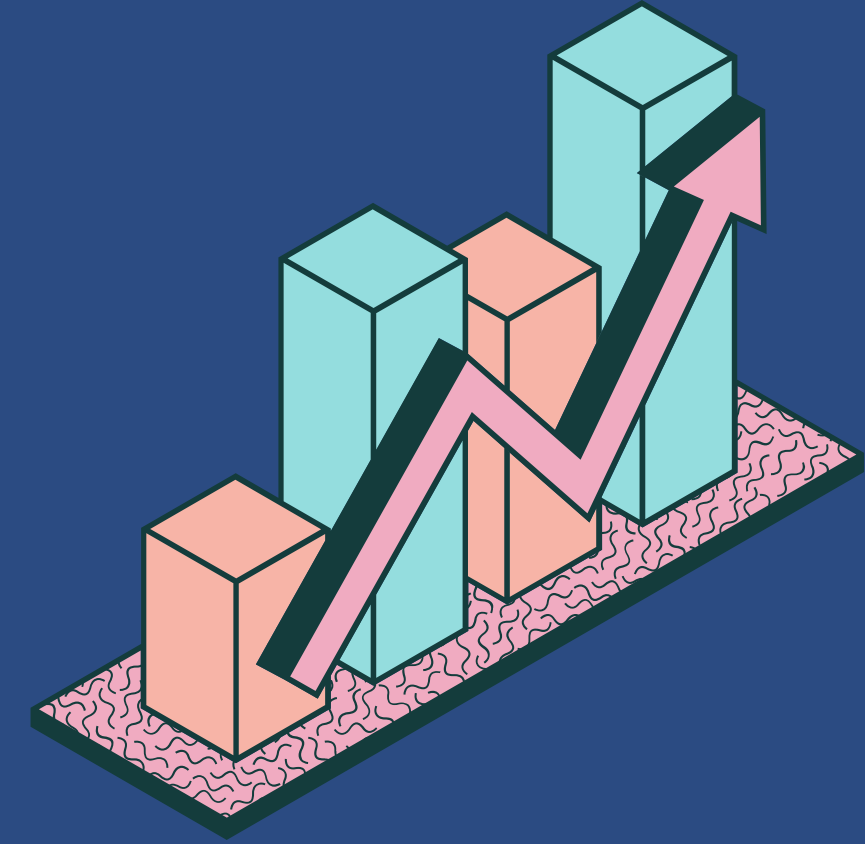
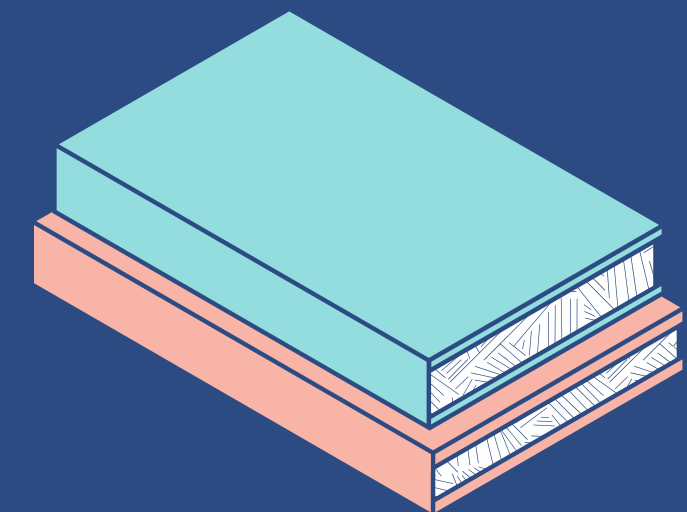


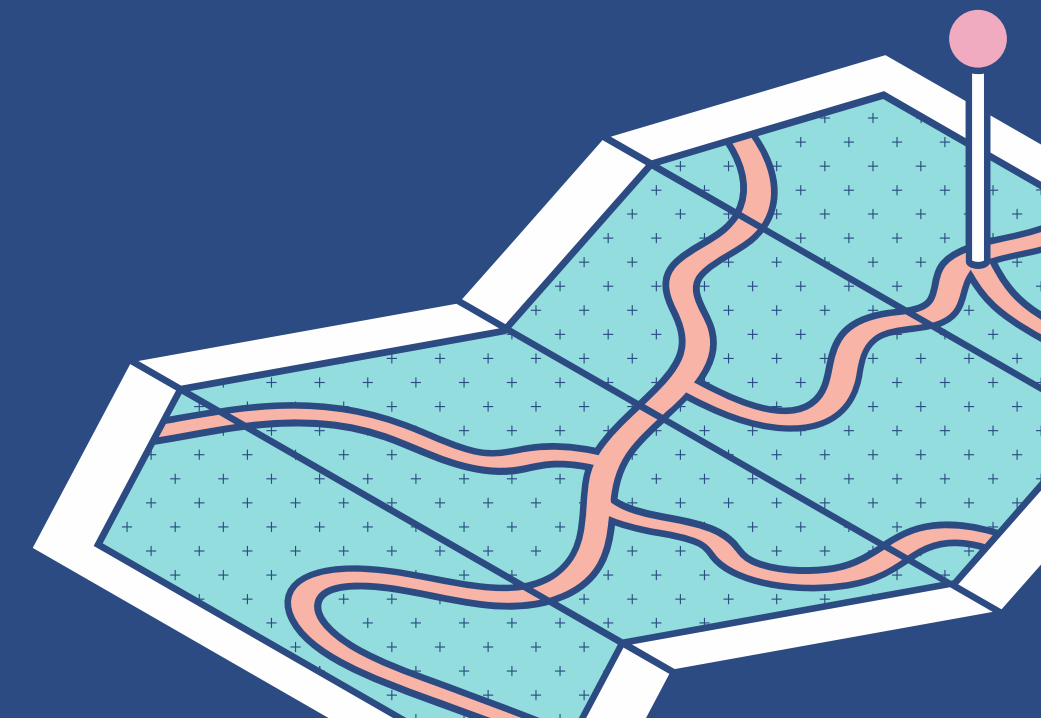
dremio



DREMIO, C'EST QUOI ?



PUBERT Elise, MARREC Guillaume
14/10/2024



Sommaire

1

2

3

4

5

PRÉSENTATION

Introduction à
Dremio

BIG DATA

Son lien avec le Big
Data

HISTORIQUE

de 1980 à
aujourd'hui

APPLICATIONS

Clients, usage et
exemples

DÉMONSTRATION



“The Unified
Lakehouse Platform
for Self-Service
Analytics and AI”

Data Lakehouse

DATA LAKE

- Référenciel de données
- Quantité
- flexibilité
- Exploitation lente



DATA WAREHOUSE

- Référenciel de données
- Quantité
- Hiercharisation
- Fonctions de gestion



DATA LAKEHOUSE

Combine flexibilité et
hierarchisation

DREMIO

CARACTÉRISTIQUES

- Moteur SQL rapide
- Service de gestion Lakehouse
- Flexibilité de déploiement de cloud hybride.



SIMPLIFIE L'ACCÈS AUX DONNÉES

Sans entrepôts de données, datamarts, extraits BI ni de cubes. Centralisation des données.



COMPATIBILITÉ POUR L'INTÉGRATION

Fusion des données ADLS avec des données d'autres référentiels.



ACCÉLÉRATION DES REQUÊTES

Optimisations comme l'indexation ou le caching

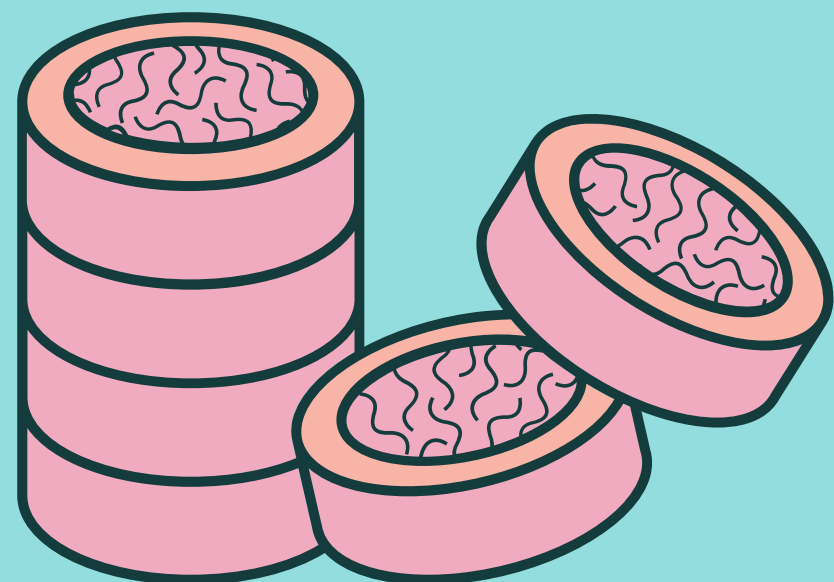


CONTRÔLE D'ACCÈS

Pour + de sécurité

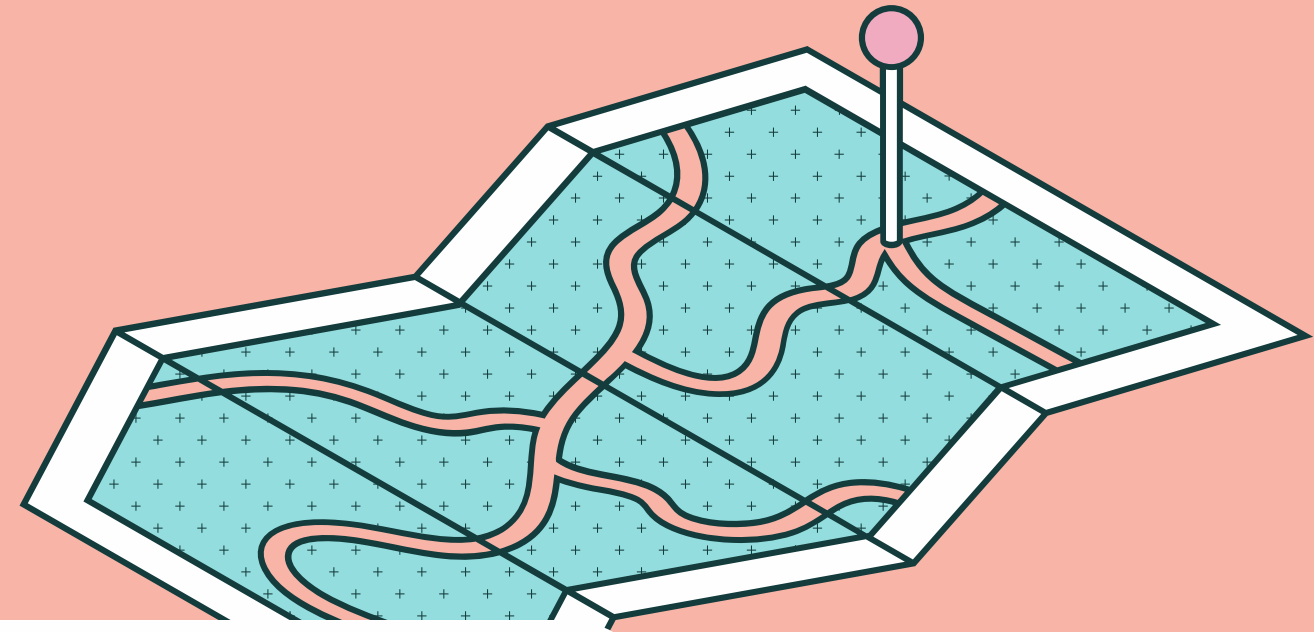
Big Data

Dremio s'intègre à plusieurs branches du Big Data en répondant aux 5V



Volume	Dremio gère de grandes quantités de données sans déplacement
Vitesse	Traitement accéléré avec Data Reflections pour des résultats en temps réel.
Variété	Dremio connecte facilement toutes les sources de données, sans transformation complexe.
Valeur	Démocratisation des données : tout le monde peut exploiter les données rapidement.
Véracité	Accès direct aux données à la source, sans risque d'erreur.

HISTORIQUE



- **1980 Bases de données relationnelles**
 - **1990 Data Warehouse**, stockage de données structurées
 - **2000 Hadoop et Spark** permet le traitement de données non structurées
 - **2010 Data Lake**, stockage massif de données brutes
 - **2020 Data Lakehouse**
-
- **Dremio 2017** : Pilote l'évolution vers le Data Lakehouse avec :
 - Dremio Arctic : Gouvernance et performance optimisées pour les données hybrides.
 - Dremio Sonar : Accès SQL performant sur les données non structurées.

CLIENTS



CAPGEMINI



SOPRASTERIA



MICHELIN



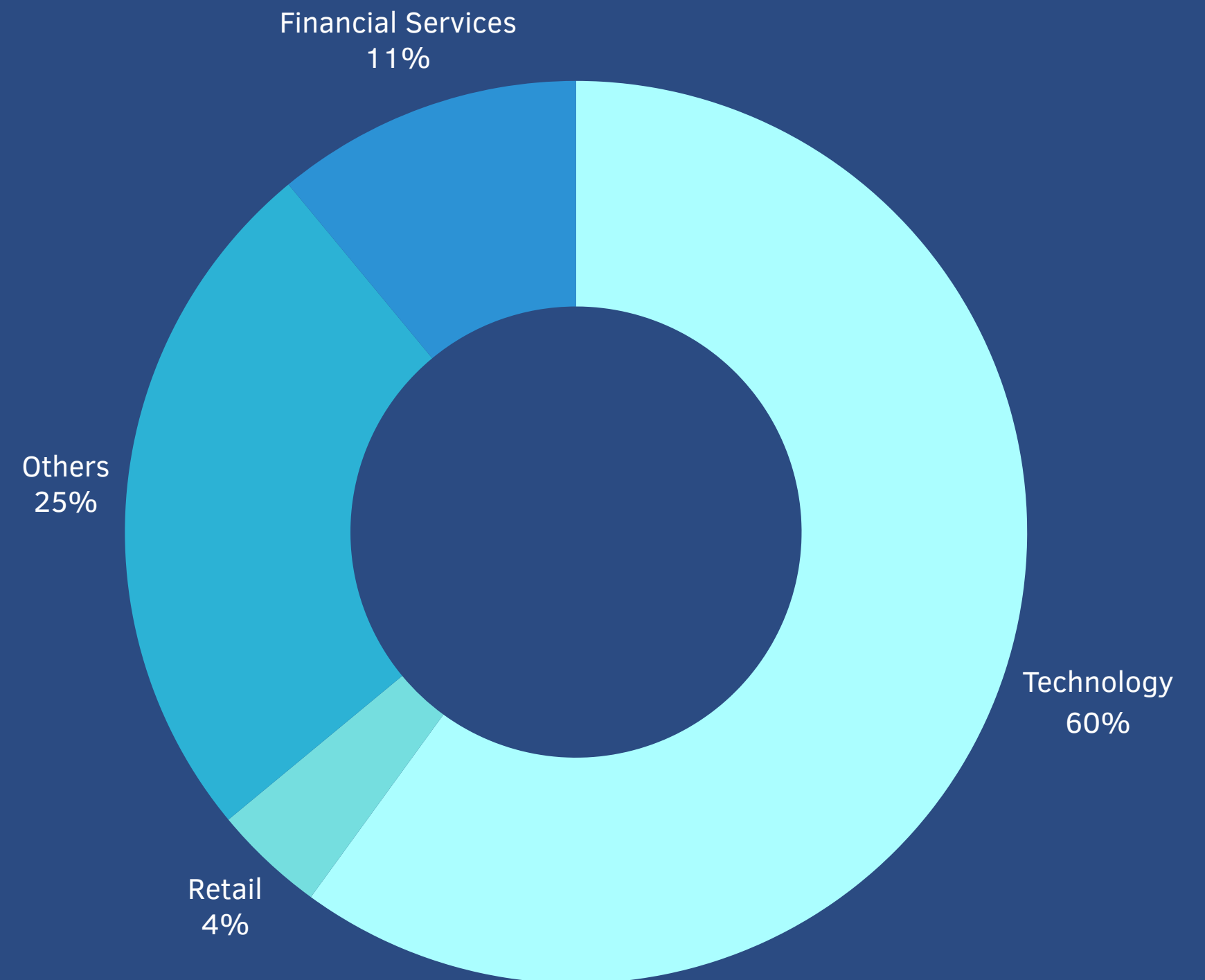
RWE



AMAZON

Répartition par secteur d'activité

“57% of Dremio customers are in United States, 5% are in Brazil and 5% are in United Kingdom.”



source: <https://enlyft.com/tech/products/dremio>

Applications et Utilisations

Analyse et Rapport de données

Représente 50% des cas d'usages

Intégration des données et ETL

30% des cas d'usages

ETL (*Extract-transform-load*) = Synchronisations massives d'information d'une source de données vers une autre.

Business Intelligence (BI)

20% des cas d'usages

Processus et outils pour analyser les données de gestion, les visualiser et aider à prendre des décisions.

Cas d'utilisation

Scenario : Un e-commerce souhaite analyser les performances des produits pendant les soldes du Black Friday.

PROBLEMATIQUE

Stockage des données dans un Datalake.

- Augmentation exponentielle des données
- Difficultés à les gérer pour l'analyse

SOLUTION AVEC DREMIO

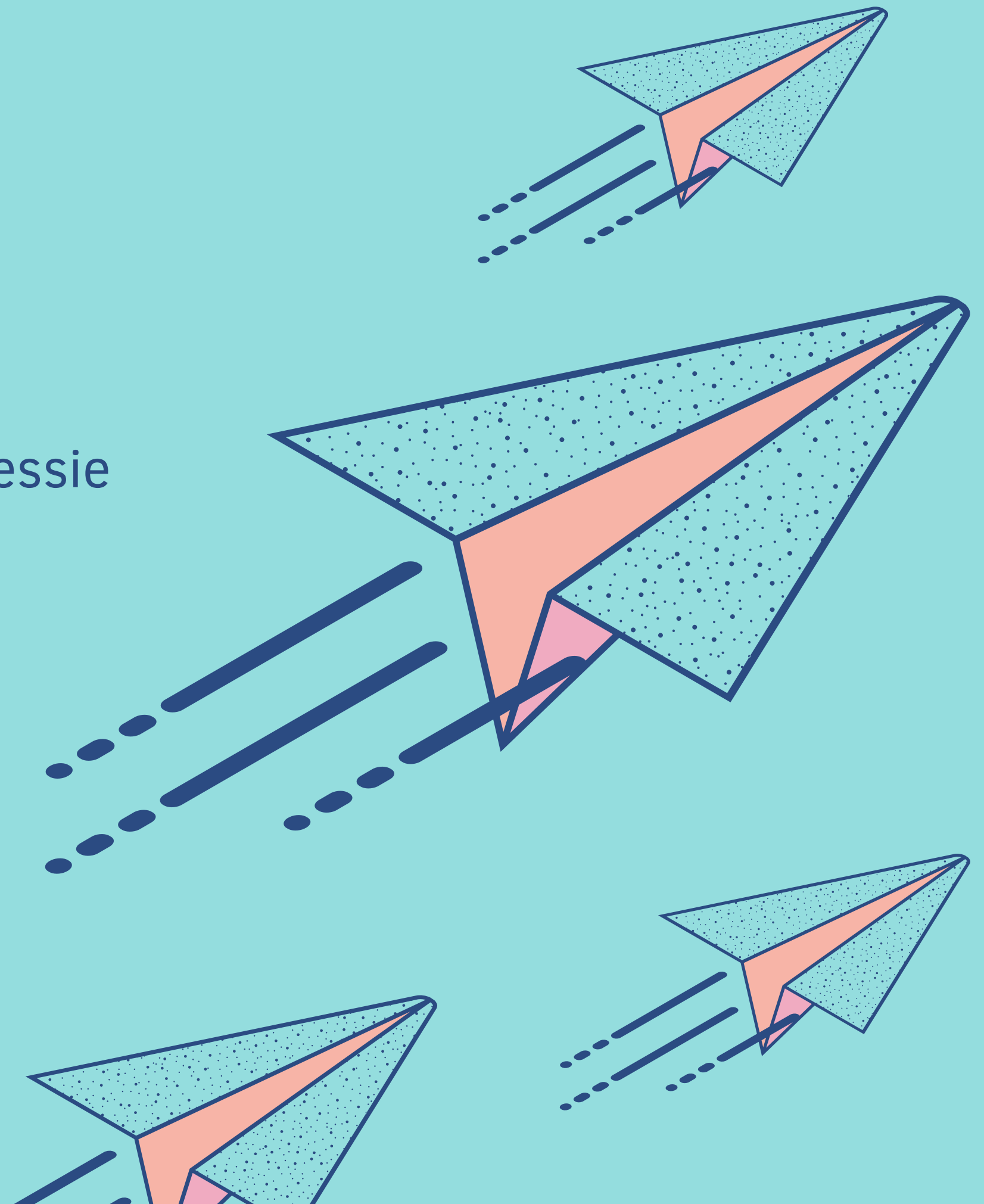
- Récupérer les données en temps réel
- Connaître les pages qui génèrent le plus d'engagement (clickstream logs)
- Interpréter les résultats (inventaire)
+ Interface SQL unifiée pour les analystes

Tip: Use links to go to a different page inside your presentation.

How: Highlight text, click on the link symbol on the toolbar, and select the page in your presentation you want to connect.

Démonstration

Exemple d'utilisation avec MongoDB et Nessie



Merci pour votre attention

Avez-vous des questions ?