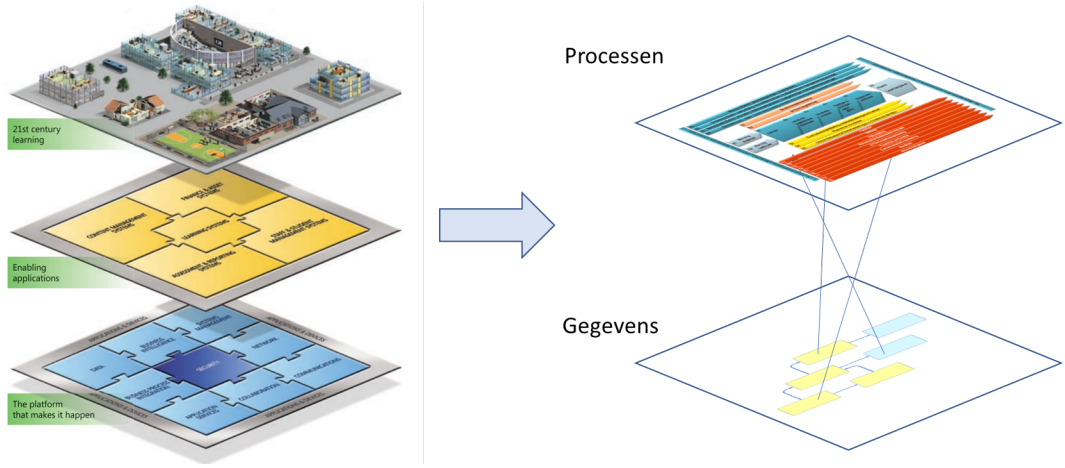
Business Intelligence maakt intern onderscheid in onderdelen die zich richten op laden van data uit bronapplicaties, onderdelen die zich richten op het koppelen en uniformeren van brondata in een canoniek model en onderdelen die zich richten op het aggregreren en verrijken van data om informatie te leveren. De functionele component waarmee de Business Intelligence wordt gerealiseerd in de datalaag wordt binnen mboRijnland Canonieke Data Store (CDS) genoemd.

Systeem Integratie
Het integratie onderdeel biedt applicaties ontkoppeling van andere applicaties en het dataplatform door middel van generieke integratiecomponenten.

Het doel van Systeem Integratie is om mboRijnland te helpen om regie te voeren op (informatiestromen in) haar Informatievoorziening. De samenhangende verzameling van alle applicaties en cloudservices die door mboRijnland worden gebruikt, wordt ook wel het IV-landschap genoemd (IV: Informatievoorziening). Dit IV-landschap is complex en continu in verandering door digitalisering in de maatschappij en door nieuwe wensen van de gebruikers. Ook de ambitie van mboRijnland om een knooppunt te vormen in het Lerend Regionaal Netwerk leidt tot nieuwe wensen. Veranderingen in het IV-landschap leiden momenteel te vaak tot verstoringen. Een goede integratielaag beoogt die verstoringen significant te verminderen.

4.2 Informatie Architectuur

Het doel van de informatie Architectuur is dat overal in mboRijnland, alle betrokkenen die informatie raadplegen of invoeren, dezelfde definitie en context gebruiken.

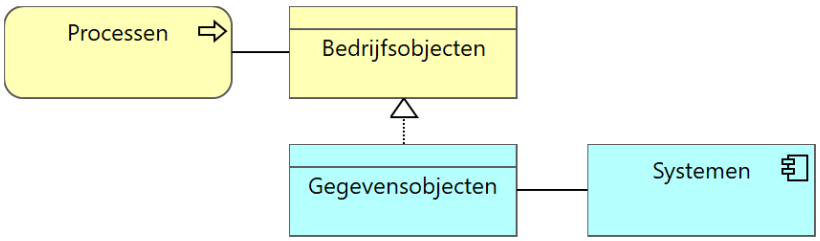


Figuur 3 – Regie informatiestromen

Binnen mboRijnland wordt het beheer voor de datalaag gevoerd binnen het "Proces- & Data- Competence Center" - een groep van mensen die kennis en vaardigheden heeft om regie te voeren over de samenhang tussen processen en gegevens. Uitgangspunt is dat goed verlopende processen bijdragen

- aan een aangename ervaring van betrokkenen bij mboRijnland (positieve user experience), en
- aan datakwaliteit (gegevens worden juist gehanteerd. Beschikbaarheid, integriteit en vertrouwelijkheid zijn voldoende geborgd).

De informatie architectuur zal worden geïmplementeerd in het zogenaamde canonieke datamodel of bedrijfsgegevensmodel, dat een realisatie is van het bedrijfsobjectenmodel. Bedrijfsobjecten representeren de concepten of objecttypen die relevant zijn voor de processen van mboRijnland. De namen en definities van de bedrijfsobjecten liggen dus dicht tegen de processen aan, en zijn dan ook herkenbaar voor 'de business'. In het bedrijfsobjectenmodel zijn ook de relaties tussen bedrijfsobjecten opgenomen. Zo vormt dit model een soort woordenboek waarin de betekenis en context van concepten (zoals Opleiding, Leertraject, Student et cetera) eenduidig is vastgelegd. Informatie over deze concepten is in systemen vastgelegd in de vorm van gegevensobjecten. Gegevensobjecten zijn dus de technische weerslag van bedrijfsobjecten. Elk systeem hanteert daarbij zijn eigen, interne, gegevensstructuur.



Figuur 4 – Bedrijfsobjecten en gegevensobjecten

Bedrijfsgegevensmodel

Het bedrijfsgegevensmodel is de gestandaardiseerde weergave van alle data binnen mboRijnland. Het bedrijfsgegevensmodel is gebaseerd op sector-, algemene- en marktstandaarden. Het bedrijfsgegevensmodel wordt praktisch weergegeven door een Canoniek Data Model (CDM). Dit CDM wordt, via de interfacelaag (loosely coupling), gevuld uit data uit de bronapplicaties. Bovenop het CDM ligt een laag die bedrijfsobjecten representeren (reflectie). Deze verzameling logische objecten zijn gericht op het ontsluiten van veel gevraagde objecten.

1Het bedrijfsgegevensmodel zal over tijd aangepast worden door nieuwe inzichten en door veranderingen in de organisatie. Denk hierbij aan het harmoniseren van processen over de onderwijsinstellingen heen, introductie van het concept van flexibel leren (dat een student bij onderwijsinstellingen waar zij niet is ingeschreven onderwijseenheden kan volgen) en andere veranderingen. Het meest duurzame is om het bedrijfsgegevensmodel een reflectie van de brondata te laten zijn, zodat een verandering van het bedrijfsgegevensmodel direct een impact op de transformaties en de reflecties heeft, zonder dat daar een volledig migratietraject voor nodig is.

Gestructureerde informatie & ongestructureerde data

Gegevens (informatie) die via de datalaag lopen hebben verschillende "doelen".

1. Er is data die gestructureerd opgeslagen en ontsloten dient te worden voor "near-real-time" provisioning naar afnemende applicaties toe.
2. Er is data die gestructureerd opgeslagen dient te worden en gebruikt wordt om historie op te bouwen. Deze informatie wordt gebruikt voor rapportages en voor het berekenen van gemiddelde andere functies over data-reeksen heen.
3. Er is data die niet-gestructureerd opgeslagen kan worden. Denk hierbij aan plaatjes, documenten (docx, pdf, pptx) en video's. Het eenduidige verwijzen naar deze informatie scheelt veel onnodige opslag en verhoogd de grip op de datakwaliteit. Denk hierbij aan het publiceren van jaarverslagen via verschillende kanalen, het gebruik van video's op verschillende touchpoints[22]. Op dit moment is het ontkoppelen van deze vorm van informatie out-of-scope.

Distributie en Rapportage

In de datastore die ligt onder het Business Intelligence component worden alle gegevens samengebracht die nodig zijn voor creëren van nieuwe data t.b.v. distributie en het creëren van informatie. Van die data is duidelijk waar die data vandaan komt en wat de status van die data is (actueel/historisch).

De data store bevat zowel ruwe gegevens (brondata) als afgeleide data (geüniformeerd in het canonieke model en aggregaties over die canonieke data).

[22] https://en.wikipedia.org/wiki/Touchpoint

Informatie architectuur principes

Code	Principe
I1	MboRijnland bedrijfsobjecten (gegevens-entiteiten) zijn vastgelegd in het bedrijfsgegevensmodel
I2	Gegevens worden alleen in het aangewezen bronsysteem gewijzigd, aangemaakt en verwijderd
I3	Studenten, docenten/medewerkers kunnen alle benodigde informatie vanaf één punt vinden

Deze principes zijn in detail uitgewerkt in Hoofdstuk 3.1 Principes Informatie Architectuur

5 Building blocks Datalaag

De datalaag bestaat uit vier onderdelen:

Canonieke Data Store
De Canonieke Data Store (CDS), is de technische of fysieke uitwerking van het Canonieke Data Model.

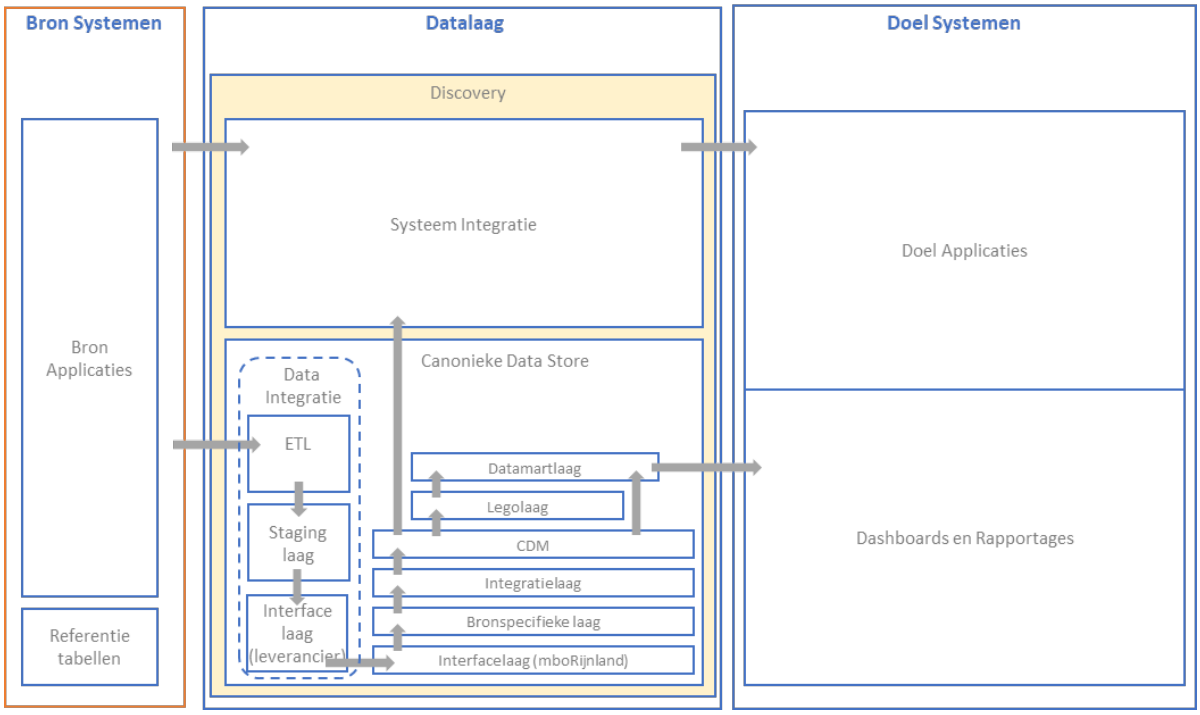
> Het Canonieke Data Model CDM is een model dat beschrijft welke gegevens worden gebruikt binnen mboRijnland en hoe deze met elkaar samenhangen. Het CDM wordt in de Datalaag gerealiseerd door een verzameling database objecten (views en tables) die gevoed worden met data afkomstig van de verschillende informatiesystemen die binnen mboRijnland in gebruik zijn. Het CDM heeft een genormaliseerde structuur (derde normaalvorm) en is o.a. gebaseerd op entiteiten die onderkend worden in de standaard datamodellen MORA en RIO. Het CDM is qua structuur dus brononafhankelijk. Het kan organisatie breed gebruikt worden als basis voor rapportage doeleinden en systeem integratie.

Het CDS bestaat uit een SQL database en bijbehorende Azure services. Deze worden gebruikt voor informatievoorziening: de productie van informatie t.b.v. mboRijnland en externe stakeholders.

Discovery
Een verzameling van componenten en tools die kan worden gebruikt om de gegevens te beheren, patronen te ontdekken, en hiervan gebruik te maken. Discovery geeft helderheid tav. definitie van metadata en technische governance. Een voorbeeld van een discovery-tool is het CDMmetamodel

Systeem Integratie
Systeem Integratie levert een middlewarelaag waarin systemen of applicaties geïntegreerd kunnen worden die, in het kader van de procesuitvoering binnen mboRijnland met elkaar integreren.

Data Integratie
Data Integratie genoemd levert functionaliteit waarin de data uit het applicatielandschap beschikbaar gemaakt wordt aan de Canonieke Data Store. Dit kan bestaan uit koppelingen met bronsystemen, maar ook gevoed worden vanuit de Systeem Integratie.



Figuur 5 – Componenten van de Datalaag

5.1 Canonieke Data Store (CDS)

De Canonieke Data Store (CDS) is een voorziening die is gericht op het optimaal aanbieden van integrale data en informatie op strategisch, tactisch en operationeel niveau. In deze voorziening wordt data over studenten, medewerkers, roosters en financiën samengevoegd beschikbaar gesteld aan doelsystemen (dashboards en power apps) aan de gebruikers. Het CDS ondersteunt de medewerkers bij: het monitoren en bijsturen op strategisch niveau, het extern verantwoorden, het nemen en toelichten van tactische besluiten en het inzicht krijgen in (operationele) processen. De combinatie van compliancy, flexibiliteit en gebruikersvriendelijkheid staan hierbij hoog in het vaandel aangezien dit een belangrijke basis vormt voor de adoptie.

Technisch gezien is het CDS een datawarehouse oplossing die is ontwikkeld is met Azure cloud technologie. De omgeving is ingericht op het snel verwerken van brondata tot mboRijnland datamodellen. Deze modellen zijn de basis voor een optimale performance van de dashboards. Het Canonieke Datamodel zorgt ervoor dat de datamodellen en de databronnen ontkoppeld zijn waardoor operationele systemen relatief eenvoudig kunnen worden vervangen.

De Canonieke Data Store kent een strikt hiërarchische gelaagdheid. Binnen elke laag wordt alleen gebruik gemaakt van gegevensobjecten uit die laag of uit de direct eronder gelegen laag. De enige uitzondering hierop wordt gevormd door referentietabellen die in alle lagen beschikbaar zijn.

Zoals aangegeven in figuur 5 “Componenten van de datalaag” bestaat de CDS uit 9 verschillende lagen:

- ETL: ETL is een term die in de markt wordt gebruikt voor “Extract Transform Load”. Deze laag levert functionaliteit waarin de data uit het applicatielandschap beschikbaar gemaakt wordt aan de datalaag. Dit kan bestaan uit koppelingen met bronsystemen, of uit componenten die database tabellen van bronsystemen kopiëren (en indien nodig transformeren naar het juiste gegevensformat).
- Staging laag: Een 1-1 kopie van de gegevens uit de bronsystemen. Vervolgens worden deze beschikbaar gesteld in de interfacelaag van de leverancier.
- Interface laag (leverancier): leverancier ontsluit iedere leverancier table of leverancier view via een leverancier interface view. Leverancier interface views bevinden zich in schema “Interface”. Ten behoeve van connectiviteit naar externe applicaties, dan wel de (volgende lagen van de) canonieke data store, is data via de Loosely Coupling laag beschikbaar.
  Noot: deze eerste 3 lagen (ETL, Staginglaag en Interfacelaag leverancier) worden gezamenlijk ook wel “Data Integratie” genoemd
- Interface laag (mboRijnland): Bovenop iedere leverancier interface view is een corresponderende mboRijnland interface view gedefinieerd in het schema mboR_Interface.
- Bronspecifieke laag: De bronspecifieke laag is vormgegeven in de vorm van views in het schema mboR_Bron. Views die binnen de bronspecifieke laag gecreëerd worden, mogen enkel gebruik maken van database objecten in de mboRijnand interface view laag. De Bronspecifieke laag dient dus om alleen die leverancier database objecten te ontsluiten die voor het CDM relevant zijn. Alleen de relevante velden worden ontsloten. Binnen de bron views moeten, indien nodig maatregelen worden genomen om ervoor te zorgen dat ieder record uniek geïdentificeerd kan worden dmv een sleutelveld. Dit betekent dat er soms records met dubbel voorkomende sleutels worden weggefilterd.
- Integratielaag: Omdat CDM entiteit objecten soms gevoed zullen worden vanuit verschillende databronnen zullen de datasets uit deze bronnen moeten worden samengevoegd tot een geheel. Dit gebeurt in de integratielaag. Normaal gesproken worden datasets afkomstig van verschillende database objecten in de bronspecifieke laag bij elkaar gebracht binnen een integratieview die verderop in de stroom weer zal dienen als bron voor het corresponderende CDM entiteit object.
- Canonieke Data Model laag: Ieder CDM entiteit object zal uiteindelijk gevoed worden vanuit een enkel database object in de integratielaag De CDM entiteiten objecten bevinden zich in de CDM laag in het schema mboR_CDM. De CDM Entiteit objecten zijn normaal gesproken allemaal views omdat in de CDM laag geen datatransformaties gedaan worden.
- Legolaag: In Lego laag (schema mboR_Lego) worden herbruikbare bouwstenen gedefinieerd. Met behulp van deze bouwstenen kan vlot een nieuwe datamart in elkaar worden gezet. Voor zaken die

telkens weer terugkomen bij het construeren van een datamart zoals bijvoorbeeld een tijdsdimensie of een organisatie dimensie of een bepaald afgeleid feit kan in de Legolaag een prototype worden gedefinieerd die vervolgens kan worden hergebruikt (al dan niet met wat extra filtering of aggregatie) in de diverse datamarts. Dit voorkomt dat zaken dubbel (of erger: in meerdere varianten) worden geprogrammeerd. Uniformiteit in aanpak wordt zo bevorderd.

- Datamart laag: Gebaseerd op entiteiten in de CDM laag, aangevuld met herbruikbare objecten in de Lego laag kunnen diverse datamarts gebouwd worden. Een datamart kan vervolgens weer de bron vormen van een informatieproduct. Een informatieproduct kan bijvoorbeeld een rapportage zijn of een dataexport.

5.1.1 Technische architectuur CDS

In onderstaand figuur zijn de Azure PaaS services gedefinieerd voor het CDS. Per PaaS component wordt het doel van dit component beschreven.

Figuur 4 - Technische architectuur

Voor deze technische architectuur geldt dat de data bronnen (in verschillende vormen, API's, Databases, Files etc.). Deze data bronnen worden middels Azure Data Factory ontsloten en de data wordt getransformeerd naar de Staging laag van de Canonieke Data Store. In de Prep en Train omgeving leven normaliter de AI modellen. Deze zijn nog niet onderkend. In de Model en Serve omgeving zijn de Analysis Modellen en Power BI rapportages en Dashboards gepositioneerd.

Data Factory v2

Azure Data Factory is een beheerde cloudservice die speciaal is ontworpen voor complexe hybride ETL- (extract-transform-load), ELT- (extract-load-transform) en gegevensintegratieprojecten.

Azure SQL Database is een volledig beheerde PaaS-data base-engine (platform as a Service) waarmee de meeste database beheer functies, zoals upgrades, patches, back-ups en bewaking, zonder tussen komst van de gebruiker worden verwerkt.

Azure Analysis Services

Power BI

Azure Analysis Services is een volledig beheerd platform als een service (PaaS) dat gegevensmodellen van ondernemingsklasse in de cloud levert.

Power BI bestaat uit een verzameling softwareservices, apps en connectors die samenwerken om uw niet-gerelateerde gegevensbronnen om te zetten in coherente, visueel aantrekkelijke en interactieve inzichten.

Keyvault Veilig sleutelbeheer is essentieel voor de bescherming van gegevens in de cloud. Gebruik Azure Key Vault om sleutels en kleine geheimen zoals wachtwoorden te versleutelen.

5.1.2 Infrastructuur CDS

Ten aanzien van de inrichting van het Azure platform voor het CDS is er een blueprint ontwikkeld waarbij alle Azure infrastructuur componenten zijn gepositioneerd en de relatie tussen de componenten is weergegeven.

Er zijn meer bronnen waarbij de infrastructurele verbinding tussen de bron en CDS anders kan zijn.

Het CDS is gebouwd op het bestaande Azure platform van mboRijnland in een eigen subscription. Dit fundament bestaat uit:

- Beheer van Azure subscriptions
- Beheer van Azure Active Directory
- Beheer van Autorisaties en toegangscontrole
- Beheer van de netwerk koppelingen, welke via een hub subscription zijn gekoppeld, waaronder de volgende diensten geleverd zijn:
    - Een netwerk peering tussen het CDS en de Hub subscription
    - Een network virtual applicatie (de centrale firewall) voor de beveiliging van het netwerk
    - Een netwerk koppeling met Osiris via een VPN (een voorbeeld databron, meerdere bronnen zijn gerealiseerd en zullen nog gerealiseerd worden)

5.2 Discovery

De Discovery-laag biedt een gecentraliseerde gegevens-beheer-functie waarmee (inzichten over) de gegevensbronnen van mboRijnland (on-premises, cloud- en SaaS-gegevens) worden bijeengebracht. Discovery geeft een overall, actueel overzicht van het gegevenslandschap van mboRijnland met geautomatiseerde gegevensdetectie, gevoelige gegevensclassificatie en een end-to-end gegevensherkomst. Discovery biedt tevens tooling voor gegevens beheer.

De discovery laag van mboRijnland zal worden opgebouwd met verschillende tools, waaronder in de toekomst mogelijk ook Azure Purview.

In de Discovery laag is zowel de metadata als technische governance ondergebracht. Voor alle bronnen wordt de metadata uitgelezen en opgeslagen in de metadata repository (CDMMetamodel) in de Discovery laag.

Deze metadata repository wordt o.a. gebruikt om de verschillende data extracties en transformaties te genereren. In de Discovery Layer wordt naast de metadata ook de technische governance opgebouwd. Dit betekent als eerste dat voor elke datatransformatie informatie wordt opgeslagen vanuit het proces. Op deze manier is het totale proces binnen de datalaag te volgen.

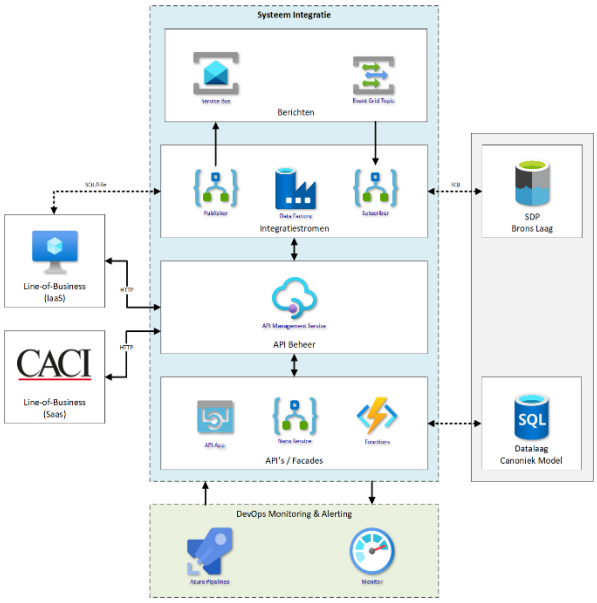
Als tweede bevat de technische governance de data repository van de velden die zijn ontsloten en de uiteindelijk beschikbaar gestelde entiteiten en attributen (middels Power BI Glossery).

De ambitie is om de dicoverylaag mboRijnland breed in te kunnen zetten, zodat zowel data integratie, het CDS als de canonieke datalaag gebruik maken van dezelfde discoverylaag. Op dit moment bevat de discoverylaag echter alleen het CDS en de aangeboden PowerBI modellen.

5.3 Systeem-integratie

Het diagram hieronder illustreert op een hoog-niveau hoe het systeem integratieplatform er uit ziet en hoe deze functioneert. Hierbij is de data van de leverancier CACI (OSIRIS) als voorbeeld genomen. De

lagen die onderdeel uitmaken van deze oplossing zijn op functioneel niveau in de volgende paragrafen beschreven. Onderstaand diagram wordt als blauwdruk gebruik voor de verschillende Systeem integratie patronen.



Figuur 5 - Systeem integratie

Berichten

De berichtenlaag in deze architectuur heeft als doel om applicaties en/of services met elkaar te integreren zonder dat deze afhankelijk van elkaar worden of dezelfde 'taal' moeten spreken. We spreken dan van een 'ontkoppelde' integratie. Dit werkt op basis van het versturen en ontvangen van berichten. Deze vorm van integratie maakt het mogelijk om meerdere afnemers van dezelfde gegevens (berichten) te hebben. Dit komt ten goede van de herbruikbaarheid van de gegevens. Ook is het mogelijk om beter te aan te sturen op de veiligheid van de gegevens omdat het inzichtelijk is welke applicaties en/of services de betreffende gegevens ontvangen.

In onze architectuur wordt deze rol vervuld door Azure Service Bus en Azure Event Grid. Beide diensten hebben hun eigen karakteristieken en use cases.

Integratiestromen

De integratielaag heeft als doel om de processen en workflows die ten grondslag liggen aan de integraties tussen de applicaties/services aan te sturen. We spreken hier over het 'orchestreren' van de integratiestromen. Daarmee wordt onder andere het plannen, starten en toezien op de integratiestroom bedoeld. Om invulling te geven aan deze behoefte worden Azure Logic Apps gebruikt. Dit is een cloudservice die het mogelijk maakt om de benodigde integratiestromen te automatiseren en daarop toe te zien.

Om deze integratiestromen te beheren en monitoren wordt de dienst Azure Monitor gebruikt. Azure Monitor is de service die Azure-breed voorziet in het verzamelen, analyseren en acteren op logging- en telemetriedata. Boven op Azure Monitor is de zogenaamde Logic Apps Management Solution beschikbaar. Deze turn-key solution kan het beste omschreven worden als een kant-en-klare weergave op geaggregeerde logging data afkomstig uit de integratiestromen. Met behulp van deze solution is het mogelijk om vanuit de Azure Portal overzicht te houden en inzichten te verkrijgen in de ontwikkelde integratiestromen.

API beheer

De API-beheerlaag heeft als doel om gegevens en functionaliteit vanuit applicaties op een gestandaardiseerde en veilige manier, vanuit een centrale plek beschikbaar te stellen. Deze werkwijze accelereert de ambitie om toe te werken naar een API ecosysteem waarbij API's op eenvoudige wijze gevonden en (her) gebruikt kunnen worden. Ook biedt het gebruik van een API Management systeem een ontkoppeling tussen de gegevens (API) en de consumerende applicaties en/of services. Door deze ontkoppeling is het bijvoorbeeld veel eenvoudiger om nieuwe versies van API's in gebruik te nemen. Hiermee neemt de wendbaarheid rondom het beschikbaar stellen van data enorm toe.
Om invulling te geven aan deze behoefte wordt Azure API Management (APIM) gebruikt. Azure API Management is het platform binnen Azure waarmee organisaties hun API's op een veilige en gecontroleerde manier API's kunnen publiceren voor zowel interne als externe afnemers.

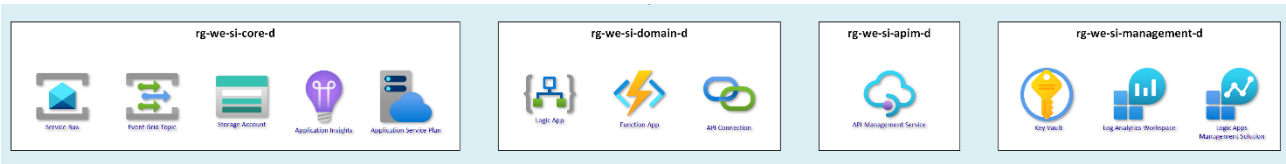
API's & façades

De API laag in de solution architectuur van het integratieplatform is ingetekend om ruimte te bieden aan API's die ontwikkeld zijn voor een specifiek doel. Dit zijn veelal façades die gebouwd worden om de complexiteit of prestatieproblemen van het onderliggende systeem te ondervangen. Ook is het mogelijk om de consumptie van data te vereenvoudigen of verbeteren, bijvoorbeeld in het geval van een legacy systeem.

Om een concrete invulling te geven aan deze behoefte biedt Azure vele mogelijkheden. Zo is het mogelijk om API's onder te brengen in onder andere Azure API Apps, Azure Functions of zelfs Azure Logic Apps. Iedere dienst heeft hierin zijn eigen voor- en nadelen. De keuze zal dan ook per geval gemaakt moeten worden. Door vooraf te anticiperen op de mogelijke noodzaak van een façade zijn we in staat om de aanpak zoveel mogelijk te standaardiseren en dPe ontwikkeling hiervan op te versnellen.

5.3.1 Infrastructuur

Applicatie-integratie is gebaseerd op Azure diensten. Hieronder staan de belangrijkste onderdelen weergegeven (de gebruikte naamgeving is voor de development omgeving):



Figuur 6 – Infrastructuur t.b.v. applicatie-integratie

5.4 Data-integratie

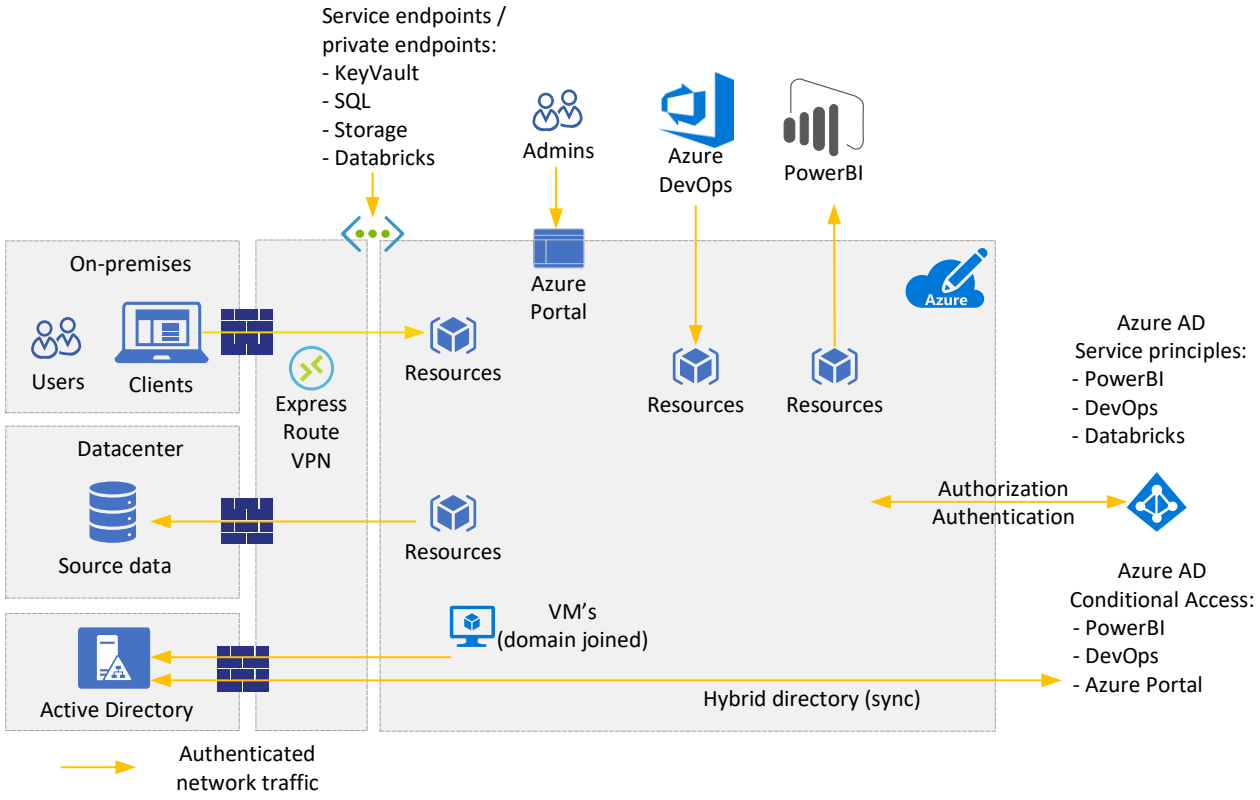
Data Integratie levert functionaliteit waarin de data uit het applicatielandschap (bronsystemen) beschikbaar gemaakt wordt aan de Canonieke Data Store. Data Integratie bestaat binnen de datalaag uit 3 lagen ETL, Staginglaag en Interfacelaag (leverancier). Data Integratie wordt door mboRijnland als dienst afgenomen (geleverd door een externe leverancier). Data-integratie is gebaseerd op Azure Data Factory. Azure Data Factory is de ETL service van Azure. ETL staat voor Extract, Transform en Load, en draait om het verenigen van grote hoeveelheden gegevens uit diverse bronnen. Voor data integratie wordt gebruik gemaakt van de diensten van een externe leverancier.

Net als Azure Logic Apps, is Azure Data Factory een orchestrator. In vergelijking met Logic Apps is Azure Data Factory daarmee in veel gevallen een geschikter middel om grote hoeveelheden data te verwerken. Een belangrijke kracht van Azure Integration Services is dat Logic Apps en Data Factory gecombineerd kunnen worden om de krachten van beide services te benutten. Een Logic App kan bijvoorbeeld orchestreren dat een dataset wordt klaargezet op een bepaalde locatie, waarna Data Factory deze op kan pakken om verder te verwerken.

6 Generieke building blocks

6.1 Security architectuur

Een belangrijk deel van de infrastructuur architectuur is de security infrastructuur. In onderstaand figuur wordt deze duidelijk gemaakt.



Figuur 7 - Security infrastructuur

De Azure subscription met de CDS-lagen is op meerdere niveaus en manieren beveiligd:

- Netwerk beveiliging
  - Publieke IP-adressen zijn alleen toegestaan wanneer deze noodzakelijk en secure zijn. Bij default zijn deze niet toegevoegd, geconfigureerd of toegestaan en wordt publiek in- en uitgaand verkeer via de firewall gerouteerd.
  - Publieke Endpoints zijn waar mogelijk beveiligd door het gekoppelde VNET en de netwerkrouteringsregels hierop (NSG) dan wel zijn endpoints beveiligd met een Azure AD Conditional Access policy
  - Alle mogelijke Azure services zijn waar mogelijk geïntegreerd met het VNET en voorzien van een privaat IPadres
  - VNET's zijn voorzien van een specifieke netwerkregels voor netwerk communicatie van/naar de CDS lagen.
  - Het gebruik en de toegang tot de Azure PaaS resources is beveiligd met Service endpoints of Private Endpoints.
- Toegangsbeveiliging
  - Resource groupen zijn voorzien van specifiek RBAC-permissies via groepen met rollen
  - Toegang is op basis van groepslidmaatschap, niet op individuele basis
  - Conditional Access policies, waaronder toepassing van MFA, reguleren externe toegang tot de volgende Azure beheer portals / endpoints :
    - Azure Portal
    - Azure Resource Manager provider

        - Classic Service Management APIs
        - Azure PowerShell
        - Visual Studio subscriptions administrator portal
        - Azure DevOps
        - Azure Data Factory portal
    - Conditional Access policies reguleren externe toegang tot de volgende SaaS portals:
        - Azure Databricks
        - Power BI portal
    - Alle Azure resources hebben onderlinge toegang op basis van expliciete permissies via de Azure Active Directory services zoals Service Principles, Managed Identity of een AAD account.

Toegang is ook altijd op naam en nooit via een gedeeld of functioneel gebruikersaccount toegestaan. Toegang is alleen mogelijk indien medewerker lid is van de juiste security groep. Meervoudige authenticatie (MFA) is ook afgedwongen via Azure AD en Conditional Access Policies.

Virtuele machines (VM) maken deel uit van het (on-premises) Active Directory domein zodat deze minimaal dezelfde beveiligingsmaatregelen hebben als alle andere servers in het netwerk. Ook kan er op de VM's alleen ingelogd worden met een domein account met de juiste permissies. VM's die worden gebruikt voor de ontwikkeling vallen hier niet onder (zie 6.3 - Ontwikkel proces en OTAP).

De volgende Azure bouwblokken zijn dusdanig geconfigureerd dat netwerktoegang alleen mogelijk is via het specifieke CDS netwerk (het VNET):

- Blueprint CDS voor mboRijnland *
- Key Vault
- Azure Storage Account
- Virtuele Machines
- Azure SQL (ingebouwde firewall)

*De blueprint voor het CDS beschrijft zowel de infrastructuur architectuur als de security architectuur en de benodigde Azure Services voor het CDS.

6.1.1.1 Beleidslijnen, richtlijnen en standaarden

Zie de centrale lijst met referentie architecturen en standaarden in 2.7 - Kaders en standaarden.

6.1.1.2 Globale Beveiliging Architectuur

Voor de beveiligingsarchitectuur gelden de eerder gemelde architectuur uitgangspunten.

1. ISO27001 & ISO27002, dit betekent onder andere
    a. Onderkennen van de scheiding tussen Identificatie, Authenticatie en Autorisatie
    b. Onderkennen dat productiedata alleen in een productie omgeving toegankelijk mag zijn (zie ook BIV en AVG)
    c. Onderkennen dat maatregelen het resultaat zijn van een risicoanalyse.
    d. Er wordt vanuit rollen geredeneerd die dynamisch toegewezen worden op basis van attributen (Attribute Based Access Control).
2. Er wordt gewerkt met BIV-classificering.
3. De AVG is van kracht die onder andere stelt:
    a. Privacy-by-design,
    b. Security-by-design,
    c. Opslag van data is gerelateerd aan het doel, wanneer het doel behaald is vervalt de doelbinding en is opslag niet meer legitiem. Deze data dient verwijderd of geanonimiseerd te worden.
    d. Toegang tot data is gerelateerd aan het doel in combinatie van de rol van de persoon in de aanvraag. Als de rol geen toegang tot deze gegevens vereist in de specifieke situatie (proces stap), dan zal er geen toegang gegeven worden.

6.1.1.3 Security Requirements

Autorisatie / Authenticatie / Identificatie

1. Authenticatie is geschikt voor federatieve authenticatie.
2. Autorisatie is op basis van de Active Directory & Azure AD van mboRijnland.
3. Autorisatie is gebaseerd ABAC toewijzing (incl. RBAC)
4. Er wordt gebruik gemaakt van OpenID connect (SURFconext)
5. Voor de toegang tot persoonsgegevens met 'vertrouwelijkheid hoog' is sterke[23] authenticatie voorgeschreven.

BIV

6. Alle persoonsgegevens hebben een BIV-classificatie.
7. De BIV classificatie van gegevens vindt plaats binnen de dataregisters van mboRijnland.

Privacy

8. Alle (digitale) communicatie is geëncrypt.
   NB: ook tussen de componenten van de oplossing.
9. Voor developers die gebruik maken van de datalaag is via de productie API's alleen geanonimiseerde data beschikbaar.

De geanonimiseerde data levert een onderlinge consistentie van gegevens voor functionele testen met data vanuit mboRijnland.

6.2 Data classificatie

Ten behoeve van informatie beveiliging binnen mboRijnland, wordt data geclassificeerd. Voor de classificatie wordt BIV classificatie gehanteerd.

> *Een BIV-classificatie of BIV-indeling is een indeling waarbij beschikbaarheid, (continuïteit), integriteit (betrouwbaarheid) en vertrouwelijkheid (exclusiviteit) van informatie en systemen wordt aangegeven.*

De classificaties worden ingedeeld in drie categorieën. Te weten hoog, midden en laag.

Vanuit het CDS wordt via Metadata & Technische governance per attribuut vastgelegd wat het bijbehorende informatie beveiligingsniveau is (BIV). Het beheer van de classificaties en de mapping aan de attributen is belegd bij de afdeling informatiebeveiliging**.**

De classificaties worden zoals aangegeven vastgelegd in het CDMMetadatamodel (discovery). Wanneer er in het canonieke model nieuwe attributen worden gecreëerd, wordt de classificatie voor deze attributen opgeslagen in het metadata-model van het canonieke model.

Indien data niet volgens BIV geclassificeerd is, dan gelden de hoogste waarden tav beschikbaarheid, integriteit en vertrouwelijkheid.

6.3 Ontwikkel proces en OTAP

6.3.1 OTAP-strategie voor de datalaag

De hele datalaag is beschikbaar in een OTAP-straat, waarbij conform goede ontwikkelstandaarden in de ontwikkelomgeving oplossingen worden gerealiseerd bestaande uit componenten die elk al op hun

[23] Sterke authenticatie is om LOA te garanderen uit naam van mboRijnland. Multi factor is self assigned. Dit is niet gekoppeld aan een betrouwbaarheidsniveau namens de instelling. Zie: Surf SecureID: https://www.surf.nl/surfsecureid-beveilig-je-diensten-extra-met-tweefactorauthenticatie

werking getest zijn. Vervolgens worden die oplossingen op (🌈Test en) Acceptatie getest om te valideren dat ze succesvol in productie genomen kunnen worden. De scenario's die hierbij gehanteerd worden, staan in meer detail beschreven in Bijlage F : OTAP Scenario's.

Noot: binnen mboRijnland is van de OTAP voor de datalaag momenteel alleen OAP gerealiseerd. Een aparte T omgeving wordt momenteel niet noodzakelijk geacht, en is daarom vanuit kostenbewustzijn achterwege gelaten. Indien dit in een later stadium wenselijk wordt geacht, kan een T-omgeving alsnog worden toegevoegd.

Bijlage A : Algemene architectuur principes

De architectuur basisprincipes en een toelichting op de implicaties zijn te vinden in het document "basisprincipes architectuur"[24]

Code	Principe
A1	mboRijnland stelt de student, zijn leervraag, ontwikkeling en resultaat centraal
A2	mboRijnland biedt gepersonaliseerd onderwijs dat aansluit op de behoeften van de arbeidsmarkt.
A3	mboRijnland faciliteert samenwerken en doorstromen in het Lerend Regionaal Netwerk.
A4	De ICT-voorzieningen vergemakkelijken contact en informatievoorziening met en naar studenten, medewerkers en derden
A4	Informatie en services zijn makkelijk te vinden en intuïtief te gebruiken
A6	mboRijnland stimuleert het gebruik van nieuwe digitale ontwikkelingen voor innovatie in het onderwijs.
A7	mboRijnland gaat vertrouwelijk om met gegevens en heeft verantwoordelijkheden expliciet belegd

-

[24] https://mborijnland.sharepoint.com/sites/Architectuurportaal/SitePages/Basisprincipes-architectuur.aspx

Bijlage B : Applicatie architectuur principes

Informatiesystemen zijn 24/7 beschikbaar
Informatiesystemen zijn 24/7 beschikbaar, makkelijk vindbaar en actueel voor een ieder die er bij moet kunnen
Implicaties
• Plaats en tijd onafhankelijk kunnen werken en leren wordt steeds belangrijker. Dit vraagt dat systemen continu beschikbaar zijn. Dit vraagt om een datalaag die ook 24/7 beschikbaar is.

Aansluiting op MORA
De ICT architectuur sluit aan bij de proces kaders van MORA en FlexID25
Implicaties
• De architectuur van het primair proces is in beginsel leidend omdat deze invulling geeft aan de wijze van uitvoering. De ICT architectuur is hier ondersteunend aan.

Uitwisselen van informatie
Het is mogelijk om eenvoudig informatie (uit systemen) uit te wisselen met belanghebbenden. Het merendeel van het werk binnen mboRijnland bevat informatie uitwisseling. De uitwisseling wordt effectiever en efficiënter wanneer het delen van informatie zo eenvoudig mogelijk is.
Implicaties
• mboRijnland ontwikkelt en onderhoud systemen waaruit informatie geëxporteerd kan worden. Er is een autorisatie matrix waarin men op basis van rollen toegang krijgt tot relevante data.

Wet- en regelgeving
Systeem inrichtingen borgen, waar mogelijk, wet- en regelgeving vereisten. Het inbedden van wet- en regelgeving binnen systemen borgt dat (waar mogelijk) de naleving van regelgeving automatisch wordt afgedwongen.
Implicaties
• Proceseigenaren dienen op de hoogte te zijn van relevante wet en regelgeving en deze waar mogelijk binnen de functionele inrichting van de relevante systemen te borgen.

Effectieve ICT
Nieuwe Informatiesystemen (en grote wijzigingen) dienen een positieve bedrijfseconomisch effect te hebben (financieel en/of kwalitatief). ICT systemen dienen een positieve bijdrage te leveren aan de bedrijfsvoering en te passen binnen de financiële kaders van mboRijnland.
Implicaties
• Conform het projectproces dient er voor grote ICT wensen een Business Case opgesteld te worden.

Minimale beheerlast
De beheerlast van systemen is vooraf inzichtelijk en zo beperkt als mogelijk.

[25] https://mborijnland.nl/l/library/download/urn:uuid:48d2c942-e8ef-438b-a091-66cb89cbdcea/onderwijskompas.pdf?format=save_to_disk&ext=.pdf

	Vaak wordt niet stilgestaan bij de aanschaf van een nieuw systeem de beheerlast. De beheerlast is daardoor niet inzichtelijk of wordt vaak onderschat.
Implicaties	
• Bij het aanschaffen van een nieuw systeem dient de beheerlast mee genomen te worden bij de overwegingen die leiden tot een aanschaf.	

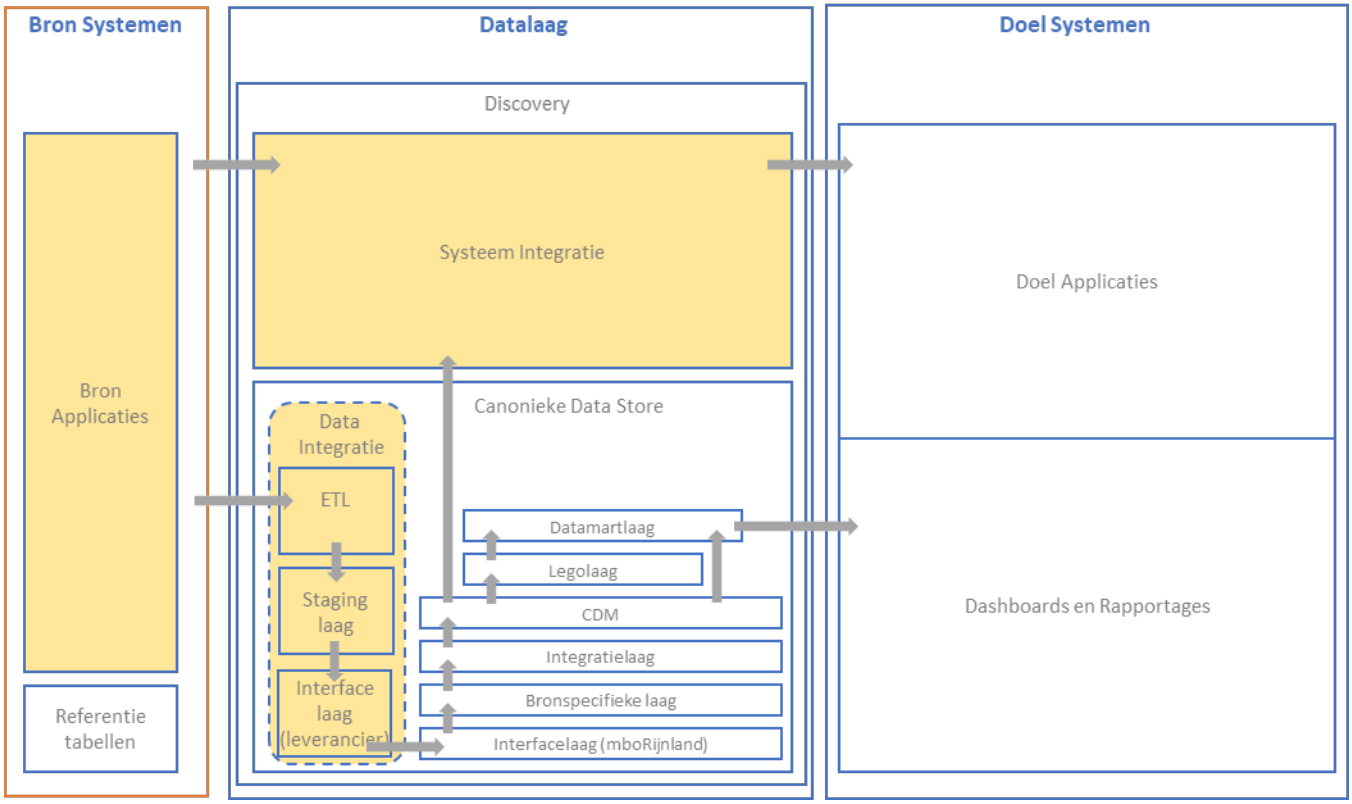
Informatie beveiliging	
	Systemen voldoen aan algemeen geldende informatiebeveiligingsnormen. De data gegevens van mboRijnland dienen beschermd te worden tegen oneigenlijk gebruik. Ontvreemding van data kan tot imagoschade leiden.
Implicaties	
• Bij het opstellen van de functionele eisen dient centrum voor Informatie en Innovatie betrokken te worden om naar de beveiligingseisen te kijken.	

Bijlage C : Applicatie architectuur componenten

In deze bijlage wordt per architectuurcomponent het component beschreven en de daarbij geldende principes.

Data bronnen en integratie

De eerste componenten die worden beschreven zijn de bronnen en de integratiecomponenten (systeem integratie en data integratie).



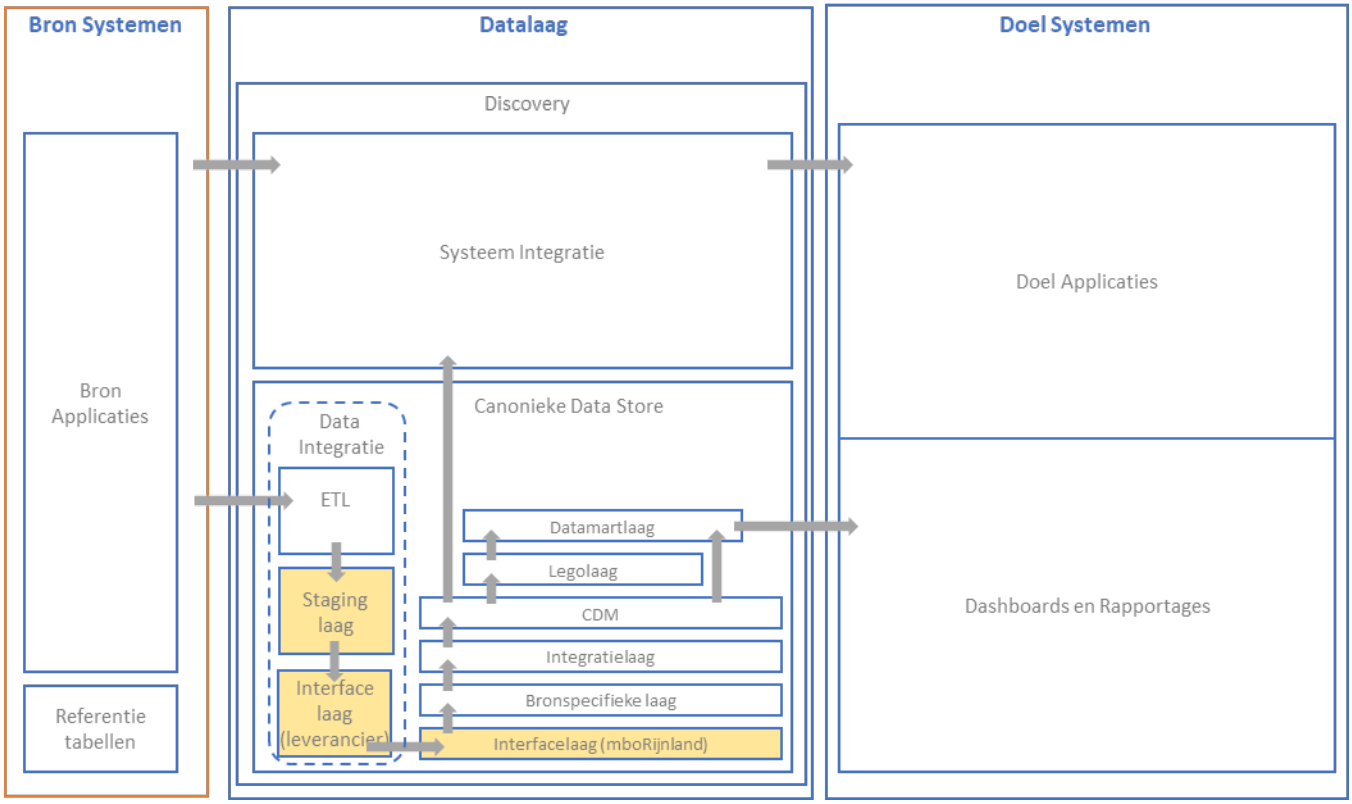
Figuur 8 – Bron- en integratiecomponenten van de datalaag

Principe	Motivatie	Omschrijving
Data wordt aangeboden middels API's	Ten behoeve van data export wordt data ontsloten via API's. Als een API niet mogelijk is dan kunnen andere protocollen overwogen worden. De voorkeur gaat uit naar open standaarden.	
Per databron wordt een datacontract afgesloten met de requirements t.a.v. beoogde aanlevering.	Het datacontract zijn afspraken tussen aanleverende en afnemende partijen. In dit contract gaat het o.a. over eigenaarschap, entiteiten, frequentie etc. Deze datacontracten worden beheerd door het datateam	
Per data afnemer wordt een datacontract afgesloten met de requirements t.a.v. beoogde verwerking.	Naast een datacontract met de aanleverende partij, wordt ook een datacontract afgesloten met de data consumerende partij. Het datacontract zijn afspraken tussen de eigenaren van de datalaag en afnemende partijen. In dit contract gaat het o.a. over eigenaarschap, entiteiten, frequentie etc.	

In de integratiecomponenten (systeem integratie en data integratie) wordt geen persistente data opgeslagen.	In de integratiecomponenten wordt geen data gepersisteerd t.b.v. data retrieval en informatie vraagstukken. Indien gewenst kan data zeer tijdelijk worden opgeslagen t.b.v. synchronisatie doeleinden. In het batch georiënteerde deel van de data- kan de data gepersisteerd worden in de vorm van bestanden voor transmissie.	
Elk data transport heeft een eigen herleidbaar geïdentificeerde eigenaar, middels een gedefinieerd account.	Ten behoeve van beheersbaarheid en controleerbaarheid wordt voor ieder data transport een eigenaar gedefinieerd (gedefinieerd account). Dit account acteert als een service account. Bijbehorende wachtwoorden zullen worden beheerd volgens de, vanuit beheer, gedefinieerde principes.	
Privacy by design	Alle data die beschikbaar wordt gesteld aan derden is getoetst en gevalideerd tegen de privacy en compliance regels zoals opgesteld door mboRijnland.	
Toetsing met BIV classificatie alvorens gegevens beschikbaar worden gesteld.	Alle data die aan derden beschikbaar worden gesteld is getoetst t.a.v. de BIV classificatie.	
Alle transacties worden gelogd in de datalaag.	Ten behoeven van herleidbaarheid van de security regels, performance metingen, volledigheidscontroles en foutafhandelingen (garbage collector) worden alle transacties naar de datalaag toe, binnen de datalaag en bij data consumptie vanuit de datalaag gelogd. Dit maakt het mogelijk om te kunnen voldoen aan de data contract met derden.	
Het integratiecomponent kan transformatie uitvoeren	Het systeem integratie component is geschikt om transformaties tussen de aanbieder en afnemer conform het gesloten datacontract te verwerken.	
De system integratie component is idempotent, stateless en herstartbaar	Een niet werkende interface tussen een aanleverende omgeving en de systeem integratie component of de interface tussen de systeem integratie component en een afnemende omgeving heeft geen invloed op de werking van de datalaag. Indien een operatie nogmaals wordt gestart, zullen de eigenschappen van de transformatie niet wijzigen.	

Staging-laag en Interface lagen (leverancier en mboRijnland)

De eerste laag waar de ruwe data landt is in de staginglaag, gevolgd door de Interfacelaag Leverancier, gespiegeld in de Staginglaag mboRijnland in het CDS



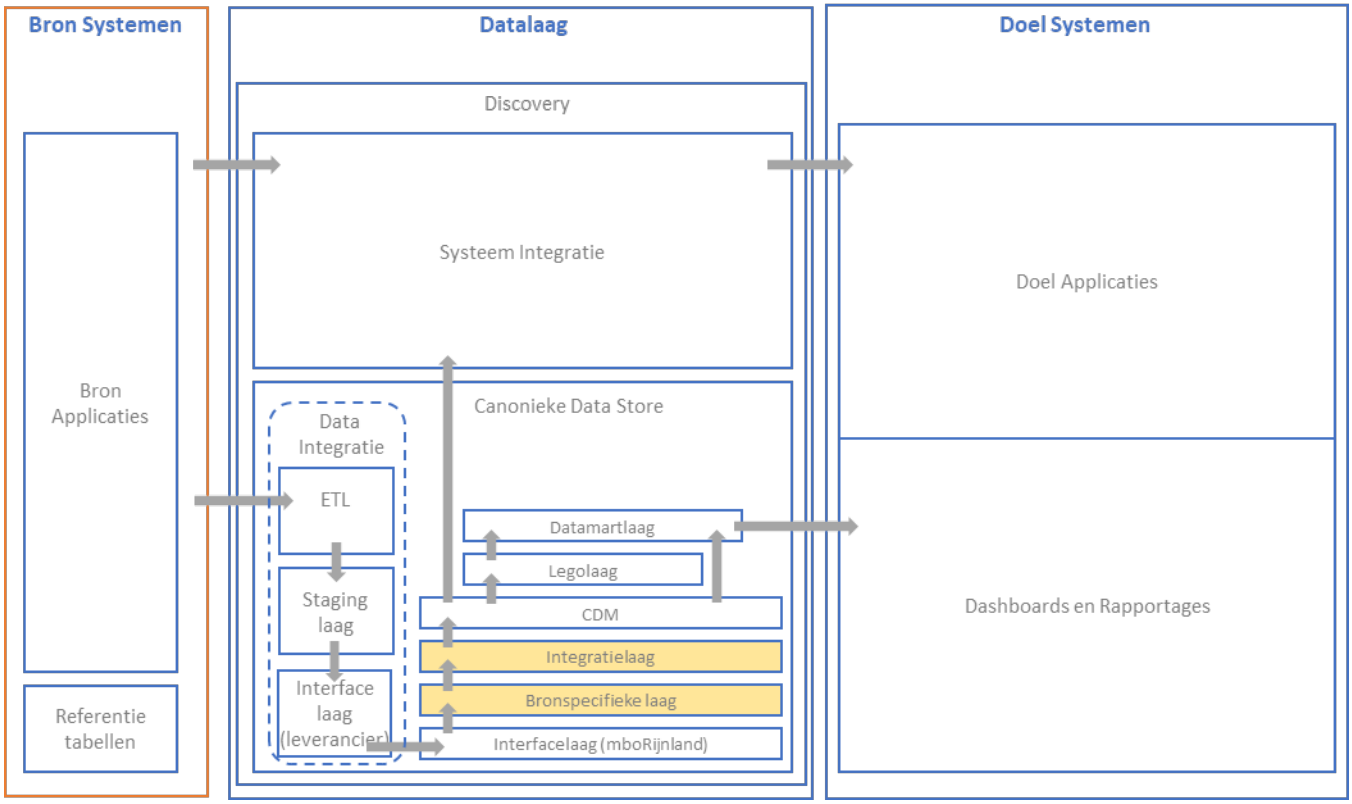
Figuur 9 – Staging en Interface lagen

Principe	Motivatie	Omschrijving
Alleen de benodigde data wordt opgeslagen (entiteiten), met alle attributen (conform het afgesloten data contract).	Per bron wordt bepaald welke data wordt getransformeerd. Alleen de benodigde data entiteiten worden opgenomen in de datalaag. Een data entiteit is benodigd indien deze gebruikt wordt in het vullen van het CDM en/of het creëren van data/informatie tbv rapportage.	
Elk data transport heeft een eigen herleidbaar geïdentificeerde eigenaar, middels een gedefinieerd account (conform het afgesloten data contract).	Data transport van en naar de Staginglaag wordt gedaan door een herleidbaar proces met een herleidbare eigenaar.	
Elk data transport heeft een duidelijke verantwoordelijke die de juiste uitvoering bewaakt.	De verantwoordelijkheid voor het bewaken van juistheid, tijdigheid, volledigheid, etc. is belegd.	
In de ruwe data laag (data store), wordt originele data bewaard.	In de ruwe data laag (data store) wordt de ongewijzigde brondata (w.o. bestanden) na verwerking bewaard, zodat indien noodzakelijk deze altijd nog kunnen worden geraadpleegd.	
Data wordt op een uniforme wijze opgeslagen binnen de Staginglaag.	Voorbeeld van verschillende data verschijningsvormen (SQL, JSON etc.)	
De staging laag bevat het laatst bekende voorkomen van een entiteit.	In de staging laag wordt het laatst bekende voorkomen vastgehouden van een entiteit. Wanneer een entiteit of	

	attribuut in een entiteit wordt verwijderd, bevat de Staginglaag het laatste voorkomen.	
Privacy by design.	Alle data die beschikbaar wordt gesteld aan derden is getoetst en gevalideerd tegen de privacy en compliance regels zoals opgesteld door mboRijnland.	
Toetsing met BIV classificatie alvorens gegevens beschikbaar worden gesteld.	Alle data die aan derden beschikbaar worden gesteld is getoetst t.a.v. de BIV classificatie.	
Alleen gegevens die voldoen aan de doelmatigheid van AVG (afnemers) worden getransformeerd en indien nodig opgeslagen.	Alleen gegevens die voldoen aan de AVG en het beleid van mboRijnland worden getransformeerd en waar nodig opgeslagen.	

Bronspecifieke laag en Integratielaag

In deze paragraaf worden de principes voor de bronspecifieke laag en de Integratielaag opgesteld.



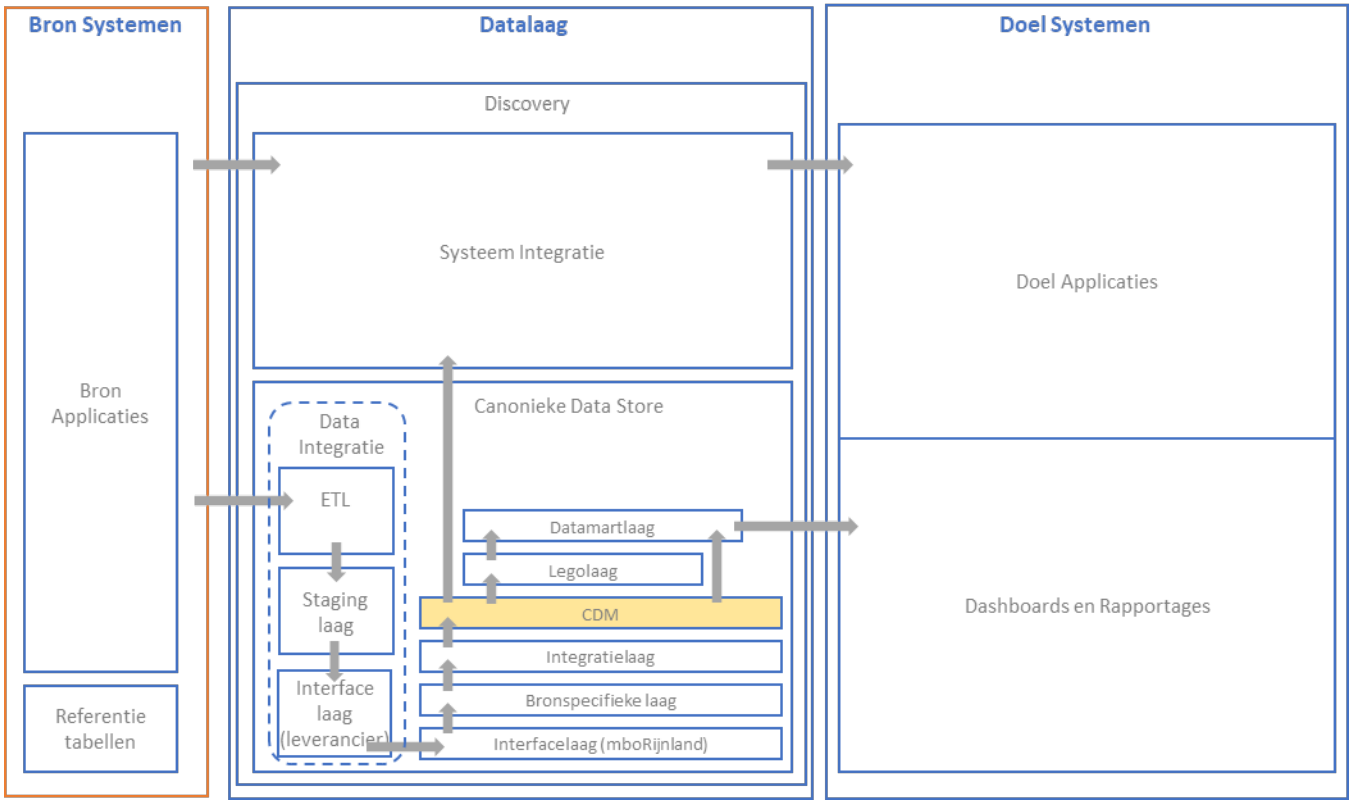
Figuur 10 – Bronspecifieke laag en Integratielaag

Principe	Motivatie	Omschrijving
Alleen de benodigde entiteiten (inclusief alle bijbehorende attributen) wordt opgeslagen.	Per bron wordt bepaald welke data wordt getransformeerd. Alleen de benodigde entiteiten worden getransformeerd (inclusief alle bijbehorende attributen) ten behoeve van de gedefinieerde afnemers.	
De Integratielaag bevat de totale (benodigde) dataset.	In de Integratielaag worden de delta's vanuit de bronspecifieke laag samengevoegd, zodat een totaal beeld van een entiteit ontstaat.	
Alle benodigde entiteiten zijn gekwalificeerd, hebben een eigenaar en zijn integer.	In de Integratielaag wordt de data integer opgeslagen.	
Security by design.	Afnemende services hebben geen directe toegang tot de Integratielaag.	

Privacy by design.	Privacy wordt direct meegewogen. Alle attributen in de Integratielaag hebben een BIV kwalificatie toegekend gekregen (middels master data management). Bij hoge vertrouwelijkheid worden beschermende maatregelen genomen.	
mboRijnland specifiek bron transformaties.	mboRijnland specifieke brontransformaties moeten ondersteund en beheerd worden, zodat specifieke brondata geregistreerd wordt in de Integratielaag (voornamelijk gericht op historische data), waarbij instelling specifieke afwijkingen uit bronsystemen weg gefilterd worden.	
Alle benodigde entiteiten en attributen zijn voorzien van een definitie en vormen gezamenlijk het CDM van het CDS.	Beschikbaarheid van definities maakt het werken met gegevens eenduidig en beheersbaar.	

Canonieke Data Model laag (CDM) mboRijnland

Het canonieke model is een verzameling database objecten (views en tabellen) die gevoed worden met data afkomstig van de verschillende informatiesystemen die binnen mboRijnland in gebruik zijn. Het CDM heeft een genormaliseerde structuur (derde normaalvorm) en is o.a. gebaseerd op entiteiten die onderkend worden in de standaard datamodellen MORA en RIO. Het CDM is qua structuur dus brononafhankelijk. Het kan organisatie breed gebruikt kan worden als basis voor rapportage doeleinden en systeem integratie.

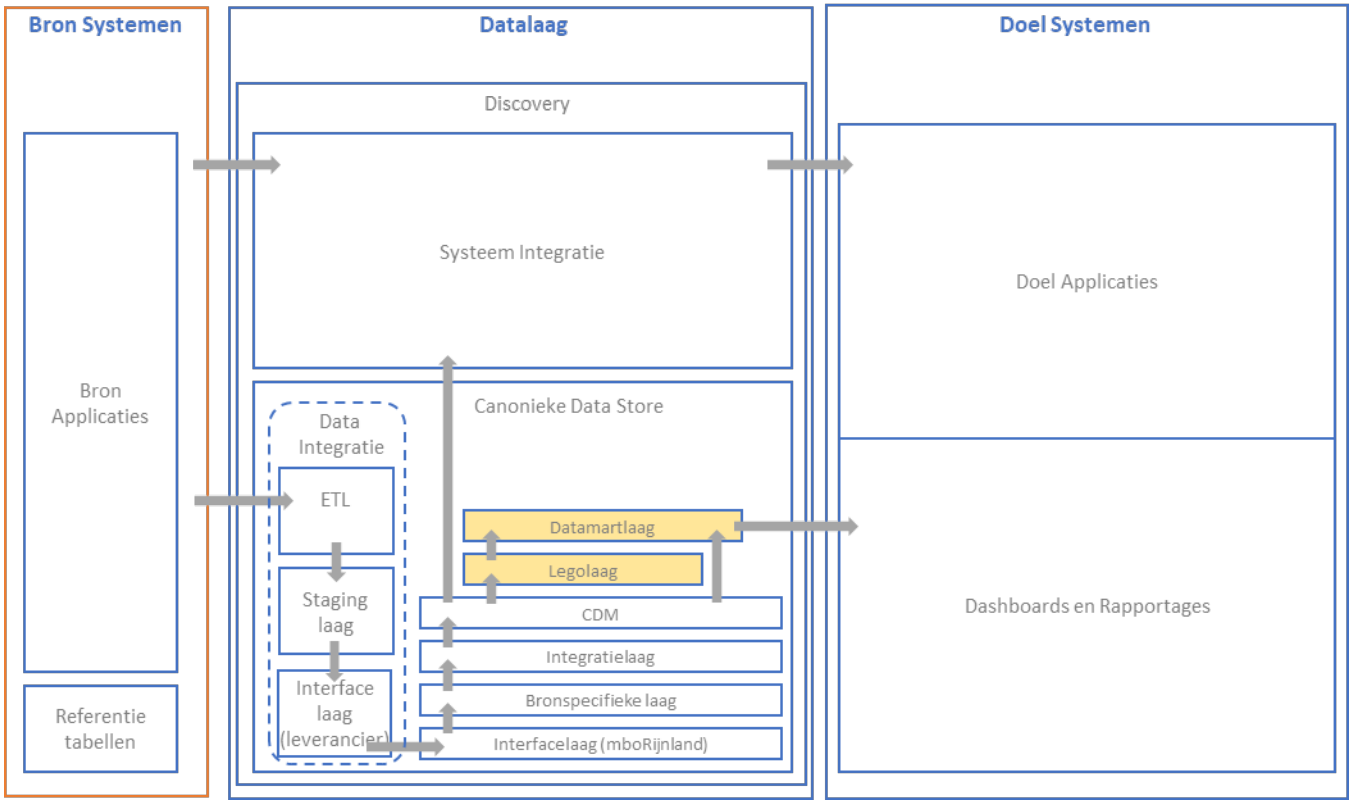


Figuur 11 CDM (Canonieke data model)

Principe	Motivatie	Omschrijving
In het CDM wordt data getransformeerd.	In het CDM kan data getransformeerd worden. Data transformaties (business logic) die nodig zijn om van brongegevens naar een CDM entiteit object te komen wordt uitsluitend vastgelegd in views	
In het CDM worden technische sleutels toegevoegd.	De genormaliseerde structuur en de bronsysteem-onafhankelijkheid van het CDM vereist dat van ieder CDM entiteit object een natuurlijke sleutel (business key) bekend is waarmee een rij uniek geïdentificeerd wordt in de gangbare communicatie tussen medewerkers van mboRijnland.	Bijvoorbeeld een locatie wordt in de wandelgangen uniek geïdentificeerd met een locatiecode (bijv GVP staat voor Groen van PrinstererSingel). In het locatie object in het CDM moet dan een veld bevatten genaamd Code waarvan mag worden aangenomen dat deze unieke waarden bevat. In de bron (momenteel Osiris) wordt een locatie echter uniek geïdentificeerd met een technisch sleutelveld genaamd ID. Dit ID (1034160739 voor Groen van PrinstererSingel) is op de werkvloer niet bekend. In de brontabel staat op het veld Afkorting (in het CDM

		herbenoemd naar Code) geen uniciteits constraint dus er kunnen dubbelingen ontstaan. Om dat te voorkomen zullen maatregelen getroffen moeten worden.
Meerdere bronnen voor een entiteit.	Het is zeer goed denkbaar dat een CDM entiteit object gevuld moet worden met behulp van data uit meerdere bronnen. In dat geval zal er dataintegratie moeten plaats gaan vinden. Dit kan normaal gesproken alleen op basis van de natuurlijke sleutel omdat technische sleutels bronspecifiek zijn.	
Alleen data vanuit de interface laag mag worden gebruikt.	Views die binnen de bronspecifieke laag gecreëerd worden mogen enkel gebruik maken van database objecten in de interface laag.	Een CDM entiteit object moet in bepaalde gevallen gevuld worden met data afkomstig uit meerdere bronnen. Om deze reden is er een apart schema (mboR_Bron) aangemaakt waarbinnen de database objecten moeten worden geplaatst die data uit een enkele bron gebruiken ten behoeve van het vullen van een CDM entiteit object. Dus als er data uit twee bronnen nodig is om een CDM entiteit object te vullen moeten er twee bronspecifieke database objecten worden aangemaakt in schema mboR_Bron die de data uit de respectievelijke bronnen gaan bevatten.
Data kan gepersisteerd worden in het CDM.	Mocht het zo zijn dat een view te traag is, dan kan besloten worden deze view te persisteren in een tabel.	
Feed forward	De data stroom verloopt uitsluitend op een feed forward wijze vanaf de interface laag naar de bronspecifieke laag naar de integratie laag naar de CDM laag. Datastromen in de andere richting zijn niet toegestaan.	

Legolaag en Datamart laag



Figuur 12 – Legolaag en Datamartlaag

Principe	Motivatie	Omschrijving
Alleen entiteiten t.b.v. rapportage en dashboards worden in de datamart laag opgeslagen	Per afnemer wordt bepaald welke data wordt getransformeerd. Alleen de benodigde data entiteiten worden getransformeerd naar de datamartlaag.	
Getransformeerd naar business entiteiten	In de Datamartlaag wordt de data opgeslagen in business entiteiten	
In de Datamartlaag wordt geen historische data gecreëerd.	Historische data wordt overgenomen vanuit Integratie/CDM laag en moet opnieuw gecreëerd kunnen worden in de Datamartlaag	
Security by design	Afnemende s hebben geen directe toegang tot de Datamartlaag.	
Privacy by design	Alle attributen in de Datamartlaag hebben een BIV kwalificatie toegekend gekregen (middels master data management).	

Bijlage D : Infrastructuur architectuur principes

Stabiele infrastructuur
De ICT infrastructuur is stabiel. In de onderwijsuitvoering speelt ICT steeds meer een rol. De beschikbaarheid van ICT systemen is rand voorwaardelijk om een groot deel van het werk te kunnen doen. Hiervoor zullen de juiste SLA's afgesloten moeten worden
Implicaties
• Er dient een robuust netwerk te zijn • Er moeten afspraken gemaakt worden t.a.v. de SLA's

Bijlage E : Beveiliging architectuur principes

Principe	Motivatie	Omschrijving
Voldoen aan externe eisen (wet- en regelgeving) ten aanzien van informatiebeveiliging en privacy	Vanuit wet en regelgeving zijn er eisen gesteld t.a.v. de informatiebeveiliging en privacy van de datalaag.	
Voldoen aan wet en regelgeving bewaartermijnen	Vanuit wet en regelgeving zijn er eisen gesteld aan de bewaartermijnen van data binnen mboRijnland en daar mee de datalaag. Daarbij geldt dat: Gestructureerde gegevens niet in datalaag gearchiveerd worden; Wet en regelgeving moet wel in acht worden genomen. Check wordt gedaan op zowel datalaag als archief system	
Koppelvlak archivering (archivering ligt buiten de datalaag)	Vaststellen requirements vanuit het project document op order. Dit is de basis om de requirements vast te stellen voor de datalaag en eventuele interfaces	
Eenduidigheid, inzichtelijkheid en control over informatiebeveiliging en privacy		
Informatiebeveiligingsmaatregelen zijn gebaseerd op het informatiebeveiligings-beleid, de BIV classificatie van de betrokken gegevens en een risicoanalyse vanuit procesperspectief	Data wordt BIV gekwalificeerd op data attribuut niveau. Data kwalificatie is verder uitgewerkt in paragraaf.	
Informatiebeveiliging wordt integraal meegenomen bij het ontwerp en de inrichting van digitale diensten en infrastructuur	Eindgebruikers hebben geen direct toegang tot de entiteiten binnen de datalaag, maar alleen via API's en/of end-user tooling. In Paragraaf 6.1 is data security verder uitgewerkt.	
Beperking van kosten voor IBP (het beperken van risico's mag niet leiden tot onnodige uitgaven – teveel sloten op de fiets)	Per maatregel moet worden vastgesteld wat de bijbehorende implementatie en beheerskosten zijn. Op basis van het risico en kosten wordt de prioriteit van IBP maatregelen bepaald	
Bevorderen risico-bewustzijn bij medewerkers van mboRijnland	Een belangrijk facet van de implementatie van de IBP beleid is de borging binnen de mboRijnland organisatie. Om dit bewust zijn te creëren is een adoptie programma noodzakelijk.	
Ondersteuning bij Defense in Depth aanpak	Defense-in-depth is een concept waarbij verschillende typen beveiligingsmaatregelen (zoals procedureel, technisch en fysiek) en verschillende soorten maatregelen (zoals detectief, preventief en correctief) worden gecombineerd. Hierbij wordt uitgegaan van een zero-trust model. Het combineren zorgt ervoor dat het falen van één maatregel wordt opgevangen door andere maatregelen. Binnen het domein informatiebeveiliging wordt de defense-in-depth aanpak toegepast bij het nemen van maatregelen. Beveiliging wordt in meerdere lagen ingericht (procedureel, fysiek en technisch (netwerk, hardware, software, etc.). Binnen de datalaag betekent dit dat er meerdere beveiligingsmaatregelen worden genomen om objecten te beveiligen: er zijn meerdere beveiligingszones in het netwerk, extern netwerkverkeer passeert minimaal twee firewalls van verschillende leveranciers, gegevens die worden uitgewisseld worden bij voorkeur op niveau van de inhoud versleuteld omdat dit	

	end-to-end werkt (in tegenstelling tot versleuteling op transportniveau) <en nog een setje maatregelen: aan te vullen vanuit expertise van Leverancier>.	
Voldoen aan AVG	Vanuit de AVG (Algemene verordening gegevensbescherming) zijn regels opgesteld hoe om te gaan met data. Om aan deze wettelijke eisen te kunnen voldoen, wordt ongeclassificeerde data (tussen en gedefinieerde Staginglaag en Interfacelagen) actief geweigerd. Alleen geclassificeerde data (en AVG "proof") data wordt getransformeerd. Classificatie geschiet in de meta data omgeving van het dataplatform. Gebruik van data van de datalaag is geautoriseerd en wordt gelogt.	
Beschermen van de privacy-rechten van gebruikers en beschermen van de vertrouwelijkheid van persoonsgegevens (Bescherming tegen onjuist gebruik van tot personen herleidbare gegevens)	Vastleggen wie welke vertrouwelijke gegevens wanneer heeft gezien (welk end-point met welke scope). Maximaal bewaar termijn van 3 maanden Detail van logging heeft een relatie met het niveau van BIV classificatie. Hoe hoger geclassificeerd, hoe gedetailleerdere logging.	
privacy-by-design	Alle data die beschikbaar wordt gesteld aan derden is getoetst en gevalideerd tegen de security en compliance regels zoals opgesteld door mboRijnland. Persoonsgegevens die niet expliciet geclassificeerd zijn worden actief geweigerd binnen de datalaag (vanaf Staginglaag wordt actief afgedwongen dat alle data is geclassificeerd)	
Security-by-design	Infrastructuur en datalaag zijn secure by design. Data in transit gaat altijd over een encrypted verbinding. Elk data transport heeft een eigen herleidbaar geïdentificeerde eigenaar, middels een gedefinieerd account.	
Beveiliging op gegevens gaat boven beveiliging op applicatie- of infrastructuurniveau	Middels RBAC/ABAC wordt vastgesteld welk natuurlijk persoon welke gegevens set mag benaderen. Waarbij gegevens beveiliging het meest restrictief is. Dit geldt voor ieder toegang tot gegevens.	
Registratie, controle en rapportering ten aanzien van processen en bewerkingen van gegevens	Alle processen en gegevens bewerking activiteiten worden gelogd. Inrichting middels het gebruik van Azure security centre.	
Onweerlegbaarheid van relevante formele processen en berichtenuitwisseling	Bij formele processen is de integriteit van belang, metadata van de processen dient vastgelegd te worden, zoals bijv. Tijdstip, wie, document zelf, etc. De datalaag draagt zorg voor de borging van de metadata (over de gegevens) van de formele processen.	
Voor alle autorisatie-objecten is aangegeven welke rollen of gebruikers geautoriseerd toegang kunnen krijgen	Onderdeel project Identity en Access management. Mapping van Autorisatie matrix op datalaag entiteiten/attributen (afhankelijk van classificatie)	
Alle toegang tot autorisatie-objecten wordt expliciet geauthentiseerd en geautoriseerd, tenzij	By default is de data lag niet toegankelijk. Op basis van de autorisatie matrix worden rechten verleent.	

deze openbaar toegankelijk is. à Alles is in principe verboden tenzij het uitdrukkelijk is toegelaten.	Dit is een formeel proces waarbij de rechten op data entiteiten/attributen geregistreerd worden in de datalaag. Jaarlijks moet worden beoordeeld of de toegekende toegang tot data entiteiten / atrributen verlopen. Deze moeten opnieuw gevalideerd en verlengd worden. Tijdige signalering is hierbij essentieel.	

Bijlage F : OTAP Scenario’s

Met betrekking tot de OTAP straat van het dataplatform zijn verschillende scenario’s gedefinieerd.
(De T omgeving is momenteel nog niet geimplementeerd bij mboRijnland)

Figuur 13 : OTAP scenario's

Use cases (ofwel scenario’s: waar gebruiken we de OTAP omgeving allemaal voor?)

- Een nieuwe bron ontsluiten
- Een bestaande bron verwijderen
- Een nieuwe versie van een API
- - gebaseerd op speed layer
- - gebaseerd op batch/mboRijnland CDM
- Een update van Microsoft Azure component
- Een leveranciers voert een wijziging door – nieuwe objecten of wijziging van definitie van object of attributen
- Testgegevens vanuit productie anonimiseren & testdata doorvoeren naar OTA
- Leverancier voert een wijziging door in het Leverancier CDM
- Valideren van een conversie
- Een proces-wijziging (die meerdere systemen omvat; dus een keten-test & keten-release)

In de volgende paragrafen worden verschillende scenario’s verder uitgewerkt. Per paragraaf worden de principes gedefinieerd.

Een nieuwe bron ontsluiten

- De nieuwe bron noemen we “Nieuw Bron Systeem” (NBS).
- De leverancier van NBS stelt een API beschikbaar vanuit een NBS-O-omgeving beschikbaar.
- Wanneer er een NBS aan het platform toegevoegd moet worden, wordt dit gedaan binnen de development omgeving van Leverancier.
- Leverancier doet een intake op de betreffende bron.
- Leverancier krijgt de beschikking over representatieve afslag van de aan te sluiten bronomgeving.
- De SLA moet worden opgesteld/aangepast.
- De extra bron wordt opgenomen in een apart schema binnen de Staginglaag (ruwe data).
- Toetsen (en mogelijk implementatie) van security en compliancy richtlijnen.
- Verschillende data attributen classificeren.
- De data uit deze extra bron wordt via de “Interface laag Leverancier” beschikbaar gesteld.
- De extra bron wordt gepropageerd over de verschillende omgevingen.
- De ontwikkel-test vindt plaats in de O-omgeving van de datalaag.

Een bestaande bron verwijderen

- De bestaande bron noemen we “Bestaand Bron Systeem” (BBS).

- Voor de bestaande bron moet worden bepaald wat er met de reeds opgeslagen data moet gebeuren.
- Koppeling wordt verwijderd uit de oplossing.
- Aanpassen Interface laag Leverancier (indien noodzakelijk).
- Aanpassen metadata/masterdata management.
- Impact op het mboRijnland canonieke model moet worden bepaald en indien noodzakelijk aangepast en gerelateerde rapporten/datasets/API's.
- De SLA moet worden aangepast.
- De aanpassingen voor verwijderen BBS propageren over de verschillende omgevingen.
- De ontwikkel-test vindt plaats in de O-omgeving van de datalaag.

Een nieuwe versie van een API

- Analyse nieuwe API, vaststellen verschillen.
  - Bepalen impact systeemintegratie
  - Bepalen impact op Data Integratie
- Aanpassen datacontract.
- Toetsen (en mogelijk implementatie) van security en compliancy richtlijnen.
- Nieuwe data attributen classificeren.
- Vatstellen impact op business behoefte.
- Aanpassen "Interfacelaag Leverancier" laag (indien noodzakelijk).
- Aanpassen mboRijnland CDM (indien noodzakelijk) en gerelateerde rapporten/datasets/API's.
- De ontwikkel-test vindt plaats in de O-omgeving van de datalaag.

Een update van Microsoft Azure component

- Analyseer de update van een Microsoft Azure component
- Bepaal de impact op de verschillende datalaag componenten
- Indien nodig:
  - Bepaal een upgrade scenario
  - Implementeer benodigde aanpassingen
- Pas de bijbehorende blueprints aan

Functionele wijzigingen in bron applicaties

- Analyse de functionele wijziging.
- Vatstellen impact op business behoefte.
- Indien ook wijzigingen aan bron applicatie attributen
  - Bepalen impact systeemintegratie
  - Bepalen impact op speedlayer
  - Bepalen impact op batchlayer
  - Aanpassen "loosely coupling" laag (indien noodzakelijk).
- Toetsen (en mogelijk implementatie) van security en compliancy richtlijnen.
- Nieuwe data attributen classificeren.
- Aanpassen mboRijnland CDM (indien noodzakelijk) en gerelateerde rapporten/datasets/API's.
- De ontwikkel-test vindt plaats in de O-omgeving van de datalaag.

Anonimiseren testgegevens

- Bepalen te anonimiseren attributen.
- Vatstellen anonimiserings strategie.
- Vatsleggen anonimiserings strategie en scripts.
- Bepalen data laag waar het data geanonimiseerd moet worden.
- Toetsen (en mogelijk implementatie) van security en compliancy richtlijnen.
- Anonimiseren van data.
- Valideren resultaat data anonimisering.
- De ontwikkel-test vindt plaats in de O-omgeving van de datalaag.

Wijziging Leverancier Staginglaag

- Wijzigingen worden besproken met de applicatie eigenaar van het mboRijnland canonieke model

- Wijzigingen in het Leverancier Staginglaag worden geabsorbeerd door de "Interfacelaag Leverancier" laag
- Wijzigingen in de Staginglaag hebben (uitzonderingen daargelaten) geen impact op het mboRijnland canonieke model.
- Toetsen (en mogelijk implementatie) van security en compliancy richtlijnen.
- Indien nodig nieuwe data attributen classificeren.
- De ontwikkel-test vindt plaats in de O-omgeving van de datalaag.

Valideren bron conversie

- Nadat een nieuwe bron is toegevoegd, kunnen de transformatie scripts worden gebruikt t.b.v. conversie.
- De Staginglaag is geschikt om data te valideren en transformeren.

Wijziging mboRijnland proces

- Analyse proces wijziging.
- Vatstellen impact op business behoefte.
- Indien ook wijzigingen aan bron applicaties en attributen
    - Bepalen impact systeemintegratie
    - Bepalen impact op Data Integratie
    - Aanpassen "Interfacelaag Leverancier" laag (indien noodzakelijk).
- Toetsen (en mogelijk implementatie) van security en compliancy richtlijnen.
- Nieuwe data attributen classificeren.
- Aanpassen mboRijnland CDM (indien noodzakelijk) en gerelateerde rapporten/datasets/API's.
- De ontwikkel-test vindt plaats in de O-omgeving van de datalaag.