

Comment construire une architecture Data-Centric avec Boomi ?

Synchroniser, qualifier et activer vos données au cœur de vos activités



Erhis
Going beyond data



boomi

Table des matières

I. POURQUOI UNE ARCHITECTURE DATA-CENTRIC ?

II. QUELQUES QUESTIONS DE DESIGN AVANT DE DEMARRER

- 1. Définir les cas d'usage en premier ? 4
- 2. Quels partis pris architecturaux ? 5
- 3. Quelles ressources ? 6

III. LES GRANDES ETAPES D'UNE STRATEGIE DATA-CENTRIC

- 1. La ou les plateformes DATA au centre de l'écosystème 8
 - 1.1. Définir un périmètre et ses connections 8
 - 1.2. Mettre les données en qualité et les gouverner 9
 - 1.3. MDM, RCU 10
- 2. L'activation des données au sein de l'écosystème IT 11
 - 2.1. Les avantages d'une plateforme d'échanges 12
 - 2.2. Propager les données au service d'applications consommatrices 13
 - 2.3. Exposer les données dans des services et même au sein de parcours clients 14
 - 2.4. Ordonner les données 15
- 3. Adresser des use-cases dans des modules dédiés 16
 - 3.1. Offrir une base pour des applications dédiées 16
 - 3.2. Spécifier des modules dédiés 17
- 4. Faciliter le déploiement de la plateforme : Des uses cases « techniques » 18

A PROPOS DE BOOMI

A PROPOS D'ARHIS

Pourquoi une **Architecture Data-Centric** ?

Il est convenu que la data est le nouvel or noir des entreprises : elle alimente l'ensemble des processus et des échanges. Beaucoup de ces données sont partagées et leur gouvernance devient dès lors une exigence forte.

Cependant, dépasser le silotage des applications métiers doit être aussi synonyme de valeur ajoutée business.

L'expression 'data-driven' symbolise cette volonté que la donnée oriente l'activité. La façon d'y parvenir reste parfois moins évidente. Chez Arhis, nous travaillons sur la data-centricity, ou comment mettre la donnée au centre, comme socle et au service des stratégies métiers.

Cette ambition passe par une ou des plateformes Data associées à des mécanismes d'échanges et de traitement des données.

Nous allons à travers ce livre blanc décrire les différents composants de cette organisation.

Résumons-les :

- Accéder aux données des applications.
- Transformer les données pour les rendre opérationnelles.
- Construire une couche de métadonnées propres à chaque besoin.
- Servir les use-cases métiers : en exposant les données, en les délivrant aux applications métiers ou en construisant des fonctionnalités spécifiques à partir de celles-ci.

Pour y parvenir, Boomi offre beaucoup de leviers qui sont autant d'accélérateurs, en particulier à travers ses capacités de plateformes d'échanges et de gouvernance.

Quelques questions de design avant de démarrer

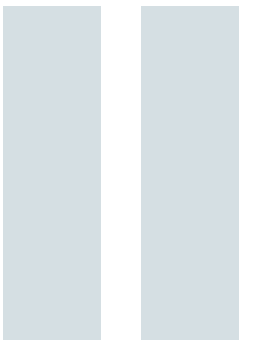
1. Définir les cas d'usage en premier ?

Il est courant de démarrer par l'identification des cas d'usage. Cet exercice permet en particulier de mesurer les pain points les plus pénalisants et d'en tirer un backlog de priorités.

Attention cependant à ne pas tomber dans l'écueil du simple regard dans le rétroviseur. La stratégie orientée data doit permettre de s'ouvrir à d'autres enjeux, et à minima à d'autres façons de faire.

Dans la pratique, la conclusion la plus fréquente de ce 'tour de piste' est de constater à quel point les cas d'usage sont divers. Devant ce constat, la conclusion qui s'impose est de s'orienter vers une architecture qui puisse adresser le plus vaste ensemble de données possibles et de pouvoir les embarquer dans une logique de mutualisation technologique. Attention en effet aux plateformes orientées sur un métier, par exemple le marketing, car une bonne part de leurs sous-jacents appartient à un socle d'entreprise commun.

Au final, la démarche pour démarrer doit s'orienter sur le choix des domaines prioritaires, qui dépendent le plus souvent de l'activité de l'entreprise. Il est évident par exemple que dans la distribution le client et le produit sont les 2 leviers les plus forts.



Quelques questions de design avant de démarrer

2. Quels partis pris architecturaux ?

Avant de démarrer un projet, il est essentiel de définir des choix de design technique qui soient pertinents sur le long terme. Traitons de quelques questions parmi les plus courantes :

TEMPORALITÉ :

Le fil de l'eau est le fil conducteur de la maîtrise des use cases. S'il convient de tenir compte de l'existant et des exigences métiers, il semble que 2 paradigmes s'imposent au regard des technologies aujourd'hui accessibles :

- Pour les flux internes : l'Event messaging
- Pour l'exposition des données : l'API

A la cible, on aura intérêt à s'appuyer sur une plateforme d'échanges unique qui gère l'ensemble des flux, le plus souvent à travers un ou des pivots de données. Nous détaillerons plus loin comment la plateforme Boomi apporte une réponse à ce challenge.

PERSISTER LES DONNÉES OU NON ?

Lorsque la modélisation est un enjeu métier fort, en particulier pour les données de références, l'option de persister les données dans la plateforme s'impose. Pour autant, 3 autres scénarios complémentaires sont possibles :

- Collecter / diffuser les données au fil des besoins via des connecteurs dédiés, sans persistance (passe-plat)
- Stocker dans des bases dédiées à un domaine, en dehors de la centralisation d'entreprise
- Utiliser la virtualisation des données pour créer des environnements dédiés

C'est l'étude des enjeux liés à ces données qui doit permettre de faire les meilleurs choix. Notamment pour les données transactionnelles de par leur forte volumétrie et volatilité, leur aspect semi ou non structuré et la diversité des formats de stockage.

Quant à la plateforme Boomi, elle offre de gérer au mieux l'intégration des données pour chacun de ces scénarios.

ARCHITECTURE CLOUD OU PAS ?

Dans le contexte d'une gestion centralisée des données, l'utilisation du Cloud et en particulier d'outils SaaS est un atout fort. Nous décrirons plus loin comment cet atout peut être utilisé pour la reprise des données. Plus généralement, nous parlons ici de la capacité à provisionner chaque brique selon les étapes du projet et potentiellement en mode run suivant des périodicités métiers.

En tant qu'IPaaS, la plateforme Boomi connecte entre elles ces différents types d'instances et propose nativement la scalabilité nécessaire.

Quelques questions de design avant de demarrer

3. Quelles ressources ?

Le succès d'une stratégie repose beaucoup sur la mobilisation des bonnes ressources. 2 questions apparaissent :

COMPÉTENCES INTERNES OU INTÉGRATEURS ?

Il est recommandé de s'appuyer le plus possible sur des 'standards' pour chaque composant. Ceci est un gage de pérennité et de la disponibilité d'une communauté d'acteurs. Cependant, la dimension 'outil' n'est pas l'enjeu véritable de ce type de projet. Un partenaire de confiance doit être choisi suivant son expérience de la maîtrise de toute la chaîne de valeur, sous tous ces aspects : fonctionnel, intégration, architecture etc. Il ne s'agit pas ici d'un simple cadrage, l'enjeu est de maîtriser les enjeux métiers, de manier des règles de gestion nombreuses, de savoir où faire quoi pour servir les besoins et optimiser les performances, sans oublier l'accompagnement au changement, et le tout dans un savoir-faire méthodologique qui optimise le TCO.

EQUIPES TECHNIQUES OU MÉTIERS ?

Trop de projets échouent du fait des appréhensions entre des profils de type DSI et des experts métiers. La vraie réponse est de mobiliser chacun selon ses compétences. Pour cela, il n'y a pas de principes standards mais une adaptation à l'organisation actuelle, associée à une projection de rôles et processus futurs. C'est le véritable contenu de la gouvernance.

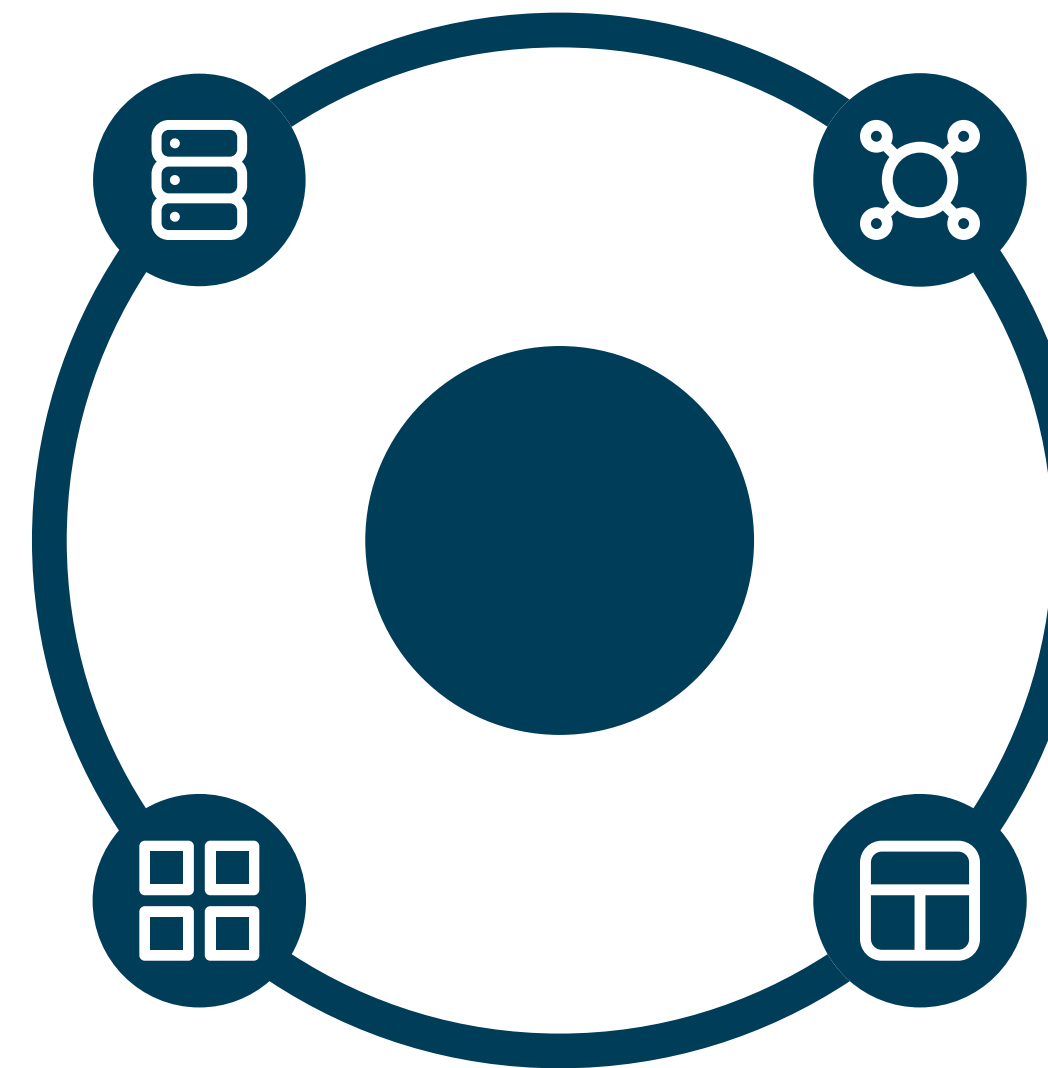
Les grandes étapes d'une Strategie Data-Centric

Il ne s'agit pas ici de reprendre les étapes d'un projet, mais de montrer les différentes facettes de ce qui peut être déployé quand les données sont regroupées au cœur de l'écosystème. Selon l'existant et les priorités de l'entreprise, on combinera ces différents éléments dans une démarche progressive, qui doit toutefois être suffisamment englobante pour adresser des premiers use cases dès les premières itérations.

Il est important de distinguer les domaines et les cas d'usage. **Les domaines** : Clients, produits, assets etc. sont chacun en partie communs à divers métiers de l'entreprise. Ainsi, un client est le sujet central du marketing mais est également un élément clé pour les départements finance, logistique, compliance etc. Au contraire, un **cas d'usage** est en rapport avec un métier, un département. Par exemple, la gestion de campagne est managée par le marketing. Les données d'un domaine sont des enjeux de partage et d'appropriation communs, alors qu'un cas d'usage nécessite des indicateurs et des fonctions, voir des outils spécifiques.

Pour aller jusqu'au cas d'usage, il faut considérer successivement plusieurs 'layers' :

- Une approche plateforme au singulier ou au pluriel qui sert de socle, et dont la vocation est de rendre les données 'exploitables' pour tous.
- L'activation des données de cette plateforme, c'est-à-dire leur exploitation par des modules applicatifs métiers ou digitaux.
- Et enfin des dispositifs spécifiques qui opèrent les données selon des couches d'indicateurs, d'opérations et de fonctionnalités métiers : dans des solutions d'éditeurs ou par le développement d'applications en propre.



Étape 1 :

La ou les plateformes DATA au centre de l'écosystème

1.1. DÉFINIR UN PÉRIMÈTRE ET SES CONNECTIONS

Pour chaque domaine, on peut distinguer 3 périmètres différents, à analyser et valider pour chaque entreprise :

- Ce qui relève du cœur de cible, et qui peut être modélisé notamment au sein de BOM (Business Object Model). On parle ici souvent de données de référence de l'entreprise.
- Ce qui relève d'un périmètre partagé par certains acteurs, et qui doit être rendu disponible pour certains scénarios, notamment transactionnels.
- Ce qui relève d'un périmètre propre à certains cas d'usage, et qui n'a pas besoin d'être partagé.

Pour constituer chacun de ces 3 périmètres, on mettra en œuvre des processus d'intégration avec des accostages différents, qui embarquent ou non des règles de gestion, et le cas échéant selon des temporalités diverses.

L'opportunité d'une plateforme de type Boomi apparaît déjà comme déterminante, offrant les atouts d'une intégration 'Anything, Anytime, Anywhere'. En tant qu'IPaaS (Integration Platform as a Service), la plateforme AtomSphere Boomi permet d'intégrer tout type d'application quel que soit le choix de l'infrastructure retenue Cloud ou On-Premise. La plateforme étant « Cloud native », il est inutile d'installer un logiciel client lourd ni de disposer de compétences de codage pointues car Boomi a l'avantage d'être un outil Low-Code pour le développement de processus d'intégration complexe.

En tant que solution iPaaS, elle supporte l'intégration temps réel et une mise à l'échelle flexible pour répondre aux besoins en volume de données nécessaires à tout type d'écosystème.

Étape 1 :

La ou les plateformes DATA au centre de l'écosystème

1.2. METTRE LES DONNÉES EN QUALITÉ ET LES GOUVERNER

Comme mentionné en introduction, l'objectif de la plateforme data-centric est de permettre de travailler sur des données 'exploitables'. Exploitable signifie d'abord exact, accessible et à jour, dans le respect des contraintes réglementaires.

Dans un premier temps, des compétences d'analyse et de design sont nécessaires. Il s'agit en particulier :

- Du data profiling, pour analyser les objets existants et la qualité des données de ces objets dans les applications sources.
- De la construction si besoin d'un modèle logique sur un périmètre donné.
- De la définition d'un glossaire d'entreprise, qui définit les métadonnées à associer à chaque objet.
- De la construction des identifiants de chaque objet et de leur lien avec ceux de chaque application.

Puis la phase de mise-en-œuvre doit traiter les données qui se doivent d'être :

- Compréhensibles et documentées.
 - Les objets sont classés et catalogués. Les objets et attributs sont typés, décrits et associés le cas échéant au modèle d'entreprise (glossaire, wiki interne etc.).
 - Les API sont décrites par leurs contrats d'interface.
- Associées à des indicateurs de qualité qualifiés : taux de complétude, répartition des valeurs, respect des formats attendus etc.
- Suivies dans leur cycle de vie : applications en charge de la collecte, traitements informatiques qui les transforment et applications qui les consomment.
- En accès sécurisé et conforme avec les exigences réglementaires.
 - Connexion à la plateforme ou API via une identification par l'annuaire d'entreprise.
 - Accès en lecture/écriture en fonction des habilitations des utilisateurs.
 - Anonymisation des données personnelles (PII) selon les profils.

Puis la phase de mise-en-œuvre doit traiter les données qui se doivent d'être :

- Description des objets (data catalog) et cartographie des traitements (data lineage) renforcées par un moteur AI propre à la plateforme qui accélère la création des liens.
- Lien avec le modèle d'entreprise (business glossary)
- Anonymisation des données personnelles (data masking).
- Description des contrats d'interface et contrôle des accès (API portal).

Étape 1 :

La ou les plateformes DATA au centre de l'écosystème

1.3. MDM, RCU

Selon les métiers de l'entreprise, la modélisation peut être traduite au sein d'un MDM pour une partie plus ou moins importante des données (ou RCU pour le domaine Client).

Il s'agit de modéliser, dédupliquer, synchroniser, nettoyer et enrichir les données de différents domaines métier et le cas échéant de les relier entre eux.

Avec Boomi Master Data Hub, on peut créer ce Référentiel (Repository) MDM pour :

- Concevoir un modèle qui représente les entités de chaque domaine.
- Déployer le modèle et identifier les sources qui vont l'alimenter ou le consommer.
- Synchroniser les données grâce à Boomi Integration pour orchestrer les flux et assurer la qualité de données.
- Gouverner les accès aux fonctionnalités mises à disposition des utilisateurs.
- Valider les données en résolvant les doublons, corrigeant les erreurs.

Finalement, on construit un point de vérité qui constitue la référence de ces données à distribuer dans tout le SI.

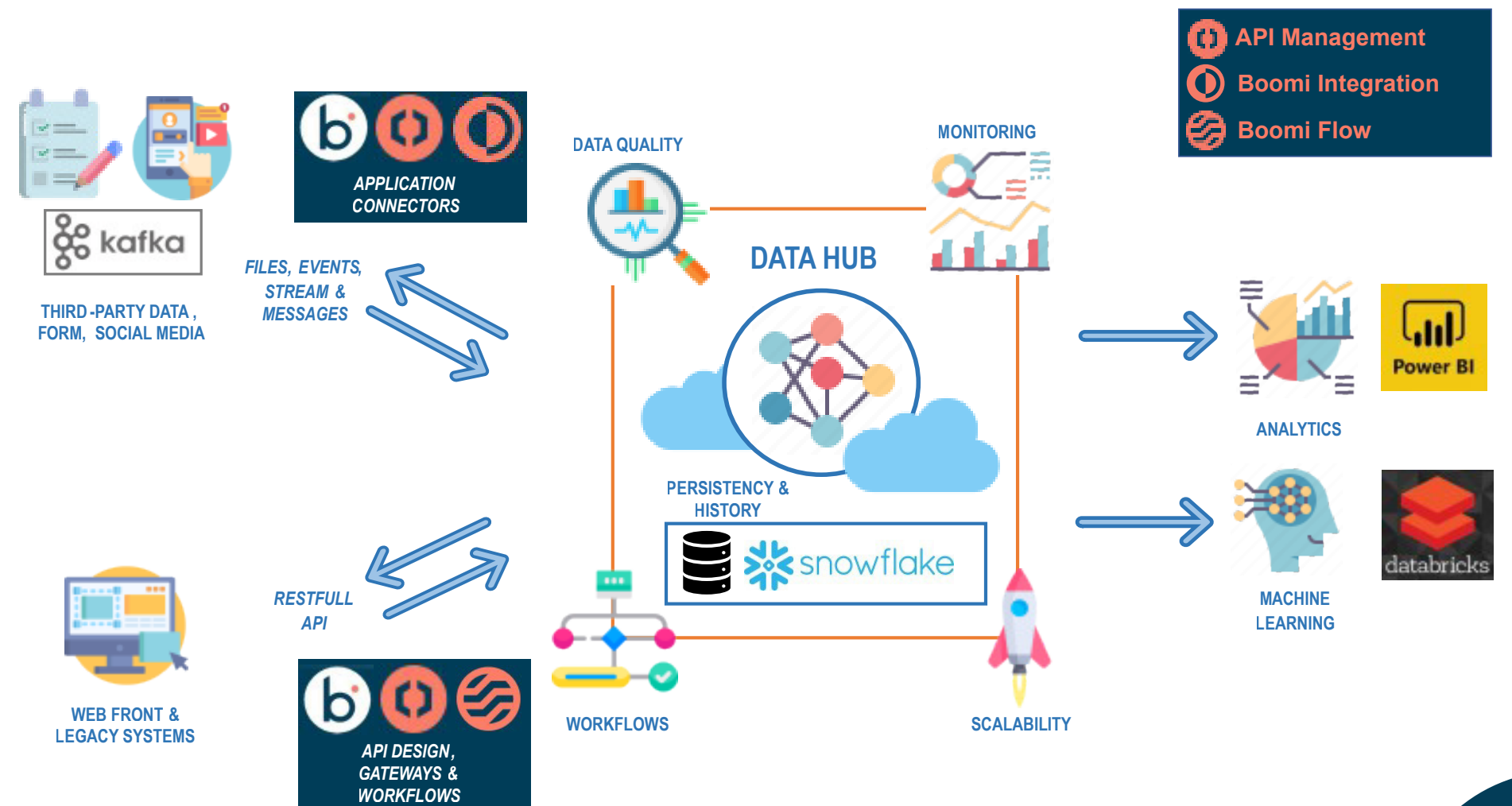
Étape 2 :

L'activation des données au sein de l'écosystème IT

Il existe de plus en plus de stratégies pour mettre les données au service de l'écosystème d'entreprise. S'il est nécessaire de travailler au cas par cas, on peut donner une vision des différentes approches :

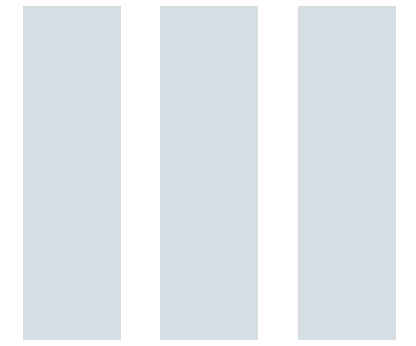
- Les données sont synchronisées au fil de l'eau avec des applications qui en sont consommatrices.
- Les données sont exposées dans des services qui sont utilisés en particulier par les fronts digitaux.
- Les données sont reprises dans des dispositifs qui les ordonnent à des fins d'analyse ou pour afficher des vues spécifiques.

Exemple de plateforme Data et de schémas d'activation.



Étape 2 :

L'activation des données au sein de l'écosystème IT

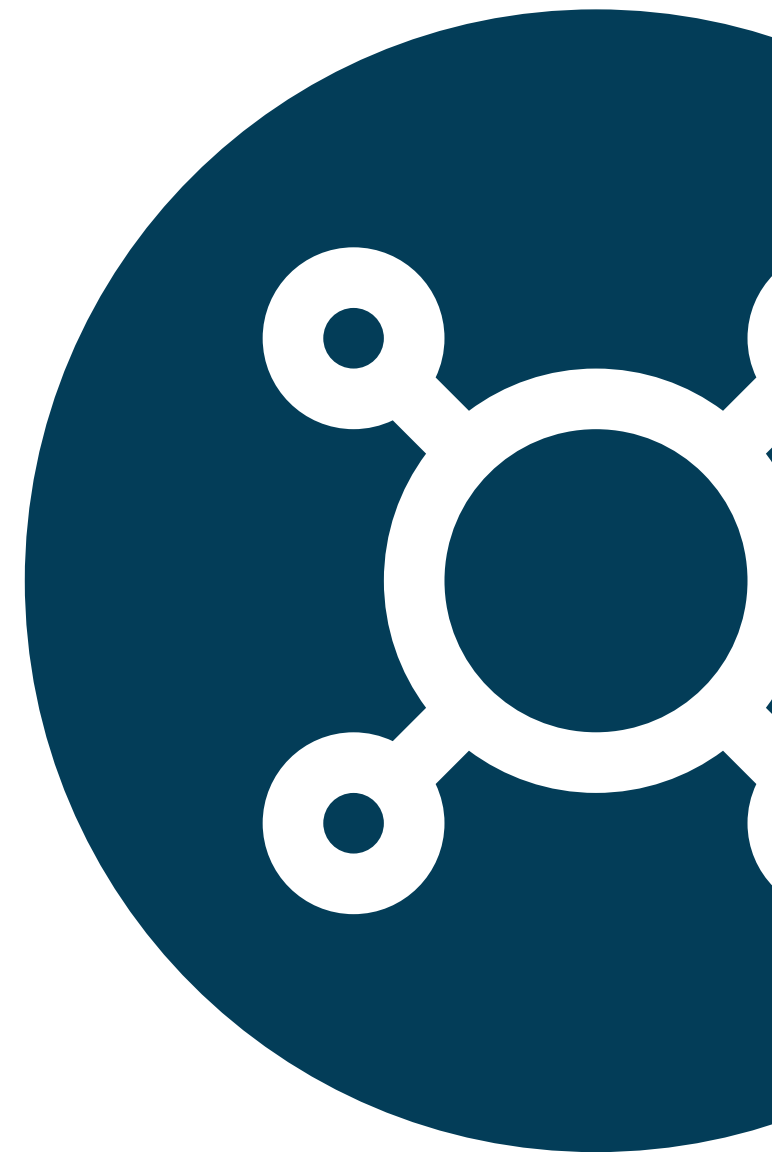


2.1. LES AVANTAGES D'UNE PLATEFORME D'ÉCHANGES

Pour pouvoir décliner les différentes stratégies d'activation des données, la maîtrise d'une plateforme unique est un gage de souplesse et d'efficacité, tout en minimisant le budget.

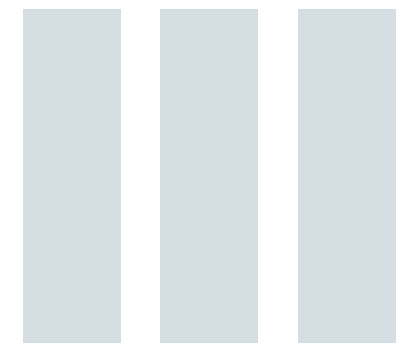
C'est ce que propose Boomi à travers une panoplie de fonctionnalités d'échanges et de services qui apportent une mutualisation pour les mises en œuvre, suivis et maintenances des connexion. Voici la liste des principales fonctionnalités :

- Pouvoir connecter des applications avec des interfaces hétérogènes en utilisant des standards interopérables ou des connecteurs applicatifs.
- Gérer le cycle de vie des services en proposant une interface permettant de créer, publier, maintenir, décommissionner une API.
- Assurer la médiation et le routage des messages en offrant des outils de transformation, gestion de règles et validation.
- Orchestrer les services et agréger les réponses pour enrichir les données d'un échange.
- Monitorer la plateforme, suivre les processus, traitements et traçage des messages entrants et sortants.
- Mutualiser des accès et gérer la politique de sécurité des applications qui accèdent à la plateforme.
- Gérer des versions pour assurer une bascule sans interruption de la continuité de service.



Étape 2 :

L'activation des données au sein de l'écosystème IT



2.2. PROPAGER LES DONNÉES AU SERVICE D'APPLICATIONS CONSOMMATRICES

La solution la plus classique pour tirer bénéfice de la plateforme est la propagation des données. Par exemple, pour le marketing, on échangera les mises à jour avec le CRM pour la gestion de la relation client et les consentements, ainsi que pour l'outil de gestion de campagne marketing.

Comme nous l'avons déjà mentionné, la meilleure pratique est de générer l'ensemble des flux en temps réel. Ceci répond aux exigences de réactivité de nombreux cas d'usage, et est facilité par les outils existants, et en particulier Boomi.

Les deux modules Boomi Integration et API Management permettent de mettre en place du temps réel selon différents patterns de flux synchrones et asynchrones qui fluidifient l'intégration avec les applications en s'adaptant à leurs interfaces.

PATTERN SYNCHRONE OU MODE CONNECTÉ

Ce pattern est utilisé dans les cas d'usage Temps réel où on va chercher la donnée à la demande quand on en a besoin. L'application émettrice de la requête appelée Client passe par l'API pour solliciter l'application destinatrice qui va effectuer le traitement souhaité et renvoyer la réponse.

Contrainte : Toutes les applications doivent être disponibles pour faire circuler la donnée, ce qui impose la mise en œuvre d'architectures hautement disponibles.

PATTERN ASYNCHRONE OU MODE NON CONNECTÉ

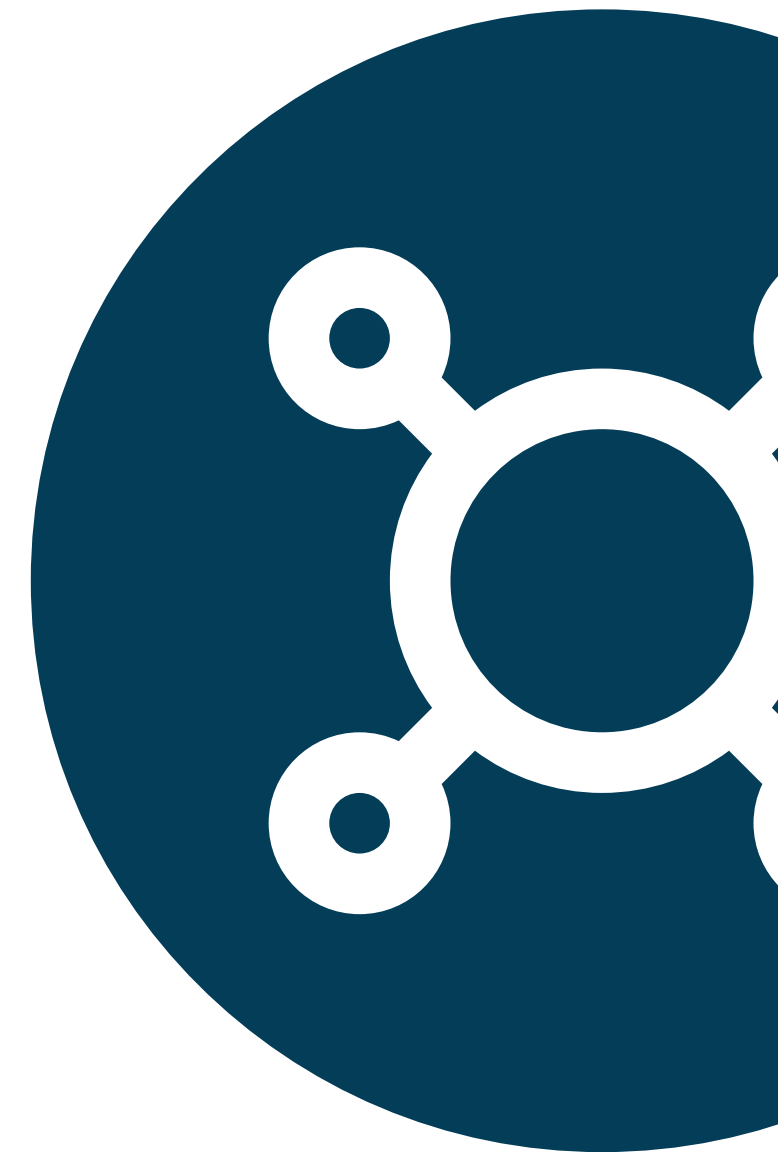
Le pattern asynchrone s'implémente en interagissant avec un MOM (Middleware Orienté Message) ou Message Broker pour assurer l'échange des données.

Tous les acteurs n'ont pas besoin d'être connectés au même moment sauf le Message Broker qui doit être hautement disponible. Ce pattern permet de réduire les contraintes de disponibilité de l'architecture et de réduire les coûts. C'est pourquoi il sera favorisé pour les échanges internes.

Il y a 2 possibilités pour mettre en œuvre un échange asynchrone :

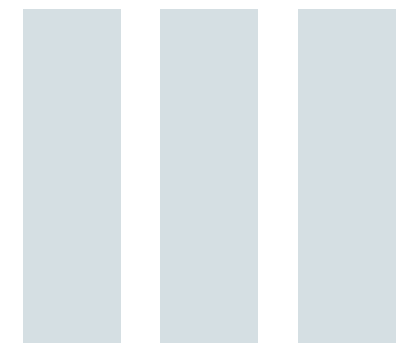
- **Publish-suscribe** : l'application productrice de la donnée publie un événement/message et les applications consommatrices s'abonnent à ce type d'événement/message et les consomment selon leurs besoins.
- **Flux d'événement (Streaming)** : l'application productrice de la donnée publie les événements/messages dans un journal et les applications consommatrices peuvent accéder au journal et rejoindre le flux d'événement à tout moment et à n'importe quel endroit selon leurs besoins.

NB : L'API peut être utilisée dans le mode asynchrone pour accéder aux données d'une application qui n'a pas de connecteur vers les MOM. La plateforme Boomi sera alors responsable de transformer cet appel en message.



Étape 2 :

L'activation des données au sein de l'écosystème IT



2.3. EXPOSER LES DONNÉES DANS DES SERVICES ET MÊME AU SEIN DE PARCOURS CLIENTS

La data-centricity doit aussi permettre de gérer les interactions entre l'entreprise et le monde extérieur. La gestion des API de Boomi dans une logique de services est ici le levier d'appui de cette stratégie.

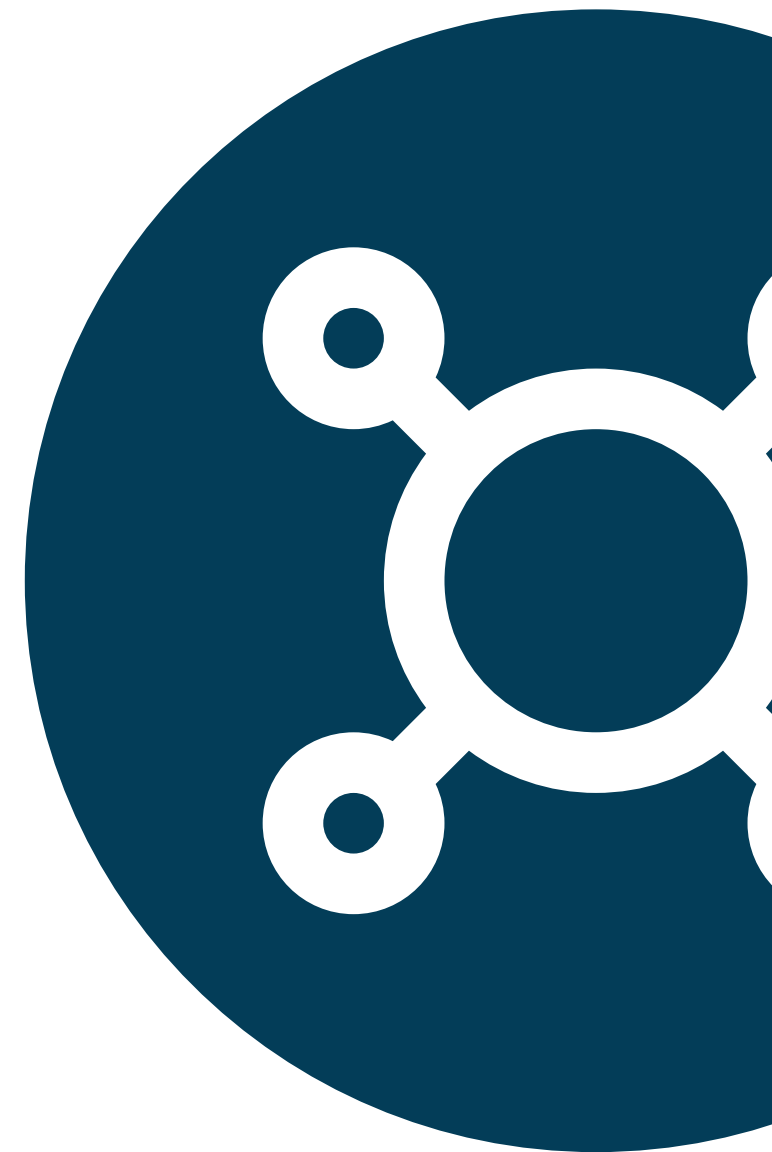
Ajoutons que la plateforme Boomi permet d'aller plus loin en permettant une logique de parcours : Boomi Flow est un environnement de développement low-code qui permet de construire des parcours clients (Customer Journeys) et d'automatiser des workflows ou processus métier. Il assure la création et le déploiement d'applications web qui peuvent manipuler et synchroniser la donnée pour l'utiliser dans d'autres applications mobiles ou web.

Ainsi, les entreprises qui disposent d'une offre digitale aboutie peuvent proposer à leurs clients d'accéder à leur services et enseignes via des interfaces digitales multicanales qui s'adaptent à tous les canaux disponibles en allant des postes fixes aux tablettes et smartphones.

Pour une expérience utilisateur optimale, l'identification de l'utilisateur devra se faire une seule fois et permettre l'accès à tous les services

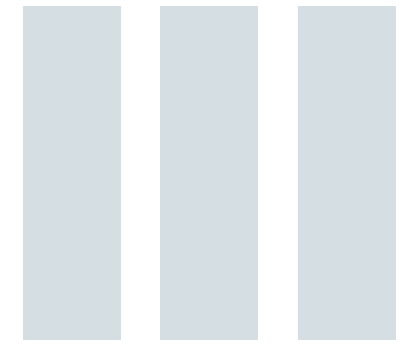
du client durant son parcours digital. On déploiera pour cela une brique en charge de l'authentification : le CIAM (Customer Identity & Access Management) qui va centraliser, gérer, suivre les accès des utilisateurs et appliquer les politiques de sécurité suivant les profils retenus. Les utilisateurs qui vont accéder aux APIs exposées par la plateforme Boomi à travers les canaux digitaux seront alors authentifiés et selon leurs habilitations pourront consommer un périmètre plus ou moins important de données.

Selon les contraintes en termes de performance et de disponibilité des données, il est recommandé de mettre en place un cache pour servir efficacement les API temps réel car les traitements de mise en qualité sont souvent complexes et lourds. Ainsi, la création ou la mise à jour via l'API dédiée sera envoyée à la plateforme et stockée dans le cache Temps Réel. Ces données seront disponibles immédiatement pour la suite du parcours client. Pour implémenter le cache Temps Réel, on peut utiliser des solutions NoSQL comme MongoDB, Couchbase ou Cassandra. Au final, le parcours digital s'affranchit des impacts de la complexité des règles ou processus métiers de l'entreprise.



Étape 2 :

L'activation des données au sein de l'écosystème IT



2.4. ORDONNER LES DONNÉES

La 3ème manière d'activer les données est de leur donner du sens en les ordonnant suivant des structures d'analyse, de reporting ou simplement pour dégager des vues spécifiques.

Citons les 2 usages les plus courants, complémentaires l'un de l'autre :

- L'Analytics qui permet d'agréger les données sous forme de rapports et d'explorer les données suivant différents axes.
- La Data Science qui permet d'analyser, interpréter un grand nombre de données et révéler des tendances, des corrélations, des logiques au-delà d'algorithmes préconstruits.

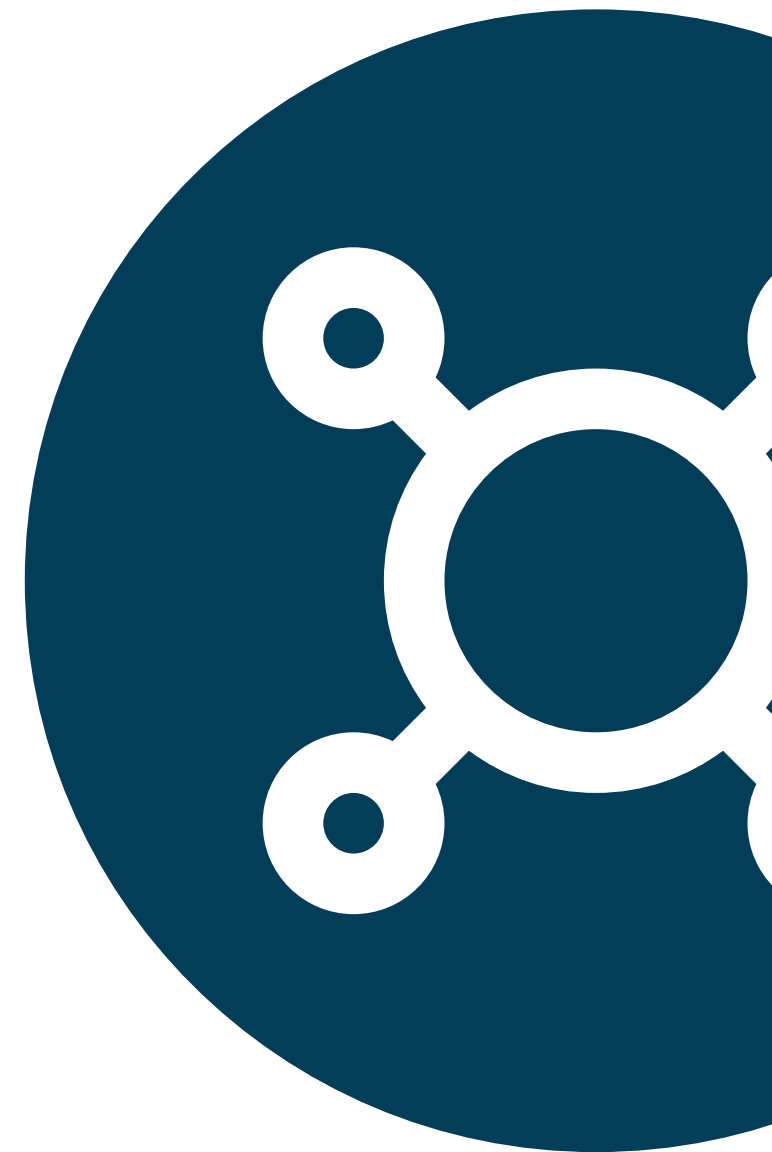
Evidemment, cette activation est fortement corrélée aux choix d'architectures sous-jacentes, comme des data warehouses relationnels, une base NoSQL ou un data lake HDFS. C'est pourquoi la plateforme peut être plurielle en termes de stockages.

Mais de plus en plus, les activités d'exploration et d'analyse s'appuient sur des outils qui unifient les accès en couvrant un périmètre large de connectivités et de langages de requête. PowerBI pour l'Analytics et Databricks pour la data science sont des solutions courantes qui offrent ce type de performances.

Pour autant, il peut être nécessaire d'ajouter une étape de transformation « à la volée » (sans recopie ni datamart) afin de présenter les données de la plateforme sous des formats qui permettent certains types d'analyse ou de visualisation.

Pour cela, les outils précédemment cités s'appuieront sur des fonctionnalités qui permettent à chaque utilisateur de façonner une image spécifique d'une donnée partagée, voire la croiser avec une donnée qui lui est propre :

- La data preparation qui permet de transformer les données à la volée, sans recopie ni datamart (PowerBI par exemple via sa librairie DAX).
- La virtualisation de données qui permet la mise en œuvre de principes de gouvernance fédérée sur des données hétérogènes. Le modèle exposé se compose de différentes couches, ce qui permet la réutilisation des entités communes. Ces entités communes sont portées par la couche mappe du modèle « Data Centric » tandis que les couches spécifiques subviennent aux besoins spécifiques des différents métiers. Cela permet aux concepteurs de réutiliser facilement des données d'autres domaines, ce qui réduit la charge d'intégration.



Étape 3 :

Adresser des use-cases dans des modules dédiés

Chaque métier adresse des besoins opérationnels en propre de consommation des données. Prenons des exemples de use cases du domaine client :

- Proposer de nouveaux critères de sélection marketing pour augmenter l'audience en croisant l'historique des campagnes avec le profil des clients, leurs valeurs etc.
- Gérer la pression commerciale en supprimant certains clients déjà sollicités selon des règles paramétrables.
- Détecter les comportements frauduleux et alerter / bloquer les services aux clients suspects, afin de prévenir les actes malveillants.
- Anticiper les mécontentements en surveillant les réseaux dans des modèles prédictifs et alerter pour augmenter la réactivité sur la gestion et la résolution de conflits.

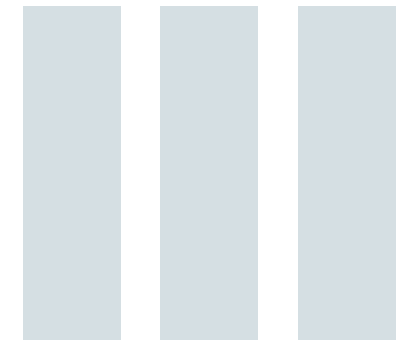
2 options permettent d'adresser ces besoins opérationnels :

3.1. OFFRIR UNE BASE POUR DES APPLICATIONS DÉDIÉES

Il n'est pas toujours nécessaire ou opportun de reproduire et synchroniser les données de la plateforme dans des applications tiers. Au contraire, il est plus simple que ces applications utilisent la plateforme comme leur propre base de données.

Ce peut être le cas par exemple pour une CDP (Customer Data Platform). Ses nombreuses fonctions : segmentation, ciblage, animation ... appellent les données de la plateforme via des webservices ou/et des connecteurs. Ainsi, les utilisateurs métiers, par exemple les marketeurs pour la CDP, peuvent utiliser toutes les données de l'entreprise comme leviers d'actions, quand des applications traditionnelles sont réduites aux attributs remodelées dans leur propre base.

NB : Précisons que la plateforme centralisée n'est pas la seule base de données support. C'est là qu'interviennent en complément des bases spécifiques à un métier, en sus de la plateforme d'entreprise.



Étape 3 :

Adresser des use-cases dans des modules dédiés

3.2. SPÉCIFIER DES MODULES DÉDIÉS

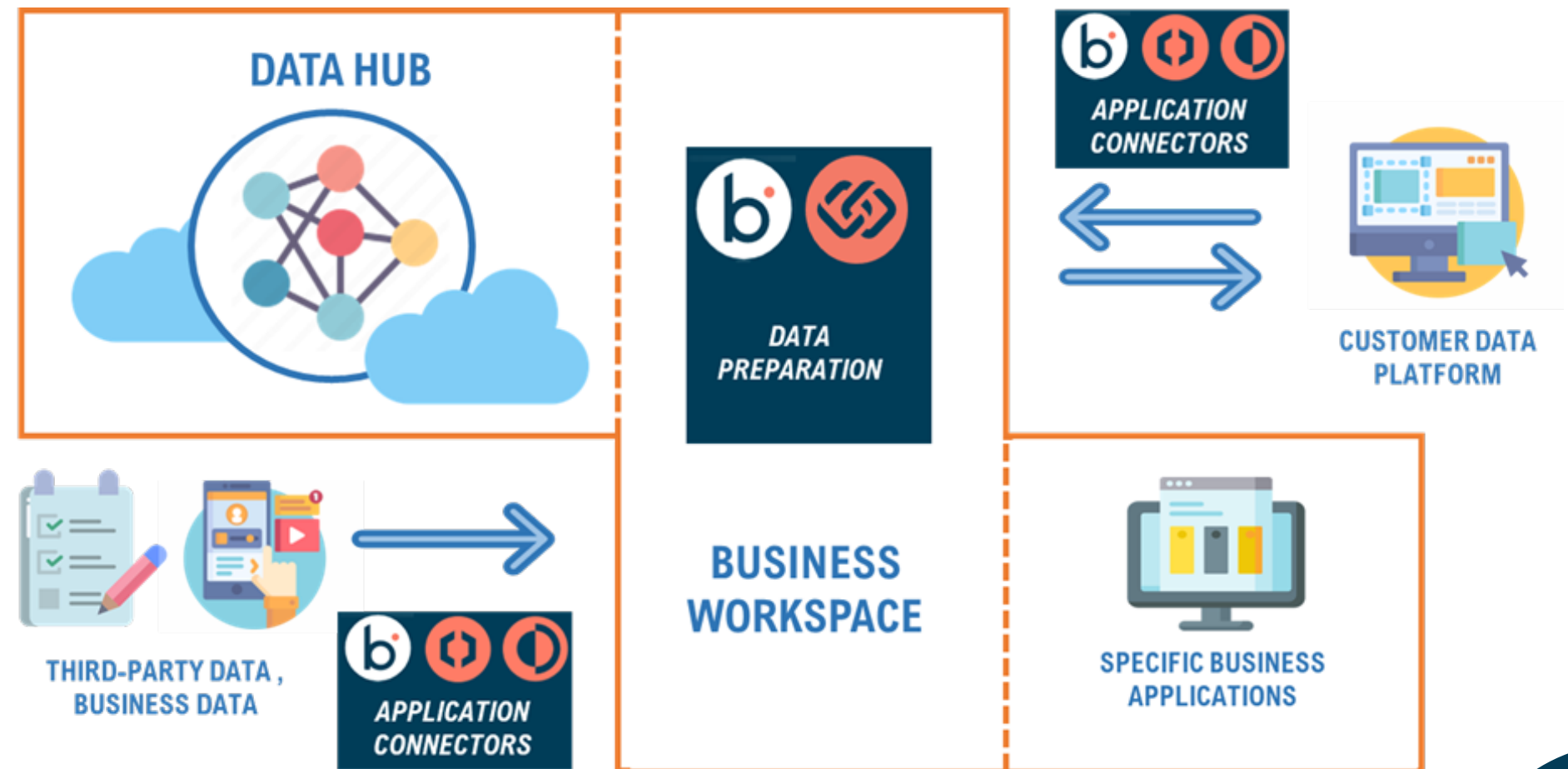
Quand aucune application métier ne couvre ces besoins, il est opportun de créer de l'intelligence et des fonctions propre à un métier à partir de couches de données personnalisées qui sont comme des extensions de la plateforme centralisée.

Il est alors pertinent de réaliser des environnements métiers spécifiques, qui s'appuie sur des visions dédiées, dans lesquelles on définit et alimente les métadonnées propres aux use cases. Le module de Data Preparation de Boomi permet justement de définir des objets propres à partir des données brutes de la plateforme centralisée ou de sources locales. En pratique, par simple clic, il est possible de :

- Joindre plusieurs data sets.
- Appliquer des filtres sur les valeurs non souhaitées.
- Ajouter des attributs calculés.
- Agréger les informations et construire leurs propres indicateurs.
- Planifier et suivre des tâches de transformation / exports en fichiers.

Sur la base de la préparation de ces données, on peut construire tout type de traitements et fonctions, y compris pour des use cases complexes, qui intègrent des interfaces utilisateurs etc.

Exemple de plateforme Data et de schémas d'activation.



Étape 4 :

4. Faciliter le déploiement de la plateforme : Des uses cases « techniques »

En plus des usages métiers, la plateforme Boomi permet d'accélérer la mise en œuvre de la plateforme data-centric elle-même. En voici l'illustration dans un use case commun à tous les projets Data.

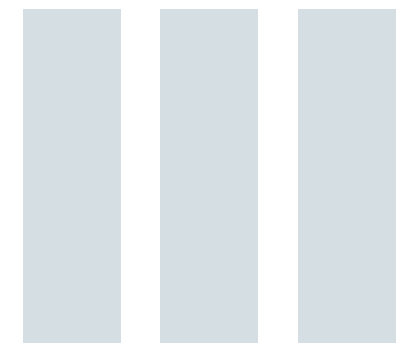
Nous avons montré comment il est préférable de privilégier autant que possible les échanges au fil de l'eau. Voyons ici comment, en s'appuyant sur Boomi, on peut traiter le chargement initial des données suivant cette logique :

Le cycle d'une architecture Data Centric s'accompagne de chargements de données au fur et à mesure de l'accostage de nouveaux systèmes fournisseurs de données.

Le chargement initial de données d'un nouveau système est bien souvent nécessaire pour maintenir un périmètre de données cohérent sans lequel l'exploitation des données en aval de l'architecture (ex. BI, gestion de campagne) deviendrait inefficace. La difficulté principale du chargement initial vient du fait des volumétries concernées dans ces échanges et du temps nécessaire pour extraire, transformer puis persister la donnée, ce qui a pour conséquence de dégrader les niveaux de services ou même de les

interrompre. Souvent les SI utilisent des outils ETL plus performants pour résoudre ces problématiques de traitements à forte volumétrie en mode batch. Cette approche a l'inconvénient de dupliquer l'effort d'intégration nécessaire car la dichotomie batch vs temps réel impose souvent de réimplémenter la même logique que celle des outils orientés messages, sans oublier les coûts de licence éventuels.

Il est donc logique de chercher une solution qui puisse adresser les 2 uses cases : chargement initial et ingestion des changements au fil de l'eau. La plateforme AtomSphere offre plusieurs avantages à cet effet car elle permet de déployer un traitement- workflow- sur un ensemble d'atomes appelés molécules, ce qui, conjointement avec les possibilités d'exécution multithread, permet d'obtenir des performances dépassant les 10.000 transactions/minute. Lorsque le chargement initial est réalisé, l'architecture peut rapidement être ramené à sa configuration nominale. De cette façon le surcoût dû au provisionnement des ressources matérielles additionnelles utilisées pendant la durée du chargement initial n'est que temporaire.



A propos de Boomi :

BOOMI alimente l'économie de la donnée en permettant aux organisations de connecter instantanément, et depuis n'importe où, les données aux outils et process, **grâce à sa plateforme unifiée ouverte et native dans le cloud**. Plus de 17 000 clients dans le monde font confiance à BOOMI pour sa **rapidité, sa facilité d'utilisation et son coût total de possession réduit**. En tant que pionnier de l'utilisation intelligente de la donnée, BOOMI permet aux clients et aux partenaires de **découvrir, de gérer et d'orchestrer rapidement et facilement les données, tout en connectant les applications, les processus, les objets et les personnes** pour obtenir des résultats commerciaux encore meilleurs et plus rapides.

Pour plus d'informations au quotidien sur BOOMI:

Suivez-nous sur notre site <https://boomi.com/fr/> et sur notre page **LinkedIn**

A propos d'Arhis :

Arhis est une société internationale de services spécialisée sur **les projets Data** depuis 2003. Les bureaux principaux sont en France, en Espagne, et à Genève où se situe également son siège social. Arhis dispose de **plus de 150 consultants intervenant sur les sujets MDM, PIM, Data Quality, Data Integration, Customer Data Platform et plus généralement Data Hub**. Ils sont devenus en quelques années un acteur incontournable du marché de la donnée depuis sa **collecte jusqu'à son exposition, en incluant sa structuration : transformation, consolidation, cataloging, segmentation ... le tout au service des applications métiers et digitales**. La force d'Arhis est de proposer une structure unifiée qui sait s'adapter au contexte et aux enjeux de ses clients selon leurs tailles, leurs existants, leurs ambitions et leurs priorités.

ARHIS en quelques chiffres :

- Plus de **19 ans d'expérience**
- Des bureaux en **Suisse, France et en Espagne**
- Plus de **60 projets de hub de données en production**
- **85%+ des projets sont également en maintenance chez Arhis**

Une stratégie de partenariat forte basée sur les technologies du moment et les méthodologies les plus adaptées

Pour plus d'informations au quotidien sur Arhis :

Suivez-nous sur notre site www.arhis.com et sur notre page **LinkedIn**.

Besoin de conseil sur la data, contactez-nous à info@arhis.com