



Whitepaper: De architectuur achter modern data platforms

Van no-code integratie tot volwassen data-operatie, van data naar impact

Datum

08/10/2025



www.intodata.nl

Inhoudsopgave

	1
Whitepaper: De architectuur achter modern data platforms	1
Samenvatting	3
Architectuuroverzicht platform	4
Data Products	5
Connectiviteit & data-onboarding	5
Federated query engine	6
Opslag: BYO-warehouse of ingebouwde DW	6
Pipelinebouw & transformaties	7
Governance & Data Mesh	7
Activatie + Data-producten	8
AI & ML integratie	9
Beveiliging & Compliance	9
Business Cases uit de praktijk	10
Conclusie & roadmap	12

Samenvatting

Data en IT blijft groeien. De hoeveelheid databronnen gaat omhoog, de vraag naar inzichten neemt sneller toe dan de capaciteit, en het tempo van de business past niet meer bij trage data-iteraties.

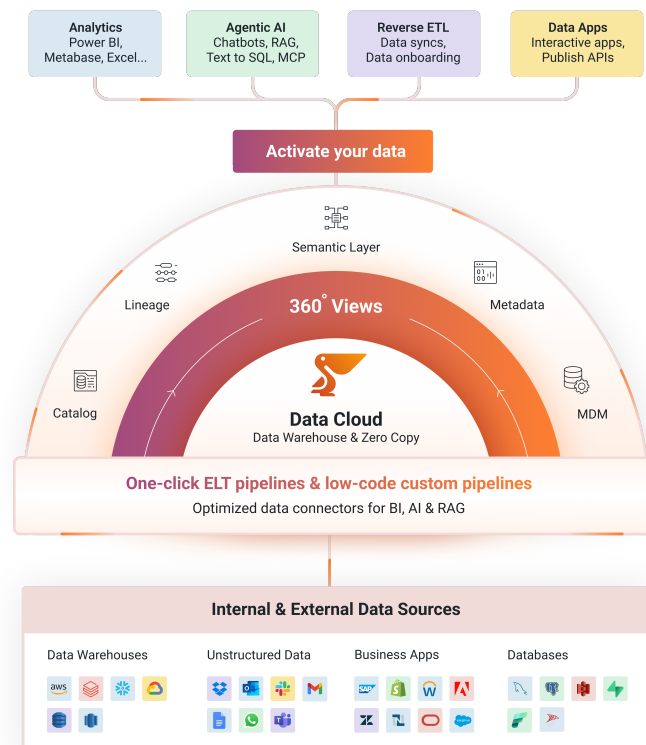
Nieuwe concepten als *self-service analytics*, *no-code integratie* en *data as a product* beloven snelheid en wendbaarheid, maar brengen ook risico's: versnippering, technische schuld en gebrek aan governance.

Deze whitepaper laat zien hoe een modern data platform, opgebouwd uit flexibele componenten, organisaties helpt om van losse pipelines over te stappen naar herbruikbare dataproducten.

Aan de hand van Peliqan als voorbeeld ontdek je:

- hoe een cloud-native, modulaire architectuur samenwerking tussen IT, BI en business versnelt;
- hoe federated querying, reverse ETL en data activation directe waarde creëren in operationele processen;
- hoe governance en het *Data Mesh*-principe controle bieden zonder innovatie te remmen;
- hoe AI & machine learning laagdrempelig geïntegreerd kunnen worden;
- en concrete praktijkcases waarin organisaties uren besparen, kosten reduceren en tientallen databronnen ontsluiten.

Met deze inzichten kun je zelf bouwen aan een toekomstbestendig data platform en groeien richting selfservice, data-gedreven besluitvorming en AI-ready integratie.



Architectuuroverzicht platform

In dit eerste hoofdstuk ontdek je hoe de architectuur van Peliquan is opgebouwd en waarom deze laag voor laag is ontworpen om **samenwerking, hergebruik en controle te ondersteunen**. Dit helpt je te begrijpen hoe een modern data platform robuust en schaalbaar kan blijven terwijl het flexibel integreert met bestaande tools.

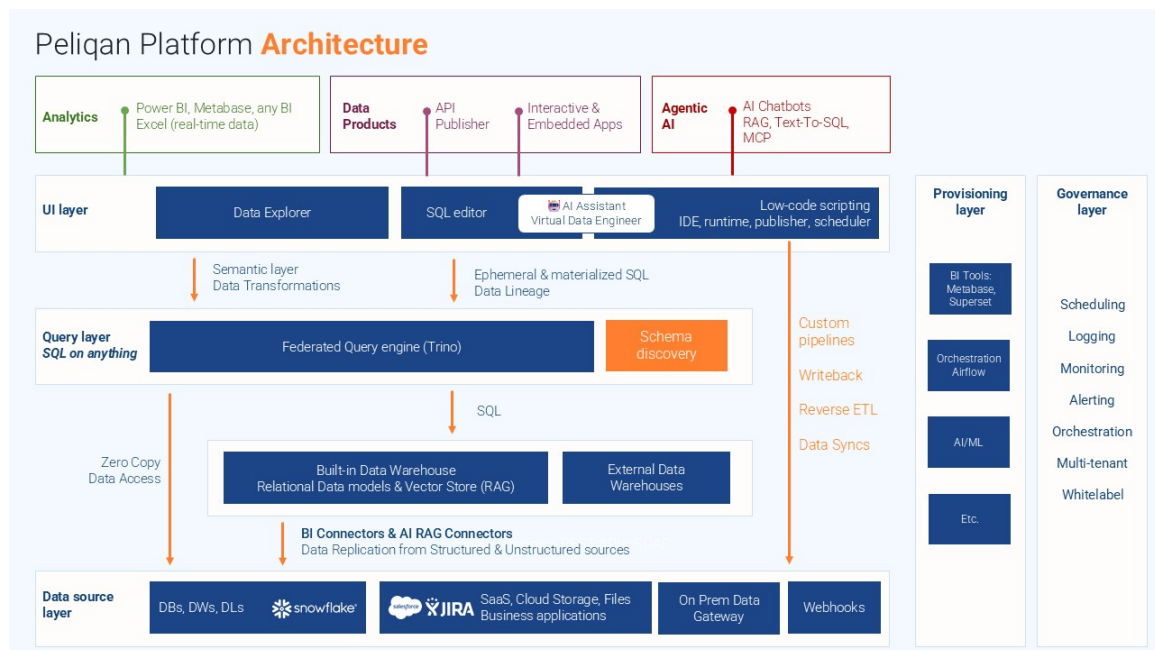
Peliquan draait op een cloud-native **microservicesarchitectuur**, doorgaans binnen Kubernetes. Elke kernfunctionaliteit — zoals connectiviteit, transformatie, opslag, metadata of monitoring — wordt afgehandeld door een aparte service. Dit maakt het platform schaalbaar en veerkrachtig.

De architectuur van Peliquan bestaat grofweg uit drie functionele lagen: de Data Interaction Layer (gebruikersinterface), de Data Processing & Orchestration Layer (transformaties, query's, logging), en de Data & Metadata Layer (opslag, catalogus, beveiliging).

Deze gelaagdheid zorgt ervoor dat Peliquan flexibel kan worden ingezet in zowel kleine als grotere organisaties. Gebruikers kunnen klein beginnen en eenvoudig **opschalen zonder technische schuld** op te bouwen.

Peliquan is bovendien ontworpen als een composable platform. Het integreert naadloos met andere tools in de modern data stack. Daarmee is het geen gesloten ecosysteem, maar een uitbreidbare en aanpasbare bouwsteen.

De Peliquan Platform Architectuur



Data Products

De eerste stap in de ontwikkeling van een dataproduct is altijd de business case. Bedrijven moeten starten vanuit een duidelijk doel. Dit vraagt om een andere benadering vanuit het data-team.

"Zij zijn eigenaar van het probleem. Jij bent eigenaar van de oplossing."

Dat betekent:

- Je luistert naar de reden achter de aanvraag
- Je diagnosticeert de werkelijke behoefte van de gebruiker
- Je begrijpt wat het bedrijfsdoel achter de aanvraag is
- Je begrijpt hoe ze hun probleem vandaag de dag oplossen
- Je begrijpt waarom hun huidige aanpak hen in de steek laat
- Je begrijpt hoe hun wereld eruit ziet als ze het probleem niet oplossen
- Je bepaalt welke data-oplossing (indien van toepassing!) zal helpen

Pas nadat dit duidelijk is volgen de volgende technische stappen voor het opzetten van de dataoplossing.

Connectiviteit & data-onboarding

In dit deel kijken we naar de cruciale eerste stap van elk data-initiatief: het ontsluiten en onboarden van databronnen. In de praktijk betekent dit: toegang krijgen tot ERP-systemen, SaaS-apps, spreadsheets, CRM-platforms, API's en databronnen waarvan niemand nog weet dat ze bestaan.

Pelican biedt standaard **meer dan 250 connectoren** naar veelgebruikte systemen. Deze connectoren kunnen met één klik worden ingesteld en ondersteunen authenticatie, incremental loading en retry-mechanismen.

Voor minder bekende systemen of klantspecifieke toepassingen kunnen datateams de connector builder gebruiken. Deze maakt het mogelijk om aangepaste connectoren te bouwen met logging, validatie en herbruikbaarheid ingebouwd.

Onboarding is niet beperkt tot alleen lezen. Met **reverse ETL** stuurt Pelican data ook terug naar operationele systemen, zoals Salesforce, HubSpot, Slack of een interne database. Hierdoor kunnen inzichten direct worden geactiveerd in de workflow.

Hiermee volgen we het paradigma dat het dashboard niet het einde is van de data pipeline, maar dat je het moet gebruiken in de systemen waar het nodig is, data activatie.

"Self service data activation is a strategy that empowers **organizations and business teams to work with data**, beyond the classical BI reporting or a dependency on data engineering teams."

Federated query engine

Als de data connectiviteit geregeld is, dan is het aan de eindgebruiker om hiermee aan de slag te gaan. Om dit te faciliteren en zo gemakkelijk mogelijk te maken wil je dat de eindgebruiker geen kennis hoeft te hebben van de onderliggende systemen.

Federated querying is de mogelijkheid om query's uit te voeren op meerdere databronnen zonder dat je deze eerst hoeft te repliceren naar een centrale datawarehouse.

Pelican maakt dit mogelijk via een engine gebaseerd op [Trino](#) (voorheen PrestoSQL). Hierdoor kunnen gebruikers data combineren vanuit bijvoorbeeld een MySQL-database, een Excel-sheet en een cloud warehouse — in realtime.

De federated query engine is krachtig voor situaties waarin data actueel moet blijven en latency cruciaal is. Denk aan audit dashboards, logistieke monitoring of realtime productiedata.

Caching en optimalisatie zijn onderdeel van de engine: query's worden waar mogelijk vereenvoudigd of gedeeltelijk geëxecuteerd per bron. Pelican beheert de afhankelijkheden automatisch.

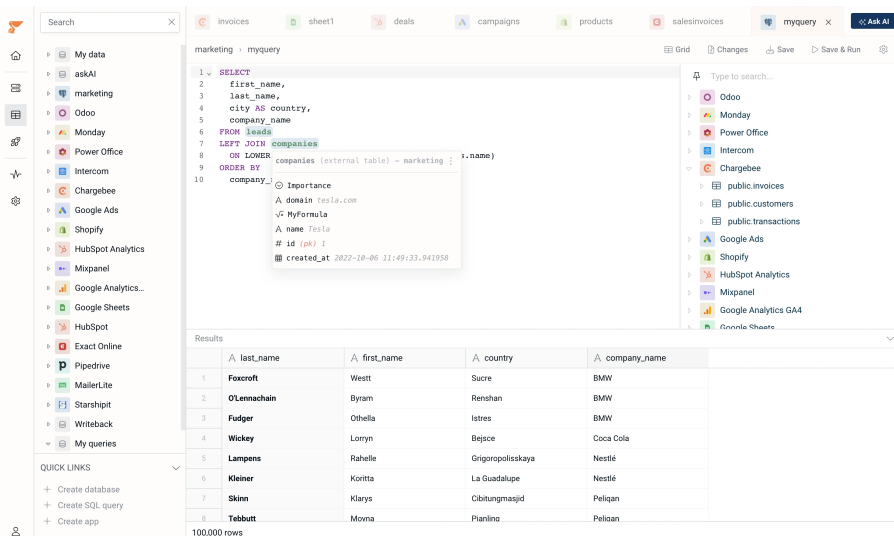
Voor data-teams en eindgebruikers betekent dit: minder ETL, minder opslag, en meer flexibiliteit om analyses direct uit te voeren waar de data zich bevindt.

Opslag: BYO-warehouse of ingebouwde DW

Niet alle use-cases lenen zich voor directe querying op de systemen. Hiervoor is het gebruikelijk om een omgeving op te zetten waarin de data is gerepliceerd, zodat de eindgebruikers zelf hun hypotheses kunnen toetsen en experimenteren.

Nog niet elke organisatie heeft een bestaand datawarehouse. Daarom biedt Pelican een [ingebouwde PostgreSQL-gebaseerde opslagoptie](#) die snel te gebruiken is voor kleine tot middelgrote omgevingen.

Voor organisaties die al een datawarehouse gebruiken, zoals [Snowflake](#), [BigQuery](#) of [Redshift](#), is Pelican eenvoudig te koppelen via veilige integratie. Dit betekent dat de rekenkracht en beveiliging van je bestaande stack behouden blijft.



Het ingebouwde datawarehouse van Pelican.

Deze flexibiliteit betekent ook dat opslag en compute gescheiden blijven, wat schaalbaarheid en kostenefficiëntie ondersteunt. Gebruikers kunnen pipelines laten draaien in het platform, maar data bewaren op hun eigen infrastructuur.

Pelican ondersteunt ook meerdere opslaglocaties binnen één project. Dit maakt het mogelijk om bijvoorbeeld historische data centraal te houden, en operationele data lokaal te verwerken.

Pipelinebouw & transformaties

Transformaties vormen de kern van data-analyse: het proces van ruwe data omzetten naar iets bruikbaar. In Pelican gebeurt dit via een visuele pipeline-editor, SQL-scripts of embedded Python-notebooks.

Het principe van self service geeft de eindgebruiker de tools om zelf met de data aan de slag te gaan.

Pipelines bestaan uit stappen: extractie, validatie, transformatie, verrijking en laadacties. Elke stap is configureerbaar en herbruikbaar. Teams kunnen standaardpijplijnen bouwen en delen binnen projecten.

Scheduling en dependency management zijn ingebouwd. Gebruikers kunnen instellen wanneer en hoe vaak een pipeline draait, afhankelijk van upstream bronnen of businessregels.

Bij fouten geeft Pelican automatisch meldingen met logging en debuggingopties. Dit maakt het onderhouden van pipelines minder foutgevoelig.

De combinatie van visueel werken en scripting ondersteunt zowel business users als technische gebruikers.

Governance & Data Mesh

Data governance draait in essentie om **vertrouwen**. Het gaat niet om poortwachtersrol; het gaat erom ervoor te zorgen dat de juiste mensen op het juiste moment toegang hebben tot de juiste data.

Data governance is essentieel om grip te houden op data, zeker als meerdere teams of afdelingen ermee werken. Pelican biedt uitgebreide functies voor datakwaliteit, eigenaarschap en auditing.

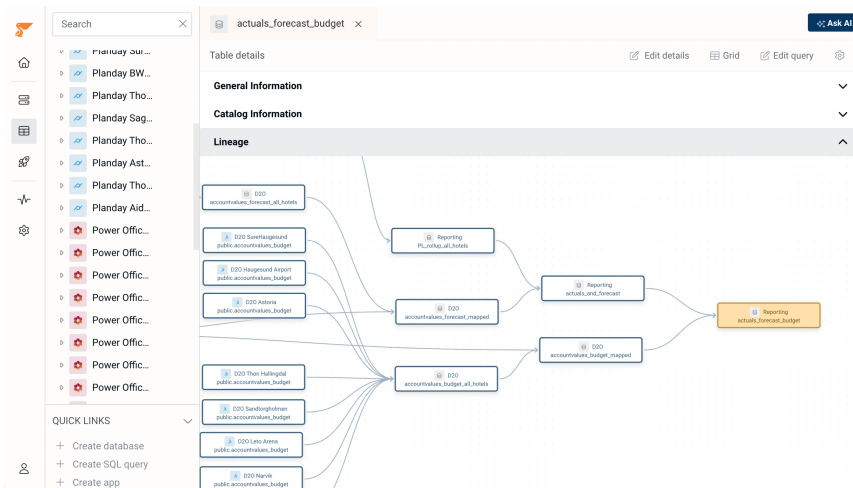
Elke dataset en pipeline heeft **metadata, lineage-tracking en versiebeheer**. Gebruikers kunnen zien waar data vandaan komt, hoe het is verwerkt, en wie het wanneer heeft aangepast.

Rollen en rechten zijn instelbaar tot op veldniveau in de aangesloten BI tool (bijvoorbeeld Power BI of Tableau). Dit betekent dat een marketeer alleen klantnamen mag zien, terwijl een controller ook omzetgegevens mag analyseren.

Pelican ondersteunt het Data Mesh-principe door domeineigenaarschap te stimuleren. Teams kunnen zelfstandig werken binnen hun eigen datadomein, zonder centrale afhankelijkheden.

De centrale catalogus zorgt ervoor dat data vindbaar en herbruikbaar is, ook over teams heen.

Data Lineage zoals in Peliqan te zien.



Activatie + Data-producten

Data is pas waardevol als het wordt gebruikt. Peliqan maakt het mogelijk om data direct te activeren via API's, dataproducten of reverse ETL.

Gebruikers kunnen inzichten direct beschikbaar stellen aan operationele teams — bijvoorbeeld via Slack-notificaties, Excel-plug-ins of updates in CRM-systemen.

Dataprodukten in Peliqan zijn mini-applicaties die op maat gemaakte inzichten tonen, bijvoorbeeld voor een specifieke businessunit of afdeling.

Ook dashboards en rapportages kunnen rechtstreeks vanuit Peliqan worden gedeeld met externe stakeholders, inclusief toegangsbeheer.

Become an **Active** player in your data driven organization

Sync data between applications, create data alerts, automate manual actions, build interactive apps, deploy ML and LLM models, Create your data API's, ..

Build interactive data apps and products.



Data is geen abstract verhaal, maar een product dat bruikbaar is in de dagelijkse praktijk.

AI & ML integratie

Pelican is niet alleen een ETL-tool. Het platform ondersteunt ook machine learning-workflows via notebooks, Python-scripting en modelintegraties.

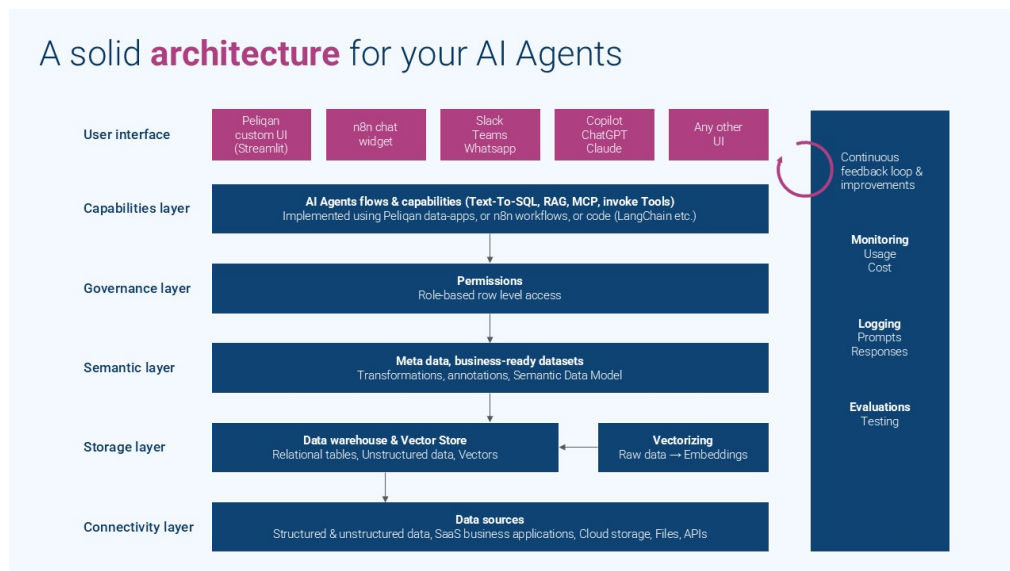
Gebruikers kunnen [voorspellende modellen bouwen binnen het platform](#), of bestaande modellen koppelen aan hun pipelines voor automatisering.

Er is ondersteuning voor vector-databases, waarmee semantisch zoeken, vergelijkingen of aanbevelingen mogelijk worden. Zo kan eveneens een business AI Chatbot opgezet worden die data ontsluit voor niet-technische gebruikers.

Een ingebouwde AI-assistent helpt gebruikers bij het schrijven van SQL-query's. Op basis van natuurlijke taal wordt de gewenste bevraging gegenereerd.

Voor veel organisaties is dit een laagdrempelige manier om te starten met data science, zonder direct een datateam te moeten opbouwen.

Data architectuur die AI en AI-agents mogelijk maakt.



Beveiliging & Compliance

Beveiliging en compliance zijn standaardonderdelen van de Pelican-architectuur. Het platform is gebouwd volgens [security-by-design](#) principes.

Gebruikersauthenticatie gebeurt via OAuth, SSO of API-tokens.

Alle data is versleuteld, zowel in opslag als in transport. Logging en audittrails zijn standaard aanwezig.

Pelican voldoet aan GDPR en ondersteunt SOC2-ready implementaties. Voor gevoelige data kunnen implementaties volledig on-premise of in een private cloud draaien.

Met deze waarborgen is het platform geschikt voor sectoren waar compliance zwaar weegt, zoals finance, zorg of overheid.

Business Cases uit de praktijk

Voorbeeld 1: CIC Hospitality

Een middelgroot hospitality bedrijf heeft een uitdaging met hun financiële en operationele rapportage.

De business case

De CFO van de groep heeft geconsolideerde gegevens nodig voor maandelijkse bestuursrapportages. Deze gegevens moeten afkomstig zijn uit een breed scala aan verschillende systemen, elk specifiek voor de betreffende hotels. Dit is een combinatie van verschillende boekhoudsystemen, planningssoftware en PIM-software.

Uitdagingen hierbij zijn:

- Het consolideren van gegevens uit diverse SaaS-platforms voor hotels is tijdrovend.
- Het standaardiseren van boekhoudgegevens (bijv. het in kaart brengen van rekeningschema's, valutaconversie) is moeilijk.
- Het onderhouden en bijwerken van rapportlay-outs in Google Spreadsheets is omslachtig en het ontbreekt aan flexibiliteit.

De oplossing

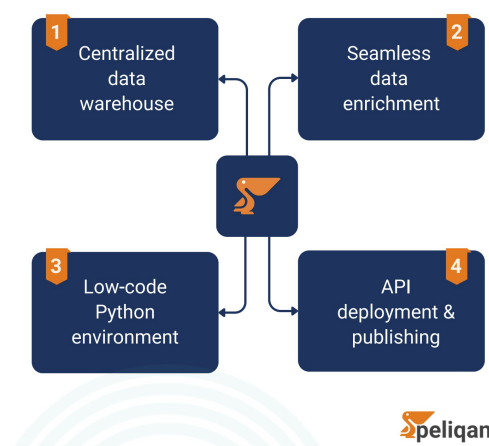
Pelican centraliseert gegevens van verschillende SaaS-platforms voor hotels, waardoor silo's worden geëlimineerd en naadloze analyses mogelijk zijn.

1. Het geïntegreerde rekeningschema in Pelican brengt transacties automatisch in kaart, wat zorgt voor consistentie en nauwkeurigheid.
2. Openbare API-verbindingen binnen Pelican halen actuele valutakoersen op, waardoor handmatige berekeningen worden geëlimineerd en nauwkeurigheid is gegarandeerd.
3. Pelican transformeert en aggregeert gegevens op basis van de Google Sheet-sjabloon en genereert maandelijks volledig geautomatiseerde bestuursrapporten.

De voordelen

- Meer dan 30 uren bespaard per maand
- Meer dan 120 SaaS-bronnen verbonden
- 50 hotels opgenomen in geconsolideerde rapportage

Link: <https://pelican.io/download-case-study-cic-hospitality/>



Voorbeeld 2: Ziggu's eigen integratie-marketplace

Ziggu biedt een klantenportaal dat de communicatie en het documentbeheer voor woningbouwprojecten tussen alle belanghebbenden stroomlijnt.

De business case

De groep heeft behoefte aan efficiënte integratieoplossingen om haar vastgoedklanten te bedienen. Klanten willen gegevens synchroon houden tussen Ziggu en de interne bedrijfsapplicaties die ze gebruiken, bijvoorbeeld door documenten te synchroniseren tussen Sharepoint en Ziggu, taken te synchroniseren tussen de Ziggu-portal en hun projectmanagementtool en projectinformatie te synchroniseren.

Uitdagingen hierbij zijn:

- Klanten van Ziggu vragen om een breed scala aan integraties.
- Het opbouwen van een integratiemarktplaats kost tijd.
- Maatwerkintegraties zijn een verkoop bevorderende factor, maar niet winstgevend.

De oplossing

Pelican is de plug & play marketplace voor vooraf gebouwde data-integraties, voor directe activering door het Customer Success-team of rechtstreeks door eindklanten (selfservicemodel).

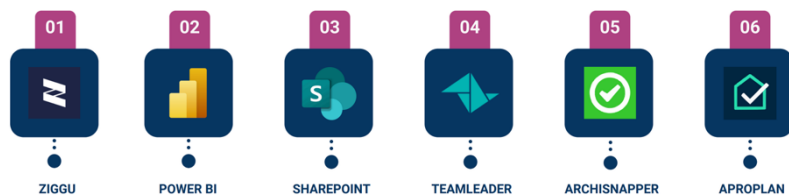
1. Framework en bouwstenen voor het bouwen van maatwerkintegraties. Gestandaardiseerde aanpak voor de implementatie van een maatwerkintegratie op basis van de behoeften.
2. Gecentraliseerd beheer van eindklanten. Hiermee kunnen supportteams assistentie verlenen, bijvoorbeeld door een eindklant uit te nodigen om een verbinding opnieuw te autoriseren.
3. Geïntegreerde front-end UI-widgets bieden inzicht in de status van de datasynchronisaties voor zowel het supportteam als de eindklanten.

De voordelen

- Meer dan 200+ uren bespaard per integratie
- 70% kosten reductie voor de ontwikkeling van interfaces
- Meer dan 20.000 project items per week gesynchroniseerd

Link: <https://pelican.io/download-case-study-ziggu/>

EXAMPLE CONNECTORS USED



Conclusie & roadmap

In deze whitepaper hebben we laten zien hoe een **modern data platform** organisaties kan helpen om hun datamanagement te vereenvoudigen, versnellen en schaalbaar te maken.

De belangrijkste les is dat **de eindgebruiker leidend moet zijn** in het stellen van vragen, experimenteren en activeren van data, terwijl data-teams hen faciliteren met de juiste architectuur en governance.

Wat zijn dan de belangrijkste overwegingen voor een toekomstbestendige data-architectuur:

- Geef **zakelijke gebruikers directe toegang** tot een goed gemodelleerde laag in het datawarehouse.
- Laat hen werken met de tools die ze al kennen en gebruiken.
- Maak gebruik van integraties en *federated querying* om snel waarde te leveren.
- Stimuleer experimenten en hypothese-gedreven analyses in **een veilige sandbox-omgeving**.
- Zorg voor governance en kwaliteitscontrole zonder innovatie af te remmen.

Pelican is ontwikkeld voor moderne data-teams die **snelheid, samenwerking en schaalbaarheid** zoeken, zonder in te leveren op controle. Het platform combineert een krachtige onderliggende architectuur (zoals **Trino, PostgreSQL en Kubernetes**) met een toegankelijke gebruikerservaring voor zowel technische als niet-technische rollen.

Door connectiviteit, transformatie, governance en activatie te verenigen in één platform, wordt het ontwikkelen van dataproducten laagdrempelig en kan de business sneller waarde uit data halen.

De roadmap van Pelican richt zich op verdere uitbreidingen in **AI, selfservice, dataproducten** en diepere integratie met open standaarden in de modern data stack, zodat organisaties toekomstbestendig blijven in een snel veranderend datalandschap.

Wil je weten waar jouw organisatie staat in haar datavolwassenheid of hoe een modern data platform zoals Pelican jou kan helpen?

- **Doe de Data Efficiency Assessment** en ontdek waar je verbeterpotentieel ligt.
- Of **plan een 30-minuten intake** met onze experts om te zien hoe je de stap kunt zetten van losse pipelines naar echte dataproducten.