

GDF SUEZ DOLCE VITA

LA QUALITÉ DE DONNÉES INFORMATICA AU SERVICE DE L'EFFICACITÉ COMMERCIALE ET MARKETING



« Informatica Data Quality nous offre une tranquillité absolue : sur 9 millions de clients, avec une vingtaine de sources de données branchées, n'importe quel système aurait eu des anomalies. En 6 mois, rien de tel ne s'est produit. Les données sous contrôle qualité sont corrigées, unifiées et réinjectées dans les systèmes opérationnels : plus personne ne peut prendre prétexte du manque de qualité des données pour expliquer de mauvaises performances. »

– David Legendre, Responsable de l'équipe Données, GDF SUEZ Dolce Vita

Evaluation des besoins

GDF Suez Dolce Vita vend du gaz, de l'électricité et des services sur une base de 9 millions de clients en France. Avec une base de clients de cette ampleur et une base de prospects tout aussi étendue, les problématiques de qualité de données appellent nécessairement des réponses de niveau industriel. « Avant la mise en œuvre du projet AMADEUS, les informations sur les clients étaient collectées de manière éparse, hétérogène et souvent non structurée, avec un silo pour chaque canal de communication », souligne David Legendre, responsable de l'équipe Données de l'équipe BI de GDF SUEZ Dolce Vita.

Contexte

Cet état de fait avait des conséquences multiples : au niveau opérationnel, la logique de silo faisait obstacle à l'obtention d'une vue 360° des clients et à une gestion véritablement multicanal ; au niveau financier, il était impossible d'effectuer certaines études, par exemple

sur les décaissements/encaissements d'une année, parce qu'aucun système n'assurait la collecte et la validation de ces informations ; au niveau managérial, l'utilisation de KPI n'ayant pas le même contenu aboutissait à une « guerre des chiffres » entre les sous-directions ; enfin, d'un point de vue tactique, il était impossible de piloter l'activité des prestataires externes gérant les envois d'emails et de SMS aux clients, et de mesurer la qualité réelle des données récoltées par les conseillers Dolce Vita alors que la qualification de données entre pour partie dans le calcul de leur rémunération.

« Ces constats et l'absence de couche de certification de la collecte de données nous ont décidés à mettre en place une structure de gouvernance de données avec, comme outillage, un entrepôt de données commun à toutes les sources capitalisant de l'information sur les clients et les prospects, et un outil industriel permettant de mesurer la qualité de données et de faire du redressement de données », explique David Legendre.



Avantages

- Liberté de choisir les technologies reconnues comme les plus performantes pour chacune des briques
- Rentabilité rapide de la solution au vu du service rendu sur le volume initial et sur les traitements récurrents
- Prêt de logiciels afin de réaliser l'audit de tous les systèmes d'information existants

Défis

- Un projet industriel mené à bien en un an
- Un entrepôt alimenté par 40 sources, diverses bases, plusieurs fichiers structurés et des référentiels externes (La Poste, Insee, Météo)
- Implémenter plus de 100 règles de qualité de données
- Mesurer les résultats
- Mener le projet communément avec les métiers et les fonctionnels

Remédier aux problèmes structurels de qualité de données

Cet état de fait avait des conséquences multiples : au niveau opérationnel, la logique de silo faisait obstacle à l'obtention d'une vue 360° des clients et à une gestion véritablement multicanal ; au niveau financier, il était impossible d'effectuer certaines études, par exemple sur les décaissements/encaissements d'une année, parce qu'aucun système n'assurait la collecte et la validation de ces informations ; au niveau managérial, l'utilisation de KPI n'ayant pas le même contenu aboutissait à une « guerre des chiffres » entre les sous-directions ; enfin, d'un point de vue tactique, il était impossible de piloter l'activité des prestataires externes gérant les envois d'emails et de SMS aux clients, et de mesurer la qualité réelle des données récoltées par les conseillers Dolce Vita alors que la qualification de données entre pour partie dans le calcul de leur rémunération.

« Ces constats et l'absence de couche de certification de la collecte de données nous ont décidés à mettre en place une structure de gouvernance de données avec, comme outillage, un entrepôt de données commun à toutes les sources capitalisant de l'information sur les clients et les prospects, et un outil industriel permettant de mesurer la qualité de données et de faire du redressement de données », explique David Legendre.

Architecture : une approche 'Best of Breed'

Pour construire cette structure, GDF Suez Dolce Vita a opté pour une approche 'Best of Breed', choisissant les technologies reconnues comme les plus abouties et les plus performantes pour chaque brique : Informatica

PowerCenter pour l'alimentation de l'entrepôt de données, Microsoft BI pour la partie restitution. Pour la qualité de données, Informatica Data Quality s'est imposé pour son interfaçage natif avec Informatica PowerCenter, son intégration poussée avec Teradata – permettant notamment de déléguer à la machine Teradata les traitements de qualité de données –, et la perspective d'une rentabilisation rapide de la solution au vu du service rendu sur le volume initial et sur les traitements récurrents.

Le challenge : démontrer la valeur du concept au management

La phase de réflexion sur ce projet a duré deux ans. « *Compte tenu de l'ampleur du projet et de ses implications organisationnelles, nous avons dû faire beaucoup de pédagogie en interne. Mais, en août 2010, nous sommes arrivés à un point de blocage, faute de pouvoir apporter au top management des preuves tangibles de la valeur d'une démarche globale de qualité des données* », se souvient David Legendre. « *Ce qui a débloqué la situation, c'est le partenariat avec Informatica qui a mis à notre disposition, en mode prototypage, ses outils de qualité de données, en particulier les outils d'audit de la qualité des données* ».

Le prêt de ces logiciels dans le cadre d'un POC (*Proof of Concept*) permet à l'équipe projet de réaliser, pendant 6 mois, l'audit de tous les systèmes d'information existants au sein de GDF Suez Dolce Vita. L'audit confirme que des millions et des millions de données sont mal qualifiées, fausses, partielles ou erronées. Le POC démontre également qu'il est possible de contrôler les données saisies par les conseillers et même de corriger des fautes de

frappe ou des erreurs de saisie sur des millions d'enregistrements au quotidien.

Un projet industriel mené à bien en un an

Opérationnel depuis l'automne 2012, le projet AMADEUS a été mis en œuvre en un an, étape par étape, en commençant logiquement par les processus d'alimentation de l'entrepôt de données, puis le déploiement de la couche qualité de données, suivie de la partie BI. En mars 2013, une étape supplémentaire est franchie avec la mise en place de la couche gouvernance de données permettant de réinjecter les données corrigées et retraitées dans tous les applicatifs de GDF Suez Dolce Vita et faisant d'AMADEUS le référentiel de l'organisation sur le périmètre des données relatives aux caractéristiques des clients, prospects et des logements qu'ils occupent.

Fin juin 2013, l'entrepôt est alimenté par 20 sources dont les deux plus conséquentes sont SAP CRM et SAP ISU (facturation/réclamations). Sont également intégrées via Informatica PowerCenter diverses bases Oracle et SQL Server, des fichiers structurés (.csv, .xml) interfacés avec les prestataires, et des référentiels externes (La Poste, Insee, Météo...). Deux sources supplémentaires sont ajoutées chaque mois pour atteindre la cible de 40 sources.

Plus de 100 règles de qualité de données ont été implémentées. Elles corrigent, enrichissent, normalisent quotidiennement les données des applications de front-office des conseillers et des prestataires. En phase initiale, elles ont permis de corriger 90% des fiches clients/prospects : 9 fiches sur 10 comportaient des erreurs et des manques, parfois minimes mais

toujours préjudiciables. Au total, 15 millions de fiches clients et prospects ont été corrigés, complétées et validées grâce à la technologie de qualité de données Informatica. La qualité des données ainsi obtenue a été un facilitateur pour les prises de contacts des conseillers clientèles, la maîtrise de la pression commerciale, une meilleure connaissance du portefeuille qui facilite la personnalisation de la relation client, le gain de temps sur la qualification des informations...

« Les algorithmes d'unification d'identité sont un point fort majeur d'Informatica Data Quality. Le rapprochement et la réconciliation automatiques permettent de reconstituer des historiques perdus et de disposer d'une vision 360° du client. »

Un chantier essentiel : le dispositif de mesure des résultats

De l'avis de David Legendre, définir les éléments de mesure des résultats des traitements de qualité des données a été assez compliqué à mettre en place : « A des fins de pilotage et dans une logique d'amélioration continue, il est nécessaire de tracer ce qui déclenche l'application de telle ou telle règle », explique-t-il. « La modélisation des règles de contrôle et de suivi a été complexe. L'exploitation est en revanche très facile, avec des reporting basés sur des requêtes SQL simples. Nous pouvons savoir, par exemple, combien de règles de normalisation des années

de naissance ou de rectification d'adresse e-mail ont été passées au cours d'une période donnée, quand, où et pourquoi ».

Ce travail de suivi et d'analyse est réalisé par l'équipe Données en collaboration permanente et itérative avec l'ensemble des métiers – une équipe de 6 personnes en charge de la gouvernance et de la qualité des données, ainsi que de la gestion des échanges entre les applications de type référentiel et avec les prestataires externes. A chaque ajout d'une source ou d'une nouvelle règle de qualité de données validée par les métiers, les analyses sont reconduites de façon à vérifier que tous les cas fonctionnels ont été prévus.

« Il est dispensable que les personnes dédiées à l'analyse des reportings connaissent les logiques métiers et les outils de qualité de données. Leur rôle est de réévaluer sans cesse ce qui a été fait et de voir ce qu'il faut affiner. »

Le début d'une vraie révolution

Avec AMADEUS, GDF SUEZ Dolce Vita a enclenché une véritable révolution en cassant les silos d'information et en se dotant des moyens techniques et humains de contrôler et piloter la qualité de données à l'échelle de l'organisation.

Quelle a été la clé du succès de ce projet ? Pour le responsable de l'équipe de Données, c'est en premier

lieu d'avoir été piloté par les métiers et les fonctionnels comme un projet organisationnel et non comme un projet de système d'information. Le second facteur clé de réussite est la mise en place d'une équipe dédiée à la qualité et à la gouvernance de données, capable d'analyser et de mettre en place des règles de qualité et de contrôle pertinentes au regard des processus, besoins et pratiques métiers. Pour conclure, le métier est (re) devenu le propriétaire de ses données dans lesquelles il peut maintenant avoir confiance : le contenant est devenu plus important que le contenu. La collaboration Business et IT constitue étalemment un facteur de succès de ce projet.

« Nous avons atteint un niveau où le système vit et peut évoluer. Nous entrons aujourd'hui dans une deuxième phase qui est, de mon point de vue, la plus intéressante : celle où l'on prend des décisions sur des données valides et non plus sur des intuitions ; celle où on a la capacité de cibler très précisément les campagnes marketing en fonction, par exemple, de l'appétence des clients au courrier, à l'e-mail ou au SMS et d'analyser finement les taux de retour. C'est là, au niveau de la performance opérationnelle des métiers, qu'est le vrai ROI du projet », conclut David Legendre.



Siège mondial, 100 Cardinal Way, Redwood City, CA 94063, États-Unis Téléphone : +33 1 42 04 89 00 (France) www.informatica.com/fr [linkedin.com/company/informatica](https://www.linkedin.com/company/informatica) twitter.com/InformaticaFr

© 2013 Informatica Corporation. Tous droits réservés. Informatica® et Put potential to work™ sont des marques commerciales ou des marques déposées d'Informatica Corporation aux États-Unis et dans d'autres pays. Tous les autres noms de sociétés et de produits sont la propriété de leurs détenteurs respectifs et peuvent avoir fait l'objet d'un dépôt de marque.