



Master Informatique - MIAGE

Urbanisation des Systèmes d'Information Architecture d'Entreprise

05 – Architecture des données

Philippe Declercq

2019-2020

Plan du chapitre

1	Définitions & enjeux
2	Les référentiels
3	La gestion des données de référence – le MDM

Urbanisation des Systèmes d'Information et Architecture d'Entreprise – Architecture des données 2 Philippe Declercq

Références

1	« Les référentiels du système d'information » de Joël Bizingre, Joseph Paumier et Pascal Rivière, éditions Dunod
2	« MDM » de Franc Régnier-Pécastaing, Michel Gabassi et Jacques Finet, éditions Dunod
3	Gartner's Seven Building Blocks of MDM : The Foundation for Successful MDM

Urbanisation des Systèmes d'Information et Architecture d'Entreprise – Architecture des données 3 Philippe Declercq



Définition et enjeux

1

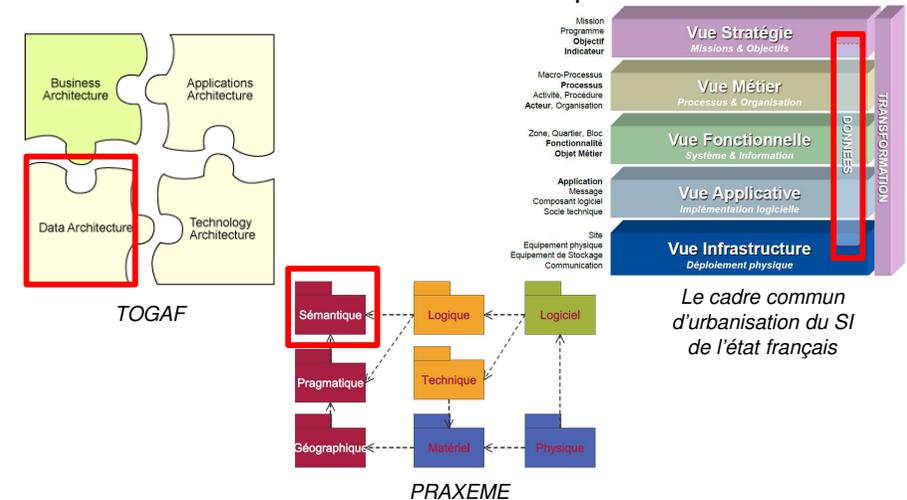
Urbanisation des Systèmes d'Information et Architecture d'Entreprise – Architecture des données 4 Philippe Declercq

Architecture des données

- ◆ L'urbanisme s'intéresse en premier lieu aux processus
 - ... mais cela n'est possible que si un travail de fond est réalisé sur la parties données,
 - Les processus manipulent des « dossiers » qui sont des agrégats d'objets métier.
- ◆ Il ne pourra y avoir de SI cohérent et agile que si :
 - Ses données sont cohérentes dans l'ensemble du SI,
 - Ses données sont correctement localisées au sein des différents blocs fonctionnels.

Architecture des données

- ◆ L'Architecture des données est identifiée dans les cadres d'urbanisation et d'architecture d'entreprise :



Architecture des données

- ◆ Au même titre que les urbanistes et architectes ont en commun un référentiel des processus métier, un référentiel des fonctions du SI, un référentiel des applications, ... il faut constituer un référentiel de données.
- ◆ Il doit être composé d'éléments structurels (c'est-à-dire un modèle de données) et d'un volet sémantique.

Architecture des données

- ◆ Quelques préoccupations des urbanistes et architectes :
 - **Unicité** : une entité du monde réel n'est-elle représentée que par un seul objet métier dans le SI de l'entreprise ?
 - **Exactitude** : les valeurs des différents attributs sont-ils exacts ?
 - **Complétude** : Toutes les données sont-elles présentes ? Est-ce que tous les attributs de mes données sont renseignés ?
 - **Cohérence / intégrité** : les données sont-elles cohérentes d'une application à l'autre ? Les liens entre elles sont-ils corrects ?
 - **Fraîcheur** : les informations disponibles sont-elles suffisamment récentes vis-à-vis des usages ?
 - **Disponibilité** : les données sont-elles accessibles à ceux qui en ont besoin ?
 - **Sécurité** : les données sont-elles protégées d'accès ou de modifications malveillantes ou simplement non-autorisées ?

Architecture des données

- ◆ Le travail sur la donnée doit porter sur l'ensemble de son cycle de vie :
 - ▶ Lors de l'acquisition, notamment dans le cas des saisies manuelles
 - Principe du garbage in – garbage out (la qualité des résultats dépend de la qualité des entrées)
 - Fiabilisation des données
 - ▶ Lors des transformations et agrégations
 - Problématiques de synchronisation post-traitement
 - Uniformité des traitements, notamment de fiabilisation et/ou dédoublonnage
 - ▶ Lors de l'analyse et la présentation des résultats.

Architecture des données

A quoi la constitution et le maintien du modèle d'entreprise répond ?

Unicité
Disponibilité
Fraîcheur
Complétude
Sécurité
Cohérence

Architecture des données

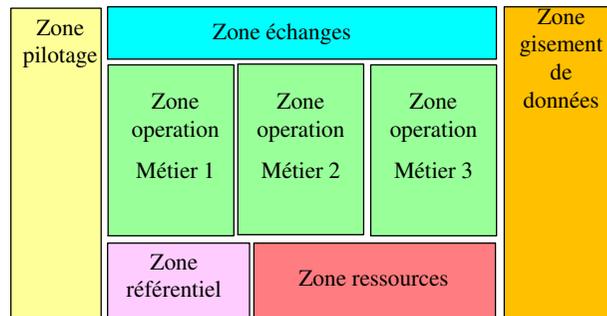
- ◆ Le modèle des données du SI doit être construit – lorsqu'il existe - à partir du dictionnaire des entités métier tenu à jour sur l'architecture métier.
- ◆ Ce modèle est généralement organisé de manière arborescente :
 - Il peut être vu comme un agrégat de modèles plus locaux,
 - Les premiers niveaux de l'arborescence peuvent être le plan d'occupation des sols (objets métier) ou les processus (dossiers).
- ◆ Cette modélisation ne doit pas entrer dans des détails d'ordre technique ou de conception fine.
- ◆ Le langage de représentation peut être UML, Merise (MCD), voire même les langages propriétaires.

Architecture des données

- ◆ Lors de la constitution d'un dossier d'architecture, les exigences fonctionnelles et non fonctionnelles doivent être capturées :
 - ▶ Confidentialité
 - ▶ Auditabilité des accès
 - ▶ Volumétries statique et dynamique
 - ▶ Durées de conservation
- ◆ Ces travaux sont nécessaires au niveau fonctionnel et applicatif. La prise en compte dans la conception doit être vérifiée.

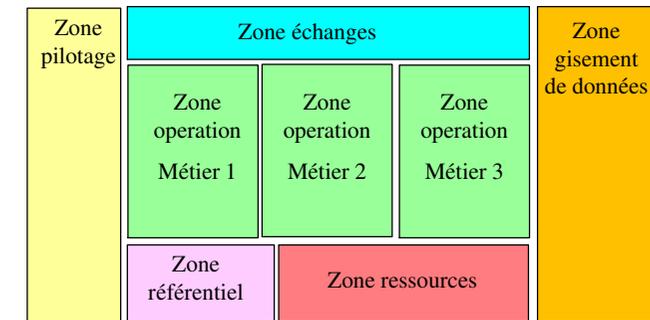
Architecture des données

- ◆ Au-delà de posséder un modèle d'entreprise pour assurer la cohérence de l'information dans l'ensemble du SI, il est nécessaire de localiser correctement les données dans le SI.
- ◆ Où se situent les données dans la structuration des zones du SI de Longépé ?



Architecture des données

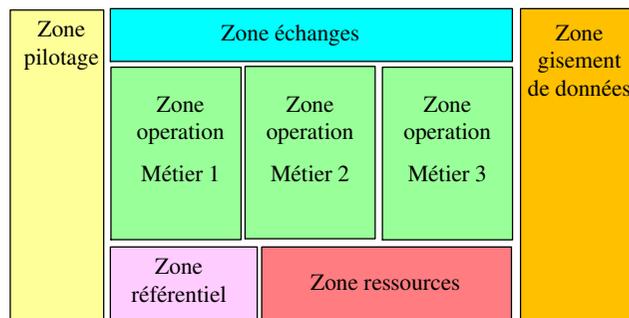
- ◆ Éléments de réponse : partout !



- ◆ Dans le cadre d'un projet, il est nécessaire de positionner chaque donnée dans les zones du POS.
- ◆ Ce travail est similaire à celui opéré sur les fonctions. Il se fait conjointement.

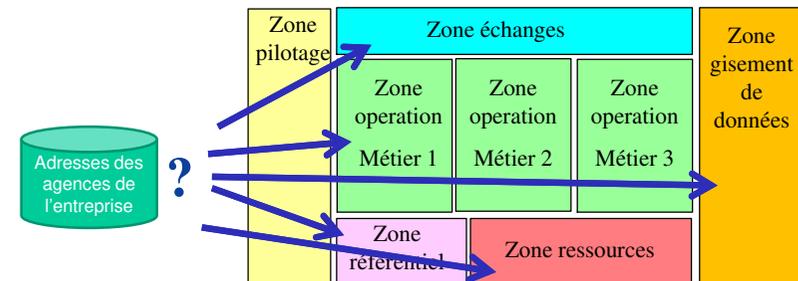
Architecture des données

- ◆ Comment décider qu'une donnée se situe dans la zone référentiel ? Dans la zone gisement de données ? Dans une zone métier ?
- ◆ Quelles en sont les conséquences ?



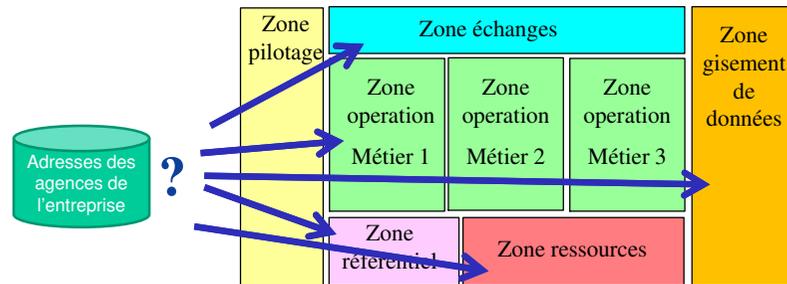
Architecture des données

- ◆ Exemple : un grande entreprise organisée sur le territoire en agences.
- ◆ Le système d'envoi des courriers aux clients a besoin de préciser le nom et l'adresse de l'agence du client sur chaque courrier.
- ◆ Ces données sont elles spécifiques aux fonctions d'édition de courriers ? Où localiser ces données ?



Architecture des données

- ◆ Pour localiser une donnée dans le SI, se poser les questions :
 - Quel est le niveau de partage de la donnée au sein des différentes zones du POS ?
 - Quel est la fréquence d'évolution de la donnée ?
 - Quels sont les processus de mise à jour de la donnée ?

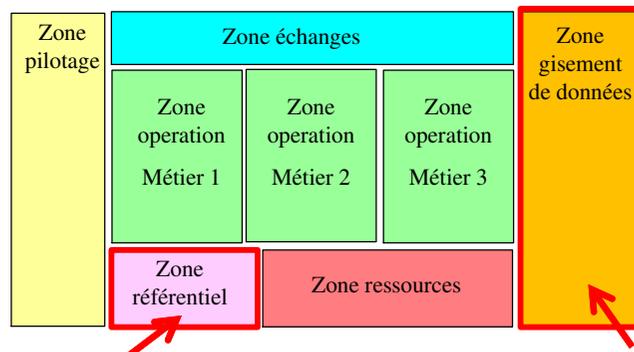


Architecture des données

- ◆ On distingue, selon le niveau de partage :
 - les données « privées » : elles n'ont pas vocation à être largement partagées au sein du SI - elles restent locales à une zone du POS ou à une application informatique
 - les données « transverses » : plusieurs zones du POS ont besoin de ces données. On peut distinguer deux types :
 - Données « stables », fréquence d'évolution faible.
Exemples : Clients, Produits
 - Données « à cycle de vie rapide », se renouvelant souvent.
Exemples : Commandes, livraisons

Architecture des données

- ◆ Les zones « Référentiel » et « Gisement de données » hébergent des données « transverses » :

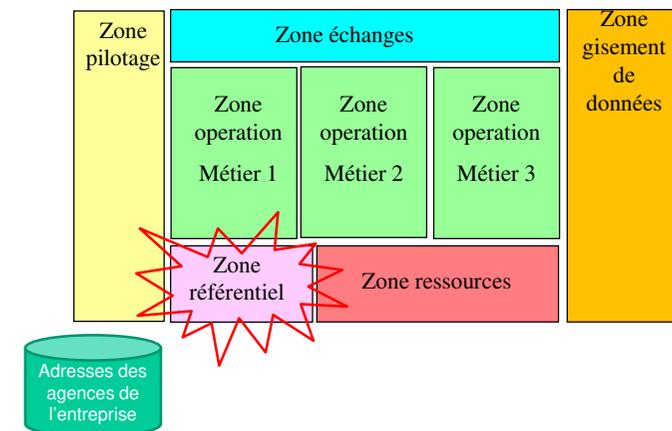


Les Référentiels sont des applications centralisatrices des informations :
 - fortement partageables sur le SI
 - stables - cycle de vie lent

Les Gisements sont des applications centralisatrices des informations
 - fortement partageables sur le SI
 - à cycle de vie rapide

Architecture des données

- ◆ Retour sur l'exemple : l'analyse des urbanistes conduit à localiser les données concernant les agences de l'entreprise dans la zone ...





Les référentiels

2

Les référentiels

- ◆ La mise en place de référentiels de données est un élément primordial et prioritaire dans la mise en œuvre de l'urbanisation.
- ◆ C'est souvent par des projets de type « Référentiel » que les entreprises peuvent urbaniser leur SI et obtenir des gains rapides.

Les référentiels

- ◆ Définitions : référentiel de données :
 - Ensemble structuré d'informations, utilisé pour l'exécution d'un logiciel, et constituant un cadre commun à plusieurs applications.
 - Ensemble cohérent de données ayant une définition sémantique commune et répondant au besoin de langage commun entre plusieurs acteurs appartenant à des entités organisationnelles différentes ou à une même entité.
 - Données qui, relativement stables et hautement partagées d'un processus à l'autre, composent les informations fondamentales autour desquelles l'entreprise structure son activité.

Les référentiels

- ◆ Le référentiel vient aussi en réponse à des problématiques :
 - Conflit d'identité. Le même objet possède des identités différentes selon les sources de données.
 - Conflit de schémas. Le même concept est modélisé de manière différente.
 - Conflit de valeur. Le même objet a des valeurs différentes selon les sources.
 - Conflit sémantique. Le même terme est interprété de manières différentes (ex : un « compte » est un compte courant bancaire dans une application et un identifiant utilisateur permettant de gérer la sécurité des accès pour une autre).
 - Plus grande complexité du SI à assurer les mises à jour. Dans une configuration de redondance de valeur, nécessité de mettre à jour toutes les bases ou fichiers contenant cette donnée via des passerelles compliquées entre les différentes sources de données.

Caractéristiques des référentiels

- ◆ Les cinq caractéristiques majeures attendues d'un référentiel sont :
 - **Centralité** – il est reconnu comme référence sur son périmètre
 - **Stabilité** – les données changent peu dans le temps
 - **Qualité** – les processus sous-jacents assurent des données fiables
 - **Unité de sens** – les données sont homogènes sur le plan sémantique
 - **Interopérabilité** – les données sont accessibles (techniquement) du reste du système d'information.
- ◆ Il existe deux types de référentiels :
 - Les **répertoires**
 - Les **nomenclatures**.

- ◆ Un **répertoire** est un ensemble d'occurrences de la même entité. Qualifient l'entité :
 - Un identifiant, de préférence non signifiant,
 - Les traits d'identité,
 - Les données de typage,
 - Les métadonnées,
 - Les données historique,
 - Les autres traits de caractérisation.
- ◆ Exemple de répertoires :
 - Liste des clients,
 - Liste des fournisseurs,
 - Répertoire des entreprises SIRENE (INSEE).

- ◆ Les **nomenclatures** sont des ensembles de catégories, organisées de manière arborescentes. Qualifient une catégorie :
 - Un code,
 - Un positionnement dans l'arborescence,
 - Un ou plusieurs libellés,
 - Des notes et/ou contraintes d'utilisation.
- ◆ Exemples de nomenclatures :
 - Listes des actes médicaux (CCAM),
 - Liste des professions (Code ROM),
 - Plan comptable.

Les référentiels

Centralité

- ◆ Un référentiel est fait pour être accessible par de nombreux clients
- ◆ Il est unique
 - Lorsque plusieurs ensembles applicatifs portent les mêmes données (c'est possible), il faut s'assurer qu'un des applicatifs est référent, et que les autres applicatifs s'en alimentent.
- ◆ Il est légitime. Il a été reconnu comme référent au travers d'une instance officielle

- ◆ *L'assurance maladie dispose d'un référentiel des personnes en tant qu'assurés, et un référentiel des personnes en tant que professionnel de santé. Respecte-t-elle ce principe ?*

Les référentiels

Stabilité

- ◆ La notion de stabilité – même si cela semble contradictoire avec la définition donnée plus haut – fait plus référence à une indépendance vis-à-vis des processus de production qu'à une notion temporelle.
- ◆ Les attributs des objets référencés – les adresses des individus dans un répertoire client, par exemple – peuvent bouger tous les jours. Dès lors qu'il reste fondamentalement le même, on peut parler de stabilité ;
- ◆ Il ne faut pas confondre « stable » et « fixe ».

- ◆ *Nous avons pris pour exemple de stabilité les adresses dans un référentiel individu. Dans quel cas peut-on considérer ces mêmes adresses comme n'étant pas stables ?*

Les référentiels

Qualité

- ◆ Les données d'un référentiel doivent être fiables. Les clients doivent pouvoir leur faire confiance.
 - ◆ On n'attend en revanche pas que les données soient sans erreurs. Le taux d'erreur doit cependant être faible, et le processus de correction défini et maîtrisé.

- ◆ *Sur le terrain, certains concepteurs veulent s'alimenter sur une application A plutôt qu'un référentiel R arguant que le référentiel R s'alimentant sur A, les données de A sont plus fiables. Ont-ils raison ?*

Les référentiels

Unité de sens

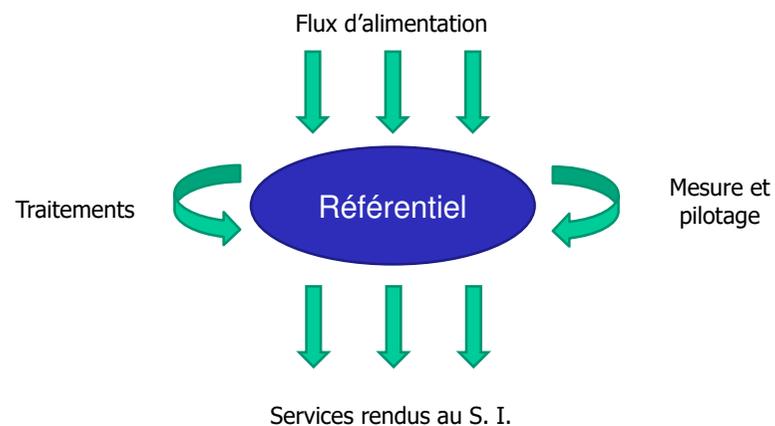
- ◆ Les attributs du référentiel doivent toujours avoir le même sens.
 - Il ne faut pas mélanger des données différentes sous prétexte qu'elle correspondent à la même « chose » du monde réel. Par exemple, mélanger « clients » et « fournisseurs » sous prétexte qu'on peut être les deux à la fois ;
 - En corolaire, il ne faut pas profiter de la présence d'un objet métier en référentiel pour lui ajouter des attributs propre à un besoin particulier ;
 - C'est un gage de lisibilité des données.

- ◆ *Une entreprise choisit d'enrichir un référentiel avec un nombre important d'attributs, confiante dans le fait que ceux-ci gardent cependant une homogénéité sémantique en son sein. Quel est le problème ?*

Interopérabilité

- ◆ Le référentiel doit être une vraie « plaque tournante », ce qui suppose des exigences d'interopérabilité technique, sémantique et fonctionnelle
 - Le référentiel est donc aussi au centre d'un système informatique d'échange ;
 - Un référentiel doit d'emblée être considéré comme un fournisseur de services, avec tout ce que cela implique en matière de performance, d'accessibilité, etc.
- ◆ *Quelle pourrait-être la conséquence pratique d'un référentiel qui ne serait pas d'emblée construit comme étant interopérable ?*

Considérations architecturales



- ◆ Les **nomenclatures** posent des problèmes spécifiques. Elles servent à « classer le monde » - à opérer une segmentation
 - ▶ La légitimité des nomenclatures est d'autant plus forte que ses données sont à la fois plus stables et plus structurantes ;
 - ▶ Pourtant, elles évoluent. Dans ce cas, que faire pour les informations passées ? Faut-il opérer un reclassement ?
 - ▶ La problématique d'historisation est donc d'autant plus importante qu'elle est nécessaire pour reconstituer les vues passées des données qui y font référence ;
 - ▶ Il peut arriver que deux nomenclatures différentes soient en conflit sur une donnée particulière. Le conflit est d'autant plus difficile à résoudre que les nomenclatures ont souvent une légitimité forte.

Les référentiels

- ◆ Les **répertoires**, quant à eux, ont pour problématique spécifique la problématique de couverture
 - ▶ Sur-couverture, lorsque le référentiel contient des doublons, les unités hors champs ou les faux actifs ;
 - ▶ Sous-couverture, lorsqu'il manque tout bonnement des occurrences.
- ◆ Il est généralement possible d'évaluer quantitativement ces situations, et ainsi prendre les mesures de correction appropriées

- ◆ La question du mode de mise à jour se pose
 - ▶ Alimentation pro-active à partir de sources référencées ?
 - ▶ Événementielle ?

Les référentiels

- ◆ L'ensemble des dimensions de la donnée doivent être prises en comptes :
 - ◆ Le concept porté ;
 - ◆ Le domaine de validité ;
 - ◆ Les méta-données ;
 - ◆ Leur cycle de vie.

- ◆ Par ailleurs, dans la constitution du référentiel, vous devez vous poser ces questions :
 - ◆ Quels sont les objets métiers que devra gérer le système ?
 - ◆ Quels sont les utilisateurs ?
 - ◆ Quels sont les processus qui devront être portés par le système ?
 - ◆ Quelles sont les exigences non fonctionnelles ?
 - ◆ Quelles sont les contraintes sur les contenus ?

Les référentiels



La méta-donnée est une donnée... sur la donnée.

Pouvez-vous en citer des exemples ?

Les référentiels

- ◆ Exemples de méta-données :
 - ◆ L'état (sous-entendu, la position dans le cycle de vie)
 - ◆ La date d'introduction dans le référentiel
 - ◆ La date de dernière mise à jour
 - ◆ La date ou la durée de validité
 - ◆ L'utilisateur ayant mis à jour la donnée

Les référentiels

- ◆ Les freins et difficultés :
 - ▶ Unité de sens difficile à capter
 - ▶ Périmètre parfois ambigu
 - ▶ Maintien de la qualité difficile
 - ▶ Responsabilités mal définies
 - ▶ Référentiel vu comme une contrainte

- ◆ Les principaux freins ne sont pas techniques, mais sémantiques et organisationnels.

- ◆ C'est à traiter ces difficultés que le MDM s'attache.



La gestion des données de référence

Le Master Data Management

3

Gestion des données de référence

- ◆ MDM = Master Data Management (Gestion de Données de Référence)

- ◆ Gartner, créateur du sigle, définit le MDM comme suit :
 - ▶ MDM is the workflow process in which business and IT work together to ensure the uniformity, accuracy, stewardship and accountability of the enterprise's official, shared information assets.

- ◆ Comme pour le terme « SOA », le terme « MDM » n'est pas déposé, et les principes ne sont pas normalisés. Il s'agit plutôt d'une « philosophie ».

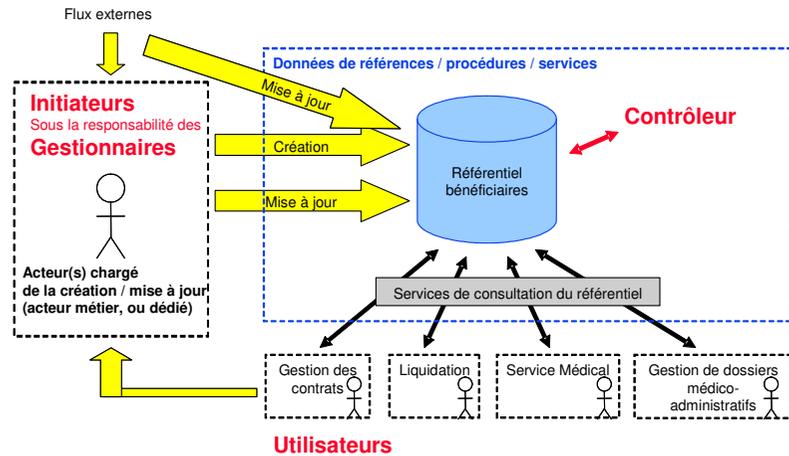
Gestion des données de référence

- ◆ Le MDM est une discipline visant à mettre en place les processus, les organisations et les outils pour rassembler, gérer et partager, de manière transverse, les données de référence.

- ◆ Le MDM offre :
 - ▶ La mise à disposition de données de référence pour les « consommateurs » de ces données dans le SI,
 - ▶ Un workflow de gestion et mise à jour de ces données de référence.

Gestion des données de référence

◆ Exemple d'acteurs liés aux référentiels :



Gestion des données de référence

◆ Les référentiels les plus courants au sein des SI :

- ▶ Référentiels clients,
- ▶ Référentiels des nomenclatures,
- ▶ Référentiels des produits,
- ▶ Référentiel de l'organisation,
- ▶ Référentiel de sécurité,
- ▶ Référentiel des règles métier.

Gestion des données de référence

Les 4 types d'architecture pour la gestion des données de référence

Gestion des données de référence

◆ Différentes architectures du MDM :

Consolidation



- **Ideal for reporting or analytics that reside in a BI/data warehouse**
- Nonintrusive to the business
- BI is the business platform
- Any industry
- Benefits dependent on success of BI strategy
- No attempt to clean up source data

Registry



- **Low control, autonomous environments**
- Nonintrusive of edge applications
- Emphasis is on remote data and application to application integration (lots of real time network access)
- Distributed governance
- Faster to implement than coexistence and centralized

Coexistence



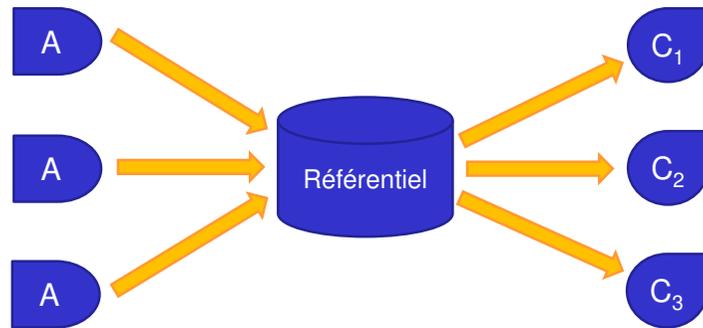
- **Large-scale distributed model**
- Largest change to information infrastructure
- Greatest need to mirror data
- Global and local governance
- Greatest risk over control, security
- Focused on shared services

Centralized



- **High control, top down environments**
- Largest change to application infrastructure
- Hugely invasive to the business
- Centralized governance
- Greatest control over access, security
- Focus on common services

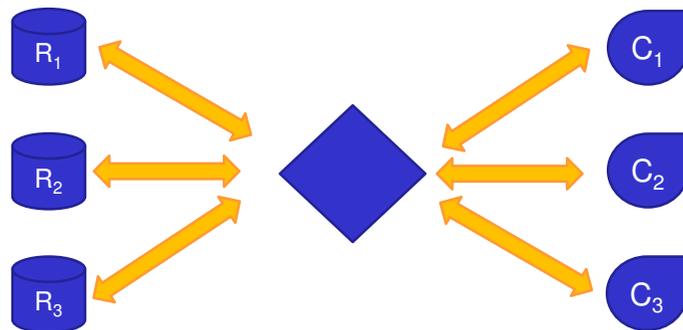
Gartner



Consolidation

Consolidation

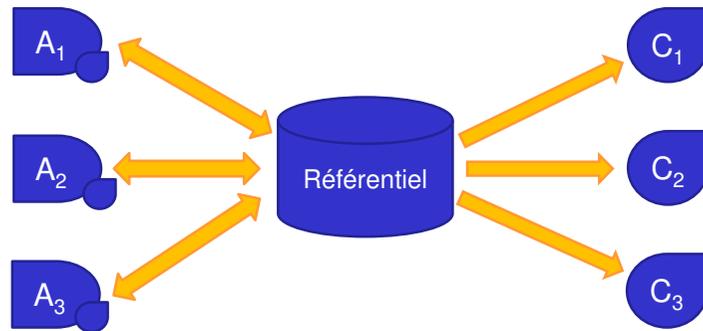
- ◆ Le référentiel est alimenté à partir de plusieurs sources (distinctes du référentiel lui-même)
 - ▶ Lien lâche entre le référentiel et ses sources ;
 - ▶ En dehors des problématiques de DQM, solution simple à mettre en œuvre.
- ◆ En revanche:
 - ▶ La solution doit embarquer l'outillage nécessaire aux traitements de la qualité des données avant validation ;
 - ▶ Pas de retour vers les fournisseurs => écart de vision entre fournisseurs et référentiel (sauf asservissement).



Répertoire virtuel (registry)

Répertoire virtuel

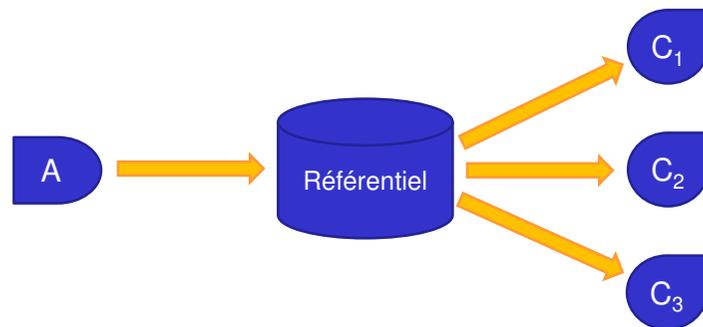
- ◆ C'est une architecture proche de l'architecture de consolidation, à ceci près qu'elle met juste en œuvre un système de type « système d'enregistrement distribué ».
 - ◆ Le référentiel peut être mis en œuvre simplement par le biais d'outils dédiés ;
 - ◆ Cf. concepts de « cache » et de « middle office ».
- ◆ En revanche:
 - ▶ Aucune facilité de gestion ou de garantie apporté par le processus référentiel ;
 - ▶ Elle peut nécessiter des outils d'intermédiation au niveau de la consommation.



Coopération (coexistence)

Coopération

- ◆ Le référentiel et les points d'acquisition interagissent
 - ▶ Le référentiel retransmet les données pour validation avant intégration au sein du référentiel
 - ▶ La validation est rendue nécessaire du fait que le référentiel peut avoir à faire pour assurer la cohérence de ses données.
- ◆ Cette architecture est très complexe à mettre en œuvre :
 - ▶ Il faut envisager les différents scénarios de synchronisation au préalable
 - ▶ Il persiste une adhérence forte entre fournisseurs et référentiels.



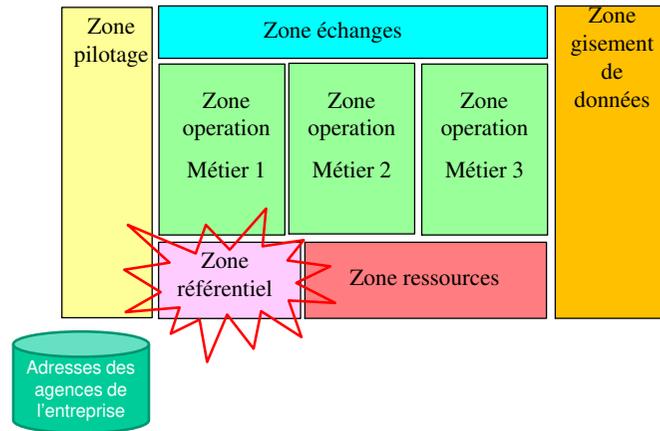
Centralisation (centralized)

Centralisation

- ◆ Le point d'acquisition est le référentiel lui-même ;
- ◆ C'est l'architecture la plus simple, la plus canonique, et certainement celle à privilégier lorsque c'est possible ;
- ◆ Elle suppose que le référentiel offre une IHM de type back office pour rendre les problématiques référentielles transparentes.
- ◆ En revanche :
 - ▶ Il faut être vigilant à l'évolutivité de la solution, qui ne doit pas devenir un frein aux évolutions du SI
 - ▶ Cela suppose aussi que la solution intègre peu ou prou des capacités de type workflow.

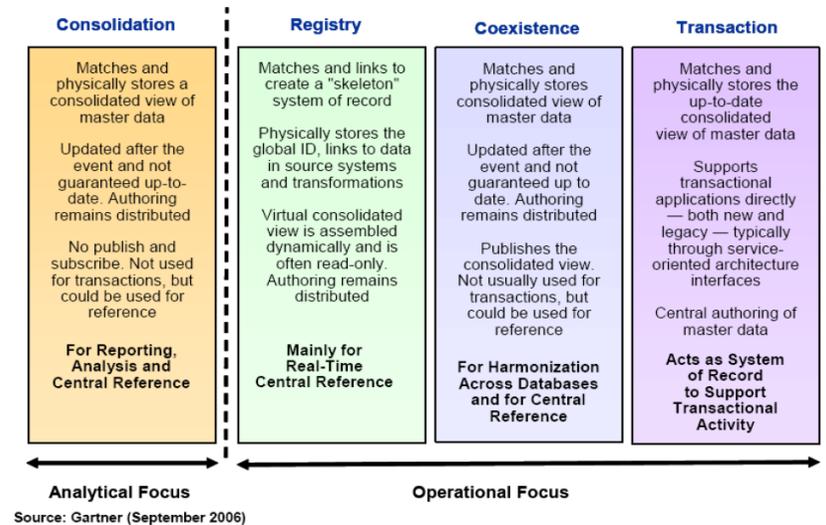
Architecture des données

- ◆ Retour sur l'exemple : dans le cas de la mise en place du Référentiel des agences de l'entreprise, quelle architecture retenir ?



Gestion des données de référence

- ◆ Différentes architectures du MDM :



Gestion des données de référence

- ◆ Plus on recherche la maîtrise des processus référentiels, plus on tend à centraliser l'architecture.
- ◆ Cela veut aussi dire que le niveau d'ambition augmente, et les risques d'échecs à la mise en œuvre aussi !
 - ▶ Sollicitations de plus nombreux acteurs,
 - ▶ Nécessité forte d'accompagnement au changement.
- ◆ Les maîtrises d'ouvrage ne sont donc pas indifférentes aux aspects architecturaux. Ils en sont à la fois à l'origine et doivent en assumer les conséquences.

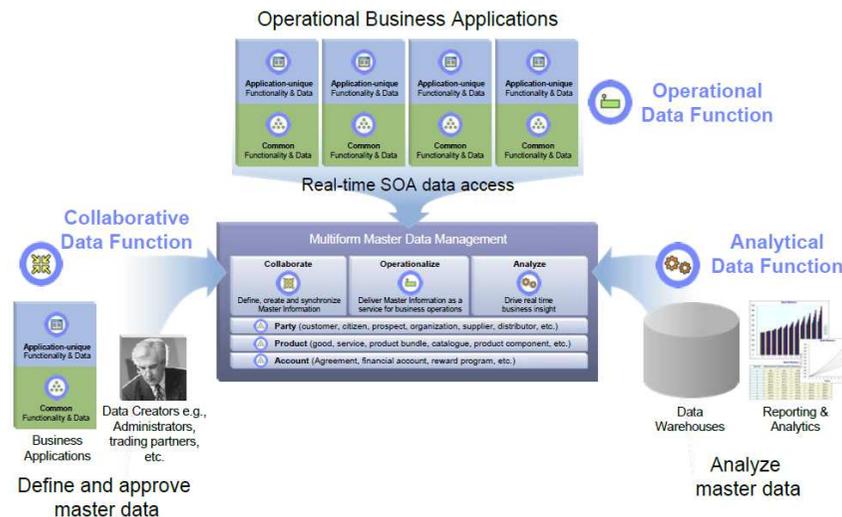
Gestion des données de référence

- ◆ Au-delà de l'architecture, il y a aussi un choix à faire de stratégie concernant l'étendue du périmètre.
- ◆ On peut choisir de mettre en œuvre un référentiel **globalisant**, qui rassemble toutes les dimensions d'un concept au sein d'un même objet (par exemple, le client vu d'un point de vue marketing, vente, comptable, etc.)
- ◆ On peut aussi faire le choix d'un référentiel **synthétique** basé sur un modèle minimum, qui ne contient que le « cœur de donnée ».
- ◆ Il s'agit de placer le curseur entre les deux :
 - ▶ Plus le référentiel est globalisant, plus il est difficile de garantir la qualité globale des données,
 - ▶ Plus il est synthétique, plus il est nécessaire de créer des liens entre référentiels.

MDM au-delà de l'architecture

- ◆ Une de force des technologies MDM est d'embarquer par construction les problématiques de pilotage, notamment la remontée d'indicateurs :
 - ▶ Complétude des données,
 - ▶ Taux de rejets,
 - ▶ Indicateurs spécifiques métier,
 - ▶ La gestion de l'obsolescence,
 - ▶ La gestion du taux de confiance.
- ◆ Les solutions MDM intègrent en outre des facilités d'auditabilité et d'analyse d'impact.

- ◆ Les fonctions associées au MDM :



- ◆ Les fonctionnalités attendues d'un MDM sont [2]:
 - ▶ Définition des rôles et droits d'accès individuels ;
 - ▶ Services de nettoyage (dédoublonnage) ;
 - ▶ Capacités de collaboration pour l'acquisition et la validation de la donnée ;
 - ▶ Service de rapprochement des données sur clef ou hiérarchie (lien entre les entités) ;
 - ▶ Fonctions de pilotage (suivi d'indicateurs)
 - ▶ Prise en charge des événements
 - ▶ Stockage et historisation.
- ◆ Les fonctions plus liées à la qualité des domaines sont plutôt du ressort d'autres types d'outils, tels que les DQM (Data quality management). Mais les frontières entre ces différents types d'outils sont parfois un peu floues.

Gestion des données de référence

- ◆ Des outils du marché (cots) existent pour la gestion des données de référence.

TYPOLOGIE DES PRINCIPAUX ACTEURS DU MASTER DATA MANAGEMENT		
Famille d'acteurs	Acteurs anglo-saxons	Acteurs français
Les spécialistes ou "pure players"	Siperian Initiate Systems Data Foundations	Orchestra Networks
Les acteurs de gestion	SAP Oracle	
Les fournisseurs d'infrastructure	IBM Tibco Sun Microsoft	
Les éditeurs de BI	SAS Teradata	
Les acteurs orientés métier	Amalto Technologies (finance) Dun & Bradstreet (commercial et marketing)	Smartco (finance)

MDM vs CRM

- ◆ Les solutions dite de CRM gère la relation avec les clients. La question peut donc se poser de savoir si les CRM peuvent être considérés comme des référentiels limités aux client.
- ◆ La réponse est « non » :
 - ▶ La priorité du CRM est le « business », pas la fiabilité des données ;
 - ▶ L'usage très opérationnel du CRM à l'échelle d'une grande entreprise va très rapidement entraîner une détérioration de la qualité des données.
- ◆ En revanche, il y a bien une synergie à trouver. On peut imaginer que les CRM s'alimentent au travers des référentiels clients de l'entreprise.

Exercice

- ◆ Etude de cas : le répertoire national inter-régime des bénéficiaires de l'assurance maladie
- ◆ Sur la base du descriptif ci-après :
 - ▶ Modéliser les informations contenues dans le RNIAM
 - ▶ Imaginer quelles peuvent être les procédures d'alimentation et de distribution de l'information
 - ▶ Quels sont selon vous les risques portant sur la qualité des informations ?
 - ▶ Le RNIAM est peu utilisé pour déterminer l'organisme. Selon vous, pourquoi ? Est-ce un problème ?

Gestion des données de référence

- ◆ Le RNIAM est un référentiel géré par la CNAV qui contient pour les assurés sociaux (assurance maladie) : le NIR (numéro identifiant, ce qu'on appelle couramment le « numéro de sécu »), les données d'état civil, la date et lieu de naissance, l'organisme d'assurance maladie (la caisse primaire pour le régime général).
- ◆ Le RNIPP est un référentiel géré par l'INSEE qui pour chaque personne enregistrée contient le nom, le prénom, le lieu et la date de naissance, de décès, le sexe, les numéros d'acte de naissance et de décès, le NIR, les mentions de notification pour les système.
- ◆ LE RNIAM est consultable par les différents organismes d'assurance maladie.
- ◆ Penser aux droits de consultation/modifications légaux, à la problématique des enfants (connus par l'assurance maladie par leurs parents), aux situations particulières (naissance à l'étranger, migrants, etc.)

En résumé

- ◆ L'architecture s'intéresse aux données dans la mesure où elles sont manipulées par les processus.
- ◆ L'architecte a un rôle important sur la qualité globale des données, notamment pour ce qui est de la cohérence, de l'intégrité, etc.
- ◆ Les enjeux sont encore plus importants pour ce qui est des référentiels, dans la mesure où les données sont massivement partagées et font l'objet d'enjeux transverses à l'entreprise ou l'organisation.
- ◆ Il ne faut pas se cantonner à une vue statique : il faut prendre en compte tout ce qui est processus de maintien de la qualité.