

Machine Learning
Data Catalog

WeeFin Finance Forward

WEEFIN

27 rue des Poissonniers
92200 Neuilly-Sur-Seine
FRANCE
<https://weefin.co>
@weefin_

Machine Learning Data Catalog

Accélérateur de Transformation Data Centric

Les Machine Learning Data Catalogs sont incontournables pour gérer le patrimoine de données mis en libre-service et se conformer aux nouvelles réglementations. Or, seules 28% (*) des institutions financières sont en train de développer une culture de la donnée. Nos clients, en s'appuyant sur notre récent panorama et sur notre démarche agile s'inspirant du Design Thinking, ciblent les éléments ayant un impact concret pour leur organisation favorisant l'adoption d'une culture Data-Driven.

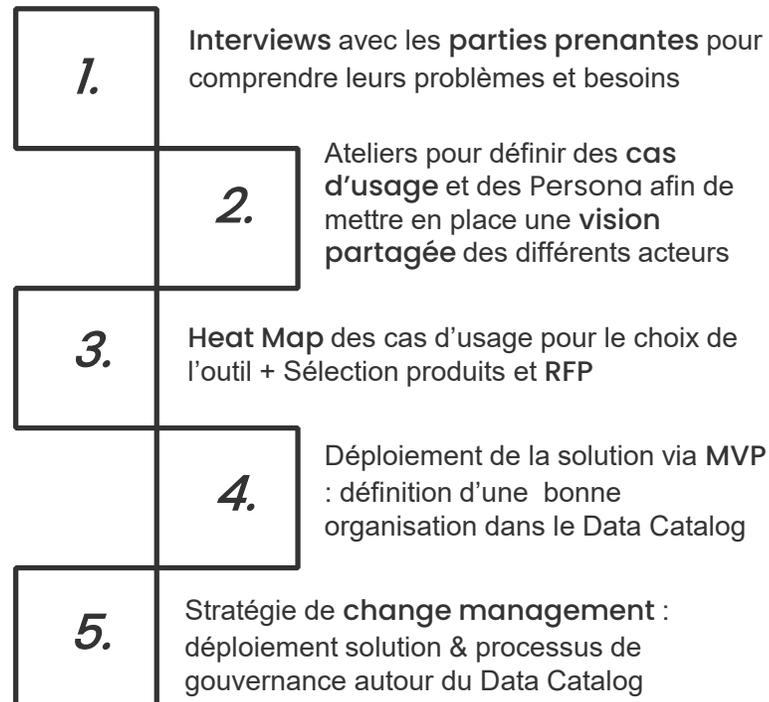
*(NVP, Big Data and AI Executive Survey 2019)

Contexte

- La **qualité** et la **sécurité** de la donnée sont des enjeux majeurs pour les **institutions financières**. Le régulateur a défini les best practices de la gouvernance avec la directive BCBS 239. Les institutions ont donc créé des **Data Offices** pour redéfinir leur processus de gestion de la donnée.
 - Jusqu'à présent, deux familles d'outils de gouvernance de la donnée coexistaient : l'un orienté processus métier long à mettre en place, l'autre lié à l'IT, dépendant des solutions ETL (**), impactant pour l'architecture.
- La nouvelle génération de Data Catalog permet une **collaboration entre les métiers et l'IT** à travers une expérience utilisateur simple basée sur un moteur de recherche. Grâce à l'intégration du **Machine Learning**, certains outils proposent un niveau d'**automatisation** encore plus poussé.

**ETL: Extract, Transform & Load

Méthode



Accompagnement

- Panorama des solutions** Data Catalogs Innovants (étude de marché, rapports, préconisations) dont Alation, Unifi, Collibra, Zeenea, ...
- Trajectoire de **mise en œuvre** d'un Data Catalog à partir de cas d'usage
- Support pour l'adoption** du projet auprès de la DSI
- Réalisation de **MVP Data Catalog** (conditions de succès, grilles d'évaluation)

Activités

- Etude d'opportunité
- Changement Management
- Etude de marché et sélection de vendeur, RFP
- Cadrage, Pilotage

Problèmes

- Mauvaise gestion et difficulté d'accès à la donnée ont pour conséquence une perte de temps et d'efficacité.
- Les projets Data peinent à prouver de la valeur : manque de collaboration entre les équipes et outils mal adaptés.

Causes

- Obstacles à l'entreprise Data - Driven, les organisations gardent **une culture et méthodologie anciennes de la donnée**.
- Méconnaissance de l'apport des ML Data Catalog comme accélérateur de projets Data.

Solution

Choisir un **Data Catalog** qui répond à de réels cas d'usage pour le métier et l'IT. En s'adaptant aux besoins spécifiques de l'organisation et à son modèle de données, il prouvera son utilité, ce qui facilitera son adoption.

Bénéfices

Bâtir une **source de référence** qui facilite l'accès à la donnée
Offrir une plus grande **autonomie** des équipes et donc plus de temps actif sur les projets
Lier les domaines métiers, stockage et flux d'informations
Permettre de mieux comprendre et faire confiance pour prendre les bonnes décisions.