TD2 — Les briques d'une architecture de données

Reprise de la séquence 1

> **Rappel — les 5 questions de la direction**
>
> * Quel est le chiffre d'affaires réel de MODA & CO, tous canaux confondus ?
> * Quels sont nos clients les plus rentables ?
> * Quels produits sont en rupture — ou risquent de l'être ?
> * Quelle campagne marketing est réellement efficace ?
> * Comment proposer des recommandations personnalisées ?
>
> Nous avons listé les problèmes. Voyons maintenant comment une architecture de données peut y répondre.

1. Définition à construire

Avant de consulter une définition de référence, prenez quelques minutes pour construire votre propre définition.

En **deux phrases**, expliquez ce que vous entendez par « architecture de données » dans une organisation comme MODA & CO. Votre réponse doit préciser, si possible :

- d'où viennent les données ;
- ce qui leur arrive avant d'être utilisées ;
- où elles sont conservées ;
- par qui et pour quels usages elles sont exploitées.

Votre définition :

```
. . . . . .
. . . . . .
. . . . . .
. . . . . .
```

Définition de référence — à consulter après votre essai

> Une architecture de données décrit la manière dont les données sont collectées, transportées, transformées, stockées, sécurisées, gouvernées et exploitées dans une organisation.

Cette définition ne désigne donc pas un outil unique. Elle décrit un ensemble cohérent de composants, de flux, de responsabilités et de règles.

3. Exercice de classement : les briques de l'architecture

Situation

MODA & CO dispose des systèmes suivants. Ils sont volontairement mélangés :

- fichiers Excel utilisés par les commerciaux ;
- application mobile s'appuyant sur Firebase ;

- ERP ;
- service client Zendesk ;
- site e-commerce Magento ;
- plateforme publicitaire Google Ads / Meta Ads ;
- WMS de gestion des stocks ;
- caisses des 50 magasins physiques ;
- CRM Salesforce.

Consigne

Classez chaque élément dans le tableau ci-dessous. Vous pouvez inscrire un même élément dans plusieurs colonnes si vous pouvez le justifier. Pour chaque choix, préparez une justification courte : « quelles données ? pour quel rôle ? »

Sources	Flux	Stockage	Traitements	Usages	Fonctions transverses

Questions de guidage :

1. Une application est-elle forcément une source, ou peut-elle aussi être un usage ?
2. Un système opérationnel peut-il jouer un rôle de stockage ?
3. Où placer un outil qui expose une API, mais qui ne fait pas nécessairement de transformation ?
4. Les fonctions transverses sont-elles un logiciel précis ou des capacités qui s'appliquent à tous les systèmes ?

Note pédagogique pour l'animateur

Certains systèmes peuvent appartenir à plusieurs briques à la fois (par exemple : **CRM = à la fois source ET usage**). C'est volontairement piégeux : l'objectif est de créer du débat et de faire comprendre qu'un classement dépend du **rôle considéré**, du périmètre et du niveau de description.

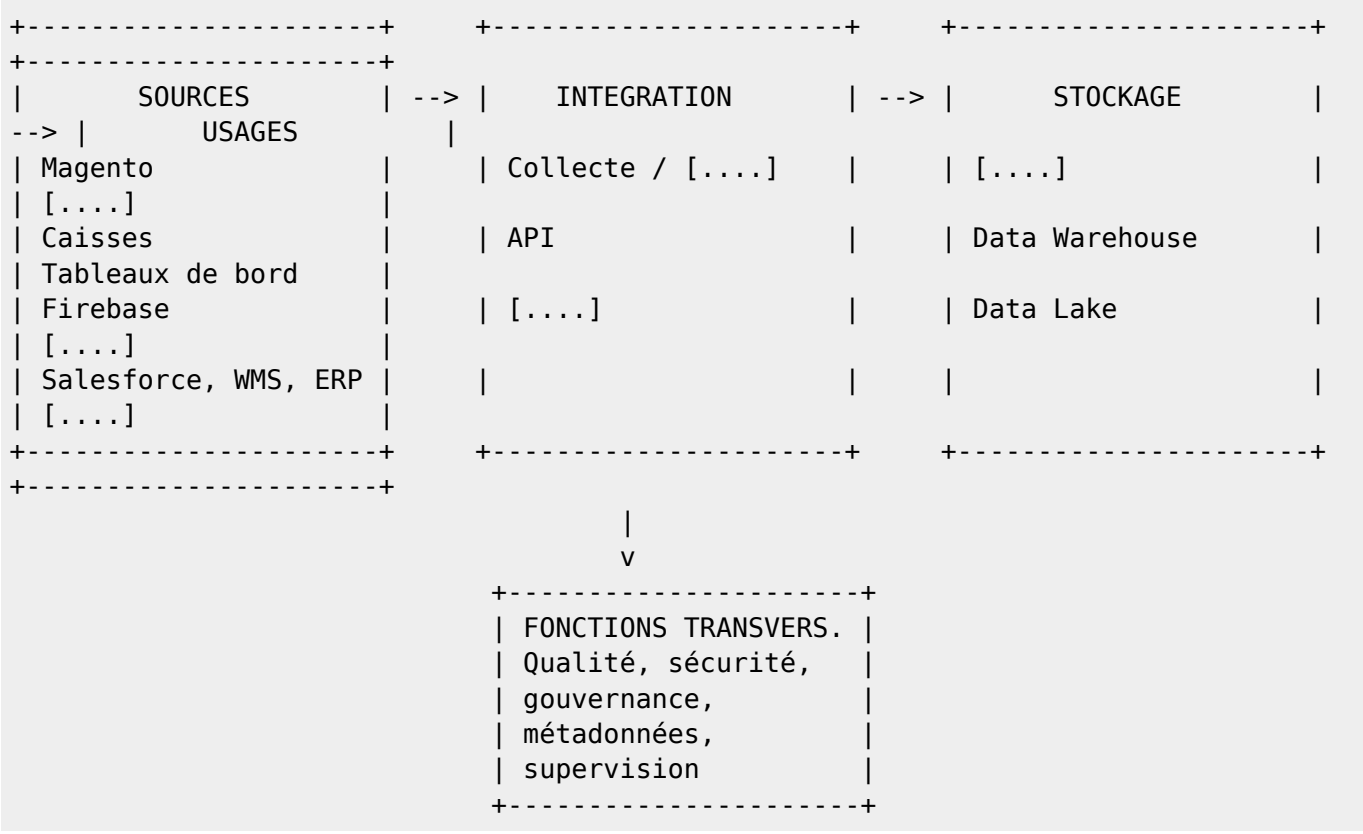
Pièges à surveiller :

- **CRM Salesforce :** source de données clients et commerciales, mais aussi usage pour les équipes de vente et éventuellement stockage opérationnel. Ne pas imposer une seule case.
- **WMS / ERP / caisses :** ce sont des systèmes opérationnels qui stockent des données et les produisent ; ils ne sont pas, à eux seuls, les « traitements » analytiques.
- **Google Ads / Meta Ads :** source de données de campagnes et usage de pilotage marketing ; les flux de retour (coûts, clics, conversions) peuvent être entrants ou sortants selon le point de vue.
- **Excel commerciaux :** source fragile, mais aussi support d'usage ; qualité, validation, versioning et sécurité doivent être discutés plutôt que classés comme un simple outil.
- **Fonctions transverses :** qualité, sécurité et gouvernance ne se rangent pas dans un seul système ; elles concernent l'ensemble de l'architecture.

Restitution attendue : chaque groupe choisit deux classements discutables et prépare un argument pour défendre son choix.

4. Schéma à trous

Complétez le schéma en indiquant des composants, des exemples d'outils ou des types de traitements. Les mots à retrouver ne sont pas nécessairement les seuls éléments possibles.



Questions de débrief :

- Où se situent les transformations : dans l'intégration, avant le stockage, après le stockage, ou à plusieurs endroits ?
- Le Data Lake et le Data Warehouse ont-ils exactement le même rôle ?
- Un tableau de bord est-il un stockage ou un usage ?
- Où placer l'IA / Data Science ?

---

5. Mini-défi : choisir les flux de données

Les données ne circulent pas toutes de la même manière. Associez chaque situation à un mode de flux plausible, puis justifiez votre choix par le besoin métier, le délai acceptable et le coût de mise en œuvre.

Rappel des modes de flux

Mode	Principe	Exemple générique
Batch	Données accumulées puis transférées par lots à intervalles planifiés.	Export quotidien des ventes vers la plateforme analytique.
Temps réel	Donnée transmise et disponible avec un délai très faible.	Mise à jour immédiate d'un stock critique.
Micro-batch	Petits lots fréquents, compromis entre batch et temps réel.	Consolidation toutes les quelques minutes.
Événementiel	Un événement déclenche la publication ou le traitement d'un message.	« commande créée », « paiement accepté », « produit scanné ».
API	Échange de données via une interface applicative exposée par un système.	Interrogation du CRM ou remontée d'une conversion publicitaire.

À vous de proposer

Pour chaque ligne, indiquez un ou plusieurs modes plausibles et expliquez pourquoi. Il n'y a pas toujours une seule bonne réponse : précisez l'hypothèse retenue.

Système ou situation MODA & CO	Mode(s) de flux proposé(s)	Justification
Caisses des 50 magasins		
Site e-commerce Magento		
Application mobile / Firebase		
CRM Salesforce		
WMS et ERP		
Google Ads / Meta Ads		
Service client Zendesk		
Fichiers Excel commerciaux		

Relance : le flux adapté à une recommandation personnalisée immédiate est-il nécessairement le même que celui adapté à un reporting financier mensuel ?

6. Fonctions transverses : rendre l'architecture fiable

Une architecture ne se résume pas à des tuyaux et des bases de données. Les fonctions transverses accompagnent toutes les briques :

- **Qualité :** exactitude, complétude, cohérence, fraîcheur, unicité ;
- **Sécurité :** contrôle des accès, chiffrement, sauvegardes, traçabilité ;
- **Gouvernance :** responsabilités, règles, définitions communes, conformité ;
- **Métadonnées :** description des données, origine, sens, propriétaire, transformations ;
- **Supervision :** suivi des flux, alertes, journaux, reprise sur incident, indicateurs de service.

> **Une architecture sans gouvernance peut fonctionner techniquement tout en produisant des résultats inutilisables ou risqués.**

Mini-défi appliqué à MODA & CO

Que se passe-t-il si personne ne valide les remises négociées à l'oral par les commerciaux dans les fichiers Excel ?

En groupe, identifiez :

```
1. au moins trois risques sur les données ou les décisions ;
2. une règle de gouvernance à mettre en place ;
3. un contrôle de qualité mesurable ;
4. une mesure de sécurité ou de traçabilité ;
5. le rôle qui devrait être responsable de la validation.
```

7. Synthèse : relier architecture et décisions

Reprenons les cinq questions de la direction. Pour **CHAQUE** question, identifiez les briques nécessaires pour construire une réponse fiable. Ne vous limitez pas au stockage : indiquez les sources, les flux, les traitements, les usages et les fonctions transverses mobilisés.

Question de la direction	Sources	Flux	Stockage	Traitements	Usages	Fonctions transverses
Chiffre d'affaires réel, tous canaux						
Clients les plus rentables						
Produits en rupture ou à risque						
Campagne marketing réellement efficace						
Recommandations personnalisées						

Question de conclusion : quelle question vous paraît la plus difficile à traiter, et quelle brique est la plus critique dans votre proposition ? Justifiez en une minute.

À retenir

Une architecture de données est une **organisation cohérente** des sources, des flux, du stockage, des traitements et des usages, soutenue par la qualité, la sécurité, la gouvernance, les métadonnées et la supervision. Elle relie les systèmes de MODA & CO aux décisions de la direction ; elle ne se réduit donc ni à un Data Warehouse, ni à un outil particulier.

Minutage indicatif : 8 min *— 5 min de remplissage du tableau, 3 min de synthèse orale.*

From:
http://slamwiki2.kobject.net/ - **SlamWiki 2.1**

Permanent link:
http://slamwiki2.kobject.net/eadl/bloc5/fm1/td2?rev=1788086941

Last update: **2026/08/30 12:49**