



FUTURES

LABS2005

LABS2005

GUIDER DANS

UN MONDE DE

DOMAINES

# FUTURE BRIGHT

LAISSEZ-VOUS GUIDER DANS  
UN MONDE DE DONNÉES



# TABLE DES MATIÈRES

- 5 Avant-propos par Bert Boers
- 7 Introduction par Jeroen Dijkxhoorn
- 10 Le prestataire de services aux RH Securex s'oriente vers une solution CRM fiable à 100 %
- 14 L'intégration des données évolue grâce aux Big Data et à l'open source
- 18 Gestion des données (Infographie)
- 20 Rijkswaterstaat bénéficie d'une visibilité complète sur ses performances
- 24 Faire de la gestion des données une question d'actualité à l'échelle de l'entreprise
- 32 Améliorer l'information sur la conjoncture économique de la Belgique
- 36 DSM prend le contrôle des données maîtresses
- 42 Qui gouverne vos données ?
- 46 La nouvelle DSI
- 48 Crédito y Caución ajoute la qualité des données à son modèle de gestion
- 52 Des données tributaires de l'amélioration de la qualité
- 56 Feu Vert utilise la technologie de Data Management de SAS pour valoriser ses données clients
- 60 Les décisions étayées par des données font toute la différence
- 64 La gestion des données maîtresses, pierre angulaire de votre activité
- 69 A propos de SAS

# AVANT-PROPOS



**Nous avons souhaité dans cette publication aborder le thème d'une «réalité orientée données». Peu d'entreprises réalisent la rapidité avec laquelle notre société guidée par les données évolue. Cette progression est due à des développements tels que l'Internet des objets, qui génère de vastes flux de données à l'origine de nouvelles opportunités commerciales. Récemment, le cabinet d'études de marché Forrester a qualifié l'époque actuelle d'« ère du client » : les clients laissent une trace numérique et attendent de leurs fournisseurs qu'ils utilisent ces informations pour leur proposer une offre plus adaptée et personnalisée.**

Ces deux évolutions ont des effets considérables dont nombre d'entreprises commencent aujourd'hui à prendre conscience. Pourtant, peu s'engagent dans une action structurée pour se préparer véritablement à cette réalité orientée données, ce qui est d'ailleurs compréhensible, puisqu'il s'agit d'une terre inexplorée. Comment s'y prendre ? Par où commencer ?

Les entreprises savent qu'elles doivent tout d'abord mettre en place des bases solides. Mais les solutions de facilité sont nombreuses, notamment les projets lucratifs d'analyse des Big Data. Comment ces deux univers sont-ils corrélés ? Quels investissements produisent le retour le plus rapide ?

Le propos de cet ouvrage est d'ouvrir de nouvelles perspectives. À partir d'entretiens avec des clients tels que Securex, DSM, Feu Vert, Rijkswaterstaat et Crédito y Cautión et des experts comme Jill Dyché et Ronald Damhof, nous expliquons les étapes nécessaires pour inscrire la gestion des données au cœur de votre organisation. Vous pourrez ainsi acquérir les bases et tirer pleinement parti de vos données pour favoriser l'innovation, la conversion et la satisfaction.

Nous espérons que cette lecture vous sera utile.

**Bert Boers**

Vice-président de la région Europe du sud-ouest  
SAS Institute



**Jeroen  
Dijkhoorn**

DIRECTEUR DU CENTRE  
D'EXCELLENCE  
PLATE-FORME  
ANALYTIQUE SAS

Depuis la création de la toute première base de données, la qualité des données a toujours été une préoccupation. Ce sujet a toutefois été longtemps négligé, pour la simple raison que l'efficacité des processus était toujours jugée plus importante que l'exhaustivité et la précision des données. Les données étaient un sous-produit des processus. Ce temps est révolu. Nous entrons désormais dans une réalité guidée par les données.

Résultat : les bases de données comportaient énormément d'erreurs et d'omissions. Pour faire face à ce problème, les données étaient systématiquement validées avant d'être utilisées. Si de nombreuses erreurs étaient trouvées, tous les efforts étaient soudain concentrés sur la recherche des données manquantes, la correction des inexactitudes et/ou le nettoyage des bases contaminées. L'intervention humaine était constamment nécessaire.

### Déclenchement automatique des processus par les données

Aujourd'hui, cette méthode pose des problèmes car les sources de données sont de plus en plus liées alors que les processus sont susceptibles de démarrer à tout moment. Si la date de lancement d'une campagne marketing par e-mail était auparavant fixée par un professionnel du marketing, elle est désormais déterminée par des déclencheurs reçus du client auxquels vous pouvez réagir. Mieux vous comprendrez le parcours du consommateur, plus votre entreprise réagira facilement à ces déclencheurs et s'adaptera à ses clients. Vous devez pour cela instaurer des règles précisant comment votre entreprise doit réagir lorsque le client ou prospect demande certaines informations ou s'inscrit à un bulletin d'information par e-mail.

Le processus se poursuit de manière entièrement automatique, sans intervention humaine et donc, sans l'étape de validation auparavant indispensable. L'exactitude des données doit par conséquent être vérifiée automatiquement, au moyen de services pouvant être déployés tout au long du processus. À ce stade, il est possible de distinguer la validation (correction technique des données dans un flux physique) de la qualité des données (vérification de l'exactitude fonctionnelle).

### Réalité orientée données

Les entreprises étaient jusqu'à présent organisées en fonction des processus. Aujourd'hui, ce sont les données qui fixent le cap. Cela signifie que la moindre défaillance dans l'identification des erreurs peut avoir immédiatement un impact considérable. La correction manuelle n'étant plus possible, l'erreur apparaîtra à plusieurs endroits, ce qui rend le suivi de la qualité des données beaucoup plus important. C'est également la raison pour laquelle les règles et réglementations de conformité imposent de nouvelles exigences de qualité des données. Aujourd'hui, les superviseurs veulent des données, et non des rapports, ce qui nécessite une organisation orientée données. Nous nous dirigeons rapidement vers cette réalité.

Le problème n'est pas tant la technologie que le manque de suivi centralisé de la cohérence des définitions de données – autrement dit, de la gouvernance des données. Cette tâche incombe généralement au directeur des données, qui fait défaut à la plupart des entreprises.

### Rôle moteur de l'ère du client et de l'Internet des objets

Il est urgent de passer à l'action, car l'ère du client impose de réagir rapidement aux déclencheurs des consommateurs. Une vision à 360° du client est donc indispensable. S'il en est question depuis des années, la répartition des données client dans différents systèmes l'a toujours rendue impossible, d'autant plus que le manque de supervision des définitions de données empêche le regroupement de ces dernières.

Les développements liés à l'Internet des objets constituent un autre moteur. Celui-ci entraînera un nouveau flux de données que vous pourrez utiliser pour optimiser vos processus et les automatiser en grande partie. Ce phénomène nécessite également une vision appropriée de la gestion des données.

### Combinaison de différents types de données


Que votre entreprise soit mue par l'une ou l'autre de ces réalités, il devient de plus en plus important de combiner des données fiables à 100 % à des données présentant un certain degré d'incertitude. Les prévisions météorologiques ou les analyses du ressenti sur les réseaux sociaux en sont des exemples. Comment associer correctement ces données non structurées, souvent stockées dans des clusters Hadoop, à des données structurées parfaitement exactes, comme la planification des trajets d'un chauffeur routier ou les données de comportement d'achat ?

**« Les entreprises étaient jusqu'à présent organisées en fonction des processus. Aujourd'hui, ce sont les données qui fixent le cap »**

Jeroen Dijkxhoorn

En termes de coût, il est impossible de stocker toutes ces données dans la même base. Et d'un point de vue organisationnel, cela ne serait pas non plus souhaitable comme Ronald Damhof l'explique plus loin. Après tout, les données dont vous devez rendre compte à vos supérieurs sont très différentes de celles que vous utilisez pour effectuer des expériences et trouver des idées innovantes. Pourtant, ces différents modes d'utilisation doivent coexister sans regrouper physiquement la totalité des données.

Une telle complexité exige un modèle logique et des définitions de données limpides. Sans ces définitions et une bonne gestion des données, il est impossible d'exploiter les opportunités qui apparaissent sur le marché et que vos concurrents s'empresseront de saisir. La question n'est donc plus de savoir si vous allez vous lancer, mais quand. Notre réponse est sans appel : dès maintenant. Les données sont votre capital le plus précieux. Agissez en conséquence et exploitez-les avant qu'un concurrent ou un nouveau venu ne prenne les devants. ■



LE PRESTATAIRE DE SERVICES AUX RH SECUREX S'ORIENTE VERS UNE SOLUTION CRM FIABLE À 100 %

## NETTOYER DÉFINITIVEMENT LA BASE DE DONNÉES DES RELATIONS CLIENT

À l'instar de nombreuses entreprises, le prestataire de services aux RH Securex était confronté à de graves problèmes concernant sa base de données de gestion de la relation client. La principale difficulté était la mauvaise qualité des données marketing, dont la fiabilité ne cessait de se dégrader. Le prestataire a nettoyé et mis à jour sa base de données à l'aide d'une plate-forme SAS® Data Management. Par ailleurs, Securex généralise actuellement cette plate-forme comme outil de contrôle permanent afin de garantir l'exactitude et la cohérence des mises à jour en mode batch et des manipulations de données manuelles. Résultat : une base de données améliorée, contenant des informations beaucoup plus fiables.

Présente en Belgique, en France et au Luxembourg, la société Securex propose des services RH aux grandes entreprises, aux PME, aux travailleurs indépendants et aux particuliers. Elle offre ainsi des services de gestion de la paie, de gestion du personnel et du droit social, de conseil en ressources humaines et de prévention des risques et accidents. Employant environ 1 600 personnes dans près de 30 agences, Securex apporte ainsi son aide à plus de 250 000 clients.

### L'incohérence des données est source de frustration

« Nous voulons être certains que lorsqu'un membre de l'entreprise saisit ou modifie des données, les changements soient automatiquement traités et rectifiés », explique Jacky Decoster, architecte métier chez Securex. La moindre incohérence des données entraîne invariablement une énorme frustration chez les personnes concernées. Le personnel met constamment à jour les données client, ajoute et modifie des coordonnées et données contractuelles, le tout pendant que les équipes marketing chargent des informations extérieures pour de nouvelles campagnes et autres communications client. « Chacune de ces manipulations est susceptible de produire de petites erreurs ou incohérences », observe Jacky Decoster. « Dans la mesure où les données sont manipulées par une multitude de personnes et d'équipes, des problèmes – doublons, enregistrements client contenant des informations contractuelles incomplètes, coordonnées manquantes telles que le prénom, le sexe, l'adresse postale, l'adresse e-mail ou le numéro de téléphone – peuvent facilement survenir. Tout ceci est frustrant, en particulier pour les équipes marketing en plein dans une campagne : de nombreux e-mails sont purement et simplement renvoyés et certains courriers sont adressés deux fois à la même personne, sans parler de ceux qui contiennent des erreurs ou des omissions. Notre réputation en a parfois sérieusement souffert. »

Malgré une base de données CRM SAP centralisée mise en place en 2004, les problèmes n'ont fait qu'empirer au fil des années. Jacky Decoster a remarqué que les plaintes portant sur la qualité des données émanaient aussi bien du personnel que des clients. « Nous devons de toute évidence intervenir et le faire de manière efficace et convaincante. »

### Le lancement du nettoyage réussi avec SAS Data Management

Le problème de qualité des données est devenu la priorité numéro un de Securex lors du lancement de son projet Client+ global. Ce projet prévoyait la migration de la base de données CRM SAP ultra-personnalisée vers la solution cloud Salesforce.com, évolutive et normalisée. Pour faciliter cette migration, Securex a décidé de déployer SAS Data Management. Jacky Decoster explique que cette approche a porté ses fruits : « SAS Data Management nous a permis de nettoyer méticuleusement les données avant de les charger dans notre nouvelle base. Les données ont été standardisées, les doublons fusionnés et les informations manquantes, automatiquement ajoutées. Les outils intégrés à SAS Data Management – définition de dictionnaires de données, correspondance approximative, analyse du nom complet, détermination fiable du sexe, normalisation des numéros de téléphone, analyse des adresses e-mail, etc. – couvrent complètement nos besoins. Nous avons d'ores et déjà réalisé la migration de nos comptes entreprises en un temps record et le service marketing assure ne recevoir pratiquement aucune réclamation concernant la qualité des données. C'est un formidable pas en avant qui aurait été inimaginable sans SAS Data Management. Nous en sommes actuellement au stade de la finalisation des comptes travailleurs indépendants et particuliers. »

«Le service marketing affirme que la qualité des données s'est considérablement améliorée, ce qui nous semblait impossible à réaliser auparavant»

Jacky Decoster

### Un contrôle permanent de la qualité des données

Jacky Decoster souligne toutefois que l'amélioration de la qualité des données ne se fait pas du jour au lendemain et doit faire l'objet d'une attention constante. C'est l'une des raisons pour lesquelles Securex a opté pour une approche globale. Son projet Client+ prévoit la redéfinition et la rationalisation des processus marketing et commerciaux. Cette initiative consiste notamment à sensibiliser le personnel à l'impact des manipulations de données, en insistant sur l'attention et la précision. En parallèle, SAS Data Management devient l'outil de contrôle permanent de la qualité des données. Selon Jacky Decoster, « on ne peut jamais être sûr à 100 % que chaque donnée sera correctement saisie, même si le personnel est parfaitement formé et sensibilisé. Nous avons par conséquent paramétré SAS Data Management pour effectuer régulièrement – chaque semaine en l'occurrence – des contrôles de cohérence et des mises à jour. La prochaine étape vise à mettre en place un contrôle pratiquement en temps réel. Dès qu'un collaborateur de l'entreprise saisit ou modifie des données, l'enregistrement concerné est automatiquement traité et corrigé par SAS Data Management. Ce processus ne dure que quelques secondes. »

### Une architecture robuste et une grande souplesse

Jacky Decoster et le personnel de Securex ne tarissent pas d'éloges pour la robustesse de l'architecture et la grande souplesse de la plate-forme SAS Data Management. Le système peut en effet être intégré à n'importe quel environnement logiciel. Par exemple, SAS Data Management comporte des connecteurs directs certifiés avec différents systèmes – dont Salesforce.com –, ce qui évite le développement d'interfaces personnalisées. D'autre part, l'ensemble des fonctionnalités sont proposées sous forme d'applications stockées, ce qui garantit la sécurité et la fiabilité de chaque transaction.

SAS Data Management est en outre facile à déployer. « La plate-forme dispose d'un puissant profileur de données, ce qui nous permet d'examiner les données disponibles et d'en évaluer la fiabilité, ainsi que les risques encourus en cas d'intégration dans de nouvelles applications. Nous l'utilisons notamment pour analyser les données que nous achetons. »

Qui plus est, le logiciel offre un puissant outil de définition de job batch pour nettoyer et normaliser les données à partir des statistiques et informations de profilage. « L'apprentissage de SAS Data Management est très rapide : après une formation de deux jours, nous étions capables de définir tous les jobs dont nous avons besoin », se félicite Jacky Decoster pour conclure. ■

Comment réagir face à l'explosion des données  
et à l'importance de l'analyse

# L'INTÉGRATION DES DONNÉES ÉVOLUE GRÂCE AUX BIG DATA ET À L'OPEN SOURCE

« Pour moi, les big data n'existent pas en tant que volume. » Voilà une déclaration étonnante de la part d'un spécialiste de l'intégration de données. « La taille est relative. Elle ne reflète que l'origine. » Il est du reste impossible de définir un seuil minimal pour mesurer l'ampleur des big data, mais ce phénomène concerne le domaine de l'intégration de données, lui-même devenu quasiment une spécialisation banalisée.

Ces doutes sur l'étendue réelle des big data émanent d'Arturo Salazar, qui occupe le poste de Business Solution Manager pour la région Europe du sud-ouest chez SAS Institute. Arturo Salazar explique que le terme « big » (volumineux) revêt un tout autre sens pour une PME que pour une grande entreprise telle qu'un établissement financier. Selon lui, il ne peut y avoir de seuil minimal pour évaluer l'ampleur des big data.

Il est certain que ce phénomène influe énormément sur le domaine de l'intégration de données, qui doit aujourd'hui faire face à davantage de variables et de données inconnues. Arturo Salazar ajoute que l'intégration des données existe maintenant depuis un certain temps et est pratiquement considérée comme une tâche ordinaire. Toutes les entreprises ne sont pas pour autant à l'aise avec l'intégration de données : certaines ne mesurent pas l'importance d'une utilisation et d'un déploiement efficaces des données. Mais en tant que spécialisation, elle a aujourd'hui atteint le stade de la maturité.

## Sortir de la zone de confort des entreprises

A. Salazar convient que l'ampleur des big data est effectivement une dimension relative. « Les big data, ce sont toutes les données qui sortent de la zone de confort d'une entreprise. » Ils impliquent également l'exhaustivité des sources de données et, plus important encore, la rapidité d'intégration des informations pour en tirer de nouveaux enseignements exploitables. Il est important de savoir si les big data sont liés au degré de maturité d'une organisation. Si les données en question sont juste en-dehors de la zone de confort, ne s'agit-il pas simplement d'une question de développement ? « C'est effectivement une question de maturité, » conclut Arturo Salazar.

## Le reporting mensuel est inadapté

« Les entreprises sont en pleine mutation technologique par rapport au mode de collecte et d'utilisation des informations », poursuit Arturo Salazar. Il cite l'exemple frappant des débuts de l'intégration de données il y a déjà quelques années. « Prenez l'exemple d'Internet et des journaux des serveurs web. » Les machines consignent dans leurs fichiers journaux les visites et les clics enregistrés sur les pages web, notamment l'adresse IP des serveurs d'origine, les données provenant des cookies, etc.

« Tous ces clics représentent beaucoup de données. » Et toutes ces données pourraient être utiles. Selon Arturo Salazar, « la plupart des données consignées peuvent être purement et simplement éliminées, mais on ne sait pas lesquelles garder. » Le regard sur les données relatives à la navigation Internet a en outre beaucoup changé. Des données qui étaient auparavant effacées peuvent désormais s'avérer très précieuses. Savoir sur quelles pages d'un site web un visiteur cliquait ne présentait jusqu'alors que peu d'intérêt, tandis que ces informations peuvent aujourd'hui être cruciales. « Voyez les recommandations en temps réel émises par les sites marchands modernes. » Autre exemple : le suivi des habitudes de navigation lorsque les visiteurs sont connectés à un site. Le profilage est à présent le fer de lance des campagnes de fidélisation de la clientèle des entreprises présentes sur le web.

### Augmentation des volumes de données et coûts de stockage

L'exemple des journaux Internet a convaincu le plus grand nombre de l'utilité de l'intégration des données, qui a elle-même favorisé la demande de déploiement de solutions d'intégration pour un large éventail d'applications. La mise en relation des données issues des sites web avec les sources d'informations « traditionnelles » – systèmes CRM, entrepôts de données, etc. – est un gage de valeur ajoutée. Elle est depuis longtemps reconnue comme offrant de nombreux avantages pour les entreprises. Depuis, des informations en constante évolution garantissent qu'elle ne se limite pas simplement à l'importation unilatérale de journaux de sites web dans une application CRM, par exemple. L'utilisation efficace des données nécessite un trafic bidirectionnel et une portée étendue, d'où des volumes de données toujours plus importants.

Dans un premier temps, la multiplication rapide des données recueillies, stockées et corrélées par les entreprises ne semble pas poser de problème majeur. La capacité des supports de stockage ne cesse d'augmenter, tandis que le prix par giga-octet est en baisse. Les performances des processeurs connaissent une croissance exponentielle, comme si les disques durs appliquaient leur propre version de la loi de Moore. Pourtant, non seulement la courbe de la capacité de stockage augmente moins nettement que celle des processeurs, mais sa progression est en outre insuffisante pour faire face à l'explosion des données.

### Extraire des pépites de données inconnues

Les logiciels, en particulier ceux de base de données, constituent un autre problème d'intégration des données face à l'explosion des informations. La majeure partie de l'imposante montagne de données ne peut pas être simplement stockée dans une base de données relativement onéreuse ou un coûteux entrepôt. Même si ces énormes volumes de données recèlent un véritable trésor, nul n'en connaît encore la taille et l'emplacement. Selon les experts SAS, il s'agit en fait du problème de la poule et de l'œuf : la valeur encore inconnue des données opposée à leur coût de recherche. Mais un espoir se dessine pour cette forme de data mining. Grâce aux nouvelles technologies, les précurseurs n'ont plus à dégager à la main des pépites noyées dans les flux s'écoulant de la montagne des données, ni à y creuser laborieusement des puits de mine.

### Suivre la même voie que Linux

C'est là qu'intervient le logiciel Open Source Hadoop, une solution abordable, capable de stocker et de traiter plusieurs pétaoctets de données, qui s'exécute sur du matériel standard. Quelle est sa puissance ? Hadoop repose sur une technologie développée par le géant des moteurs de recherche Google pour indexer Internet. « Hadoop suit la même voie que Linux », explique Arturo Salazar. Le marché l'adopte progressivement pour des applications plus sérieuses. « Personne ne souhaite stocker ses journaux dans une coûteuse base de données. » Toutefois, Hadoop n'en est qu'au début du chemin, tandis que Linux a déjà parcouru une longue distance, ce qui pose problème à de nombreuses entreprises. L'un et l'autre sont issus d'univers très différents de ceux auxquels la plupart des sociétés sont habituées et requièrent une certaine connaissance technique, y compris côté utilisateurs.

« Au début, les gens avaient tout aussi peur de Linux », rappelle Arturo Salazar. Depuis, des entreprises comme Red Hat ont associé le logiciel central du système à des applications métier, et en proposent les résultats sous forme de packages. Hadoop n'en est qu'au début de ce processus. Notre



Arturo Salazar

BUSINESS SOLUTION  
MANAGER SAS RÉGION  
EUROPE SUD-OUEST

## « Même si ces montagnes de données recèlent un véritable trésor, nul n'en connaît encore la taille et l'emplacement »

Arturo Salazar

expert cite Cloudera et Hortonworks ; selon lui, ces programmes joueront pour Hadoop le même rôle que celui exercé par Red Hat dans l'adoption de Linux. « De nombreuses entreprises jugent encore Hadoop intimidant et trop compliqué », décrypte Arturo Salazar. Normalement, elles emploient des spécialistes pour ce type de logiciel Open Source, que ce soit pour l'installation, la configuration ou la maintenance, mais aussi pour l'utilisation au quotidien. Quelles sont les compétences nécessaires ? Des programmeurs expérimentés, dotés de compétences en programmation et d'une certaine aisance administrative, mais aussi des connaissances et de l'expertise habituellement réservées aux analystes de données. Cette somme de qualités est rare et par conséquent, coûteuse.

### Mettre Hadoop à la portée du plus grand nombre

Hadoop connaît une popularité grandissante malgré sa complexité. « Il offre d'innombrables avantages », souligne Arturo Salazar. L'éditeur de solutions décisionnelles SAS suit également cette tendance. Selon A. Salazar, SAS utilise un socle technologique basé sur Hadoop. La complexité de ce logiciel étant ainsi dissimulée dans des processus et programmes avec lesquels le client est déjà familiarisé, les entreprises peuvent utiliser efficacement des outils d'intégration de données sans faire appel à des spécialistes du logiciel sous-jacent.

En février 2015, SAS lançait un nouveau produit de gestion des données pour accroître la convivialité de Hadoop. Arturo Salazar assure que la nouvelle application web baptisée SAS® Data Loader for Hadoop permettra de creuser encore plus profondément dans la montagne de données. Cette application permet aux utilisateurs lambda comme aux analystes de préparer, puis d'extraire les données stockées dans Hadoop. Bientôt, nous serons tous capables de trouver un trésor ! ■

**POURQUOI EST-ELLE IMPORTANTE ?**

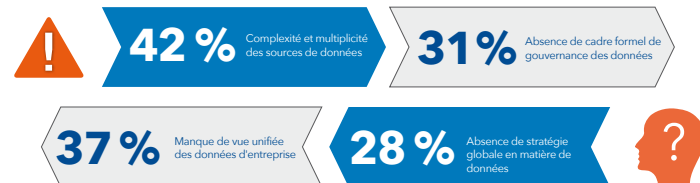
Une gestion efficace des données offre l'assurance de données exactes, alignées sur les objectifs de l'entreprise et garantes de meilleures décisions.

**DES DONNÉES CORRECTEMENT GÉRÉES SONT TRAITÉES :**

<p><b>AU BON ENDROIT</b></p> <p>Déplacements efficaces d'un système à l'autre</p>	<p><b>AU BON MOMENT</b></p> <p>Réactions immédiates, transmission de données de capteurs en continu, traitements par lots nocturnes</p>	<p><b>DANS LE BON FORMAT</b></p> <p>Validées, normalisées ou enrichies ; exploitables</p>	<p><b>POUR TOUS LES UTILISATEURS</b></p> <p>Utilisation régie par des règles ; application de la sémantique d'entreprise</p>
---	---	---	--

**DÉFIS**

Voici les principales difficultés citées dans le cadre d'une enquête MarketPulse réalisée en 2014 auprès des DSI :



**TECHNOLOGIES**

<b>ÉCONOMIE</b>	<b>INFORMATIQUE</b>
-----------------	---------------------



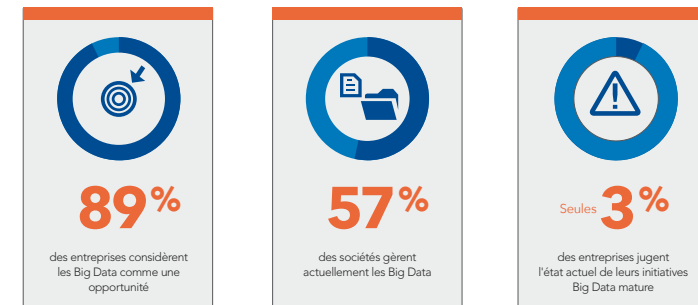
Une stratégie de gestion des données efficace repose sur un large éventail d'outils et de technologies. Pour y parvenir, la collaboration entre les secteurs économique et informatique est essentielle.

SAS Data Management facilite cette collaboration et offre :



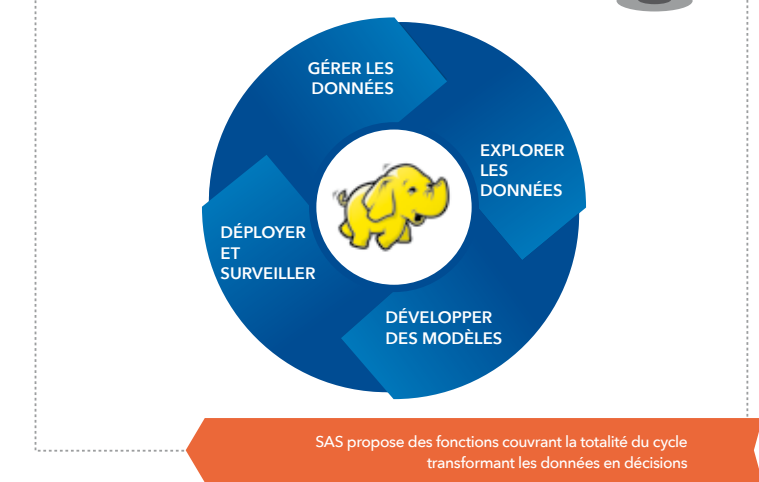
<p><b>ACCÈS AUX DONNÉES</b></p> <p>Mettre en forme, consulter et distribuer les données n'importe où</p>	<p><b>INTÉGRATION DES DONNÉES</b></p> <p>Transformer et compléter le jeu de données en combinant et en interprétant des données issues de différentes sources</p>	<p><b>FÉDÉRATION DES DONNÉES</b></p> <p>Intégrer les informations de manière logique sans avoir à rendre les magasins de données physiques disponibles</p>
<p><b>QUALITÉ DES DONNÉES</b></p> <p>Rendre les données précises, cohérentes et utilisables</p>	<p><b>GOVERNANCE DES DONNÉES</b></p> <p>Veiller à ce que toutes les données stratégiques pertinentes soient gérées de manière formelle</p>	<p><b>GESTION DES DONNÉES MAÎTRESSES</b></p> <p>Créer et gérer une vue unique des données, utilisée dans tous les domaines de l'entreprise</p>

**GESTION DES BIG DATA**



Source : TDWI BP Report: Managing Big Data

La plupart des entreprises gèrent d'importants volumes de données sur des moyens de stockage à bas coût. Elles s'appuient sur de puissants processeurs utilisant des plates-formes Big Data distribuées, comme Hadoop.



**GRÂCE À SAS DATA MANAGEMENT, VOUS POUVEZ :**

- Mettre les données stockées dans Hadoop à la disposition de tout le monde, n'importe où
- Intégrer et utiliser des données analytiques sans compétences supplémentaires
- Travailler dans les environnements traditionnels à l'aide de Hadoop et contrôler la sécurité et les métadonnées de manière centrale
- Réaliser des transformations, des opérations de qualité des données et des modèles de scoring directement dans le cluster Hadoop
- Exécuter des processus SAS « sur cluster » dans Hadoop grâce à des processus intégrés et des technologies en mémoire

**Passez des Big Data aux Big Analytics**

**Jacorien  
Wouters**

RESPONSABLE DU  
SYSTÈME INFORMATIQUE  
DE GESTION DE RÉSEAUX  
CHEZ RIJKSWATERSTAAT


Une vision claire et intégrée  
du réseau routier et de celui  
des voies navigables

## RIJKSWATERSTAAT BÉNÉFICIE D'UNE VISIBILITÉ COMPLÈTE SUR SES PERFORMANCES

Rijkswaterstaat, l'agence exécutive du Ministère de l'infrastructure et de l'environnement néerlandais, est en charge des principaux réseaux routiers et de voies navigables, et du plus important réseau de distribution d'eau des Pays-Bas. Pour rendre des comptes au ministère et à la Chambre des Députés du parlement néerlandais, Rijkswaterstaat doit gérer les processus opérationnels de sa propre organisation interne, mais aussi disposer d'informations précises en temps voulu et y avoir accès en interne et en externe. C'est dans cette optique que l'agence a développé le système informatique de gestion de réseaux NIS.

Rijkswaterstaat a commencé à développer le NIS il y a quelques années. Le système était conçu pour offrir une vision intégrée des performances de l'agence, dont il permettait de contrôler plus étroitement les résultats et offrait une vue d'ensemble. Lors de la procédure d'appel d'offres, Rijkswaterstaat a choisi les solutions de SAS pour leur capacité à prendre en charge l'intégralité du processus, de la source au navigateur. Rijkswaterstaat utilise ainsi SAS® Business Intelligence et SAS® Data Management pour le NIS.

« Le NIS est aujourd'hui l'une des sources d'informations les plus importantes pour la direction de Rijkswaterstaat », affirme Jacorien Wouters, responsable du programme. « Il nous a permis de réunir les deux principaux flux d'informations sur les réseaux de notre organisation : les performances des réseaux et les données relatives aux actifs tels que les routes et les ponts, mais aussi la mer des Wadden, par exemple. Un processus intensif d'intégration de données a eu lieu au préalable ».



**« La visibilité sur les données sous-jacentes nous aide à être plus efficaces et au final, à réduire nos coûts. La méthode de reporting extrêmement claire nous permet également de gagner du temps. Désormais, il n'existe qu'une seule version de la situation »**

Jacorien Wouters

### De meilleures décisions

Lors du lancement du programme NIS en 2004, les données de différentes applications étaient disséminées dans les systèmes informatiques des dix agences régionales de Rijkswaterstaat. Aujourd'hui, le NIS récupère régulièrement les données de plus de 40 systèmes source. L'intérêt du système réside notamment dans la possibilité de combiner les données et de les présenter sous forme de diagrammes et de cartes, ce qui permet d'obtenir une vision claire et rapide des performances de chaque agence et de Rijkswaterstaat dans son ensemble. Les chiffres du NIS possèdent un statut officiel, ce qui est très important en interne, mais aussi en externe, puisque trois fois par an, Rijkswaterstaat transmet un rapport au ministère sur l'état de certains indicateurs de performance, ou PIN. Dans le cadre d'un accord de niveau de service, des engagements ont d'ailleurs été pris pour une période budgétaire de quatre ans.

### Des analyses plus complexes

Outre un contrôle accru, l'accès aux informations a été considérablement simplifié, comme a pu le constater le nombre croissant d'utilisateurs du système NIS chez Rijkswaterstaat. « Des informations qui n'étaient auparavant accessibles qu'à quelques employés depuis une source de données spécifique sont à présent disponibles via le portail NIS pour l'ensemble du personnel de Rijkswaterstaat », explique Jacorien Wouters.

### Une seule version de la situation

« La visibilité sur les données sous-jacentes nous aide à être plus efficaces et au final, à réduire nos coûts », poursuit J. Wouters. « La méthode de reporting extrêmement claire nous permet également de gagner du temps. Aujourd'hui, il n'existe plus qu'une seule version de la situation, et les interminables discussions sur des définitions ou des chiffres appartiennent au passé. Le fait que les informations soient plus facilement accessibles dans le NIS nous permet également de procéder à des ajustements plus rapides et plus précis. Nous éditons un rapport sur les performances trois fois par an et nous découvrons à cette occasion des sujets que nous aurions préféré traiter en temps réel, ce qui est parfaitement possible aujourd'hui ».

### Développements

L'implémentation de SAS a permis à Rijkswaterstaat d'aller encore plus loin en améliorant la qualité des données. L'agence a également commencé à utiliser SAS Visual Analytics. « Comme nous disposons d'une meilleure visibilité sur nos données, notre direction peut prendre des décisions à plus long terme », se félicite Jacorien Wouters. « Nous réalisons constamment des progrès dans la combinaison d'informations, ce qui nous permet de mettre au jour des liens qui n'auraient pas été visibles auparavant ». ■

Le consultant indépendant en gestion des informations Ronald Damhof est l'auteur du modèle de quadrant de données

## « Faire de la **gestion des données** une question d'actualité à l'échelle de l'entreprise »

**Ronald  
Damhof**

CONSULTANT  
INDÉPENDANT EN  
GESTION DE  
L'INFORMATION

Le domaine de la gestion des données regorge de toutes sortes de jargons. La plupart des dirigeants d'entreprise n'ont pas la moindre idée de ce que signifient tous ces mots, sans parler de les utiliser pour évoquer la valeur précise de données particulières, et la manière de les traiter. Pour permettre aux entreprises d'ouvrir une discussion sur les données, Ronald Damhof a développé le modèle de quadrant de données.

Ronald Damhof est consultant indépendant en gestion de l'information auprès d'établissements tels que Ahold, la banque centrale des Pays-Bas, l'administration fiscale néerlandaise, Alliander et des entreprises des secteurs financier et de la santé. Ces structures, qui font un usage intensif de données, prennent chaque jour conscience que la qualité de leur travail est de plus en plus liée à celle de leurs données. Mais comment transformer cette prise de conscience en stratégie efficace ? Une stratégie compréhensible par chaque membre de l'entreprise, du directeur à l'ingénieur informatique ? Ronald Damhof a développé un modèle de quadrant pour faire de la gestion des données une question d'actualité.

**Pousser ou tirer ?**

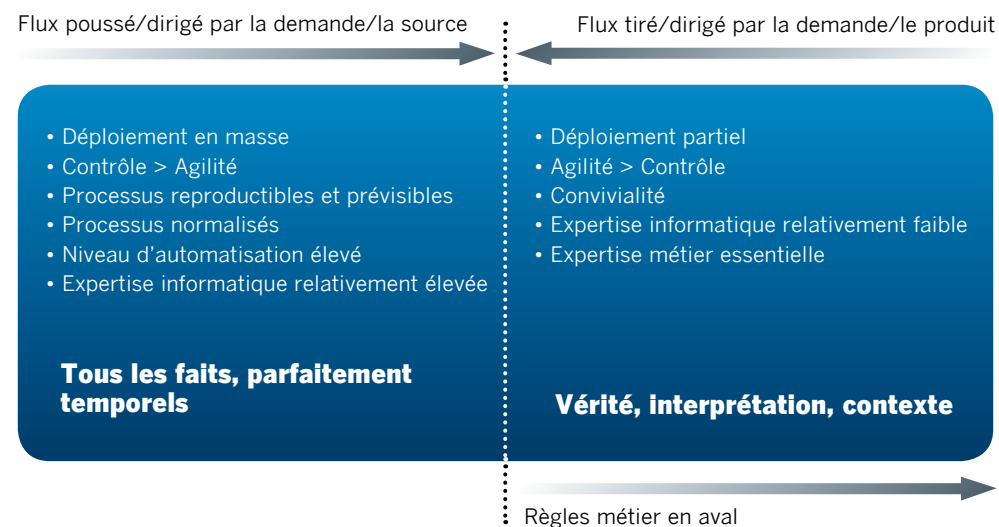
Ronald Damhof commence par expliquer un concept dont chacun d'entre nous a déjà entendu parler : le « pousser-tirer », qui concerne l'ampleur des effets de la demande sur le processus de production. Il prend l'exemple de la construction d'un yacht de luxe, dont le processus ne démarre que lorsque le client passe commande. Le point de découplage se situe alors au début du processus de production. Autre exemple, totalement opposé : la fabrication d'allumettes. Si un client veut des allumettes, il se rend au supermarché et en achète (sauf s'il souhaite des allumettes noires, plus difficiles à trouver). Ici, le point de découplage se situe complètement à la fin du processus de production. La construction d'une voiture, en revanche, implique des pièces standard et des éléments sur mesure. Les clients peuvent en effet demander une couleur particulière, une sellerie en cuir ou des jantes spécifiques. Dans ce cas, le point de découplage se situe quelque part au milieu du processus de production. « De même, dans la production d'un rapport, d'un tableau de bord ou d'un environnement analytique, le point de découplage se situe quelque part dans cette zone du milieu », explique Ronald Damhof.

Le point de découplage divise le processus de production en deux parties : un côté tiré et un côté poussé, également appelés parties dirigée par l'offre et dirigée par la demande. Les systèmes à flux poussé visent à réaliser des économies d'échelle en cas d'augmentation du volume et de la demande, tout en garantissant la qualité du produit et des données associées. À l'inverse, les systèmes à flux tiré sont dirigés par la demande. Chaque type d'utilisateur souhaite manipuler les données de façon à produire « son » produit, sa réalité, sur la base de sa propre expertise et de son contexte.

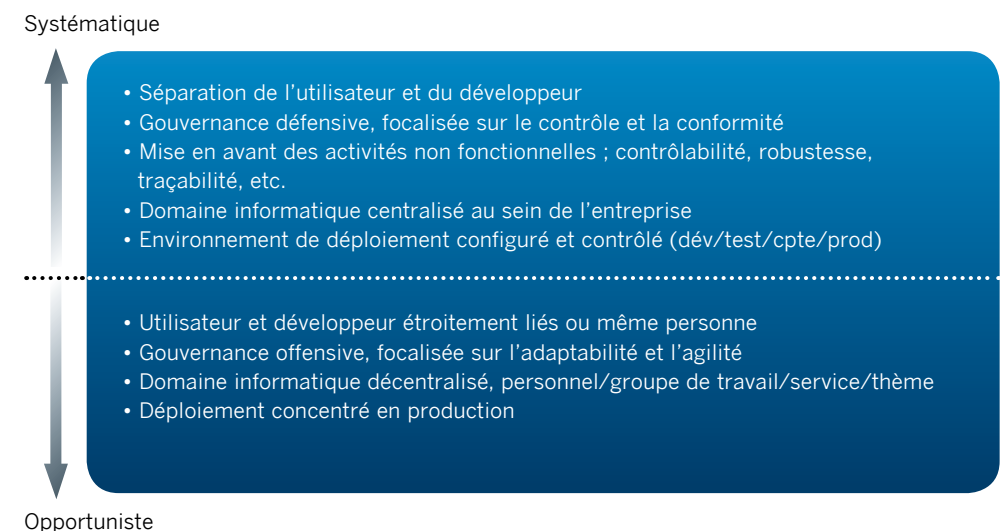
**Développement opportuniste ou systématique ?**

Sur l'axe Y, Ronald Damhof projette la dimension du style de développement. « Ce qui signifie : comment développez-vous un produit d'informations ? Il est ici possible d'agir de manière systématique, auquel cas l'utilisateur et le développeur sont deux personnes distinctes et vous appliquez une gouvernance défensive, axée sur le contrôle et la conformité. Cela met en pratique tout ce que les ingénieurs ont appris pour créer des logiciels sur une base solide. Ce schéma est souvent visible dans les données d'entreprise centralisées, comme les données financières et celles transmises aux autorités de contrôle ». Il est également possible d'adopter un style de développement opportuniste. « Dans ce cas, le développeur et l'utilisateur ne sont souvent qu'une seule et même personne. Prenons l'exemple du statisticien en quête de données innovantes pour produire et expérimenter des modèles analytiques, ou différentes situations où la rapidité de livraison est essentielle. Dans ces deux cas, la gouvernance est souvent offensive, et ce sont la flexibilité et l'adaptabilité qui priment ».

Le point de découplage des données



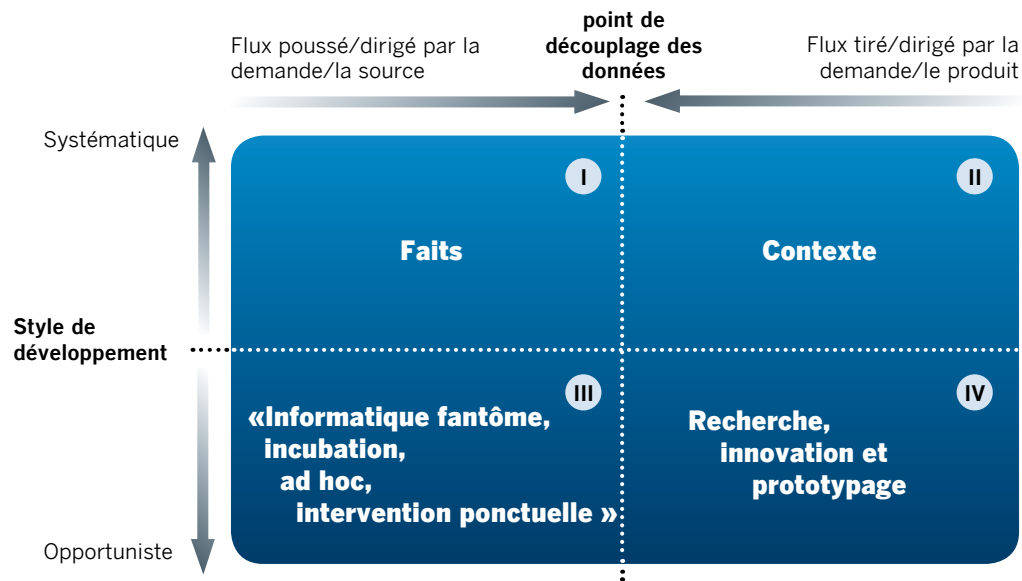
Style de développement



**Modèle de quadrant de données**

La réunion de ces deux dimensions produit le tableau suivant.

Quadrant de déploiement de données



« Les faits réels se trouvent dans le quadrant I », explique Ronald Damhof. « Ces données peuvent être intégralement fournies de manière intelligible aux quadrants II et IV. Les données du quadrant I étant produites par des systèmes et processus hautement normalisés, elles sont entièrement prévisibles et reproductibles ».

Les données du quadrant IV, diagonalement opposées, se caractérisent par l'innovation et le prototypage. « C'est le quadrant dans lequel travaillent les data scientists, qui n'ont en réalité que trois exigences : données, puissance informatique et convivialité du logiciel ». Différentes équipes sont progressivement mises en place à mesure que les laboratoires qui alimentent ces « scientifiques » en données leur donnent carte blanche pour les utiliser à des fins d'expérimentation et d'analyse, dans une optique d'innovation. « Ce type de gestion des données est nécessaire pour avoir des idées et les expérimenter. Lorsqu'un concept fonctionne, il doit remonter du quatrième au deuxième quadrant, car seules la génération et l'analyse systématiques de données permettent des économies d'échelle dans ce domaine. Le concept peut ensuite être généralisé à toute l'entreprise ».

« Je discute souvent avec des data scientists qui tirent de très précieux enseignements d'un environnement de type sandbox », poursuit Ronald Damhof. « Mais ils oublient ou sont incapables de monétiser ces connaissances dans un système de production. Ils ne parviennent pas à les faire passer du quadrant IV au quadrant II. C'est là qu'intervient la gouvernance ». Mais c'est également à ce stade que réside la principale difficulté pour bien des entreprises, phénomène bien connu de



« Les entreprises génèrent des volumes de données toujours plus importants et ne peuvent plus se permettre d'être aussi négligentes que par le passé. Elles doivent à présent s'assurer que la gestion des données et la gouvernance associée sont correctement définies. C'est précisément le rôle du modèle de quadrant de données »

Ronald Damhof

Ronald Damhof. « Si vous expliquez ce modèle aux responsables et les interrogez sur leurs priorités, ils répondent inmanquablement qu'ils doivent commencer par mettre en place des bases solides – le premier quadrant. Mais si vous leur demandez dans quoi ils investissent actuellement, dans quels domaines ils innovent, il s'agit souvent du quatrième quadrant. C'est très bien de s'engager ainsi dans une forme de gestion des données plus expérimentale et exploratoire, mais ce n'est possible que si les bases sont solides. Sinon, cela revient à être équipé de toilettes ultra-modernes qui ne sont pas raccordées au réseau d'assainissement, et c'est le chaos ». Demandez au statisticien moyen ce qui lui prend le plus de temps et il vous répondra : récupérer les données au niveau de qualité adapté – ce qui est l'objectif du quadrant 1. « Seul un data scientist disposant d'un solide logiciel analytique, d'une grande puissance informatique et de données de qualité fera véritablement la différence ».

**Fiabilité et flexibilité**

« Les dirigeants insistent sur la fiabilité et la flexibilité des systèmes, mais ces qualités sont inversement proportionnelles. Un système particulièrement fiable et robuste est moins souple, tandis qu'un système extrêmement flexible implique de revoir ses exigences de fiabilité à la baisse », décrypte Ronald Damhof. « C'est parfaitement clair dans le modèle de quadrant de données. Dans le quadrant I, la fiabilité prévaut sur la flexibilité, et c'est l'inverse dans les quadrants II et IV ». Une utilisation optimale des données exige par conséquent une expertise et des compétences différentes.

**Expertise et compétences**

Les entreprises exigent souvent qu'une seule personne offre l'expertise et les compétences couvrant l'intégralité du quadrant. Cette personne n'existe pas. Les employés du quadrant I ont un profil d'ingénieur. Ce sont des spécialistes de l'informatique et des données, ayant une formation en architecture et modélisation. « Vous remarquerez qu'il ne s'agit pas du profil informatique type. Ces ingénieurs sont capables d'effectuer un développement orienté modèle et comprennent parfaitement les besoins en modélisation conceptuelle et logique ». Un tel niveau d'expertise est très rare. À l'opposé, les quadrants II et IV exigent des spécialistes dans leur domaine respectif, ayant en outre des compétences en informatique décisionnelle et/ou en analyse.

« Dans le monde de l'entreprise, il est souvent question d' »une seule version de la vérité », mais cela n'existe pas. Il existe « une seule version des faits » et plusieurs « vérités ». Après tout, la manière dont chacun interprète les faits varie selon le type d'entreprise, les perspectives, les connaissances et l'expérience ».

Ronald Damhof

**Les faits et la vérité**

Ronald Damhof surnomme aussi le quadrant I « la seule version des faits ». Ces derniers sont mis à la disposition des employés au niveau des quadrants II et IV, et leur permettent d'établir leur propre vérité. Dans la mesure où les mêmes faits servent à créer plusieurs vérités dans la moitié droite du modèle – selon le contexte et les connaissances de l'utilisateur des données – Ronald Damhof appelle cette moitié « les multiples versions de la vérité ». Vous devez garder à l'esprit que la « vérité » change souvent au fil du temps. « Dans les entreprises, il est souvent question d' »une seule version de la vérité », mais cela n'existe pas. Après tout, la manière dont chacun interprète des faits particuliers varie selon le contexte, les perspectives, les connaissances et l'expérience ».

**Quadrant III**

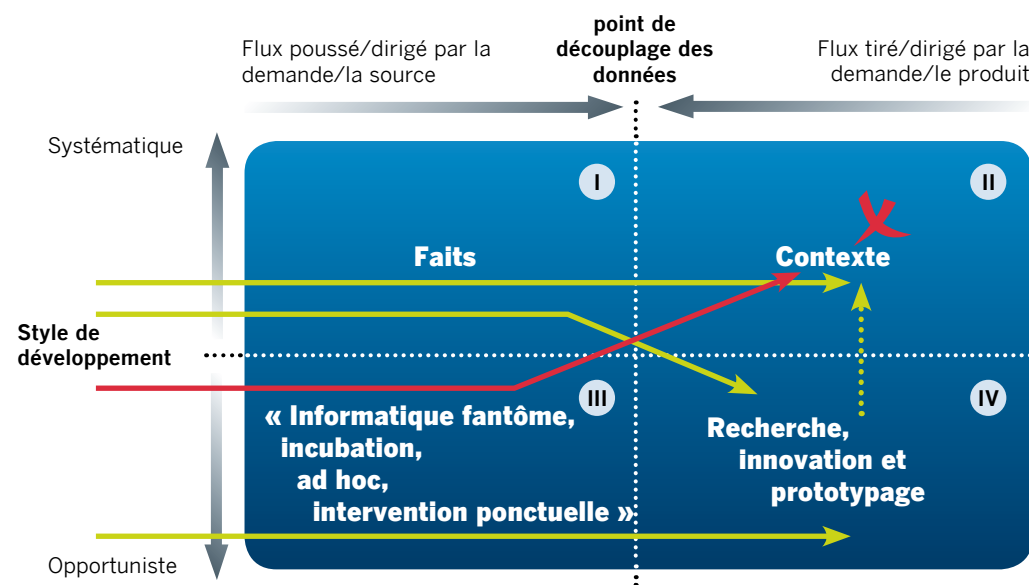
Le quadrant III, rarement évoqué jusqu'à présent, est d'une importance capitale. C'est le quadrant des sources de données qui ne font l'objet d'aucune gouvernance, comme un téléchargement ad hoc auprès d'un fournisseur de données ouvertes, une liste Excel dont vous souhaitez vous servir ou une série de données de vérification reçues sur un CD. « Vous pouvez même associer sans problème des données administrées du quadrant I à votre propre jeu de données dans le quadrant IV », affirme Ronald Damhof.

**La traversée des quadrants**

Pour valoriser vos données, vous pouvez effectuer différents déplacements dans le modèle. Vous pouvez ainsi passer d'une gestion des données factuelle à un modèle où le contexte est également important (C'est-à-dire des quadrants I à II). « C'est vraiment le parcours classique de « déblocage des données et d'élaboration d'un produit d'informations » », analyse Ronald Damhof. Il est pourtant souvent inefficace, car ce processus suppose des exigences et des attentes connues de l'utilisateur. « Or l'utilisateur ne dispose pas réellement de ces connaissances à l'avance ». La plupart des entreprises adoptent une démarche axée sur l'agilité, en passant par exemple des quadrants I à II via le IV. Demandez aux employés du quadrant IV d'élaborer un produit d'informations de manière itérative à partir des données du quadrant I ou III. Remontez ensuite le produit au quadrant II, uniquement s'il est important de le gérer.

Il est également possible de passer du quadrant III au quadrant IV. « Vous disposez de jeux de données propres et vous souhaitez faire des essais ? Parfait », note Ronald Damhof. Le seul mouvement totalement proscrit consiste à passer du quadrant III au quadrant II. « Car dans ce cas, vous utilisez des données dont vous n'êtes pas entièrement sûr, puisqu'elles n'ont pas fait l'objet d'une gouvernance éclairée, contrairement aux prescriptions. Un exemple à éviter à tout prix : une déclaration de conformité destinée à un organisme de contrôle, que vous souhaitez produire à partir de données non soumises à une gouvernance ».

Comment produire, variantes des process



**Faire de la gestion des données une question d'actualité**

Dans le cadre de son travail quotidien, Ronald Damhof constate que son modèle aide les entreprises à parler de gestion des données. « Chez mon client actuel, la Banque des Pays-Bas, j'entends régulièrement des phrases telles que « Je veux faire passer ce produit de données du quadrant IV au quadrant II », « Nous devons tout d'abord placer les données dans le quadrant I, mais les données de ce quadrant sont sous la responsabilité de l'émetteur » ou « J'ai besoin d'espace pour stocker provisoirement les données dans le quadrant III ». Tout le monde comprend ce dont il s'agit, ce qui est nouveau. L'établissement n'avait jamais envisagé les données sous cet angle auparavant, et cela peut vraiment s'appliquer à toutes les sociétés qui font un usage intensif de données. Les entreprises parlent depuis longtemps des données comme d'un « actif » mais dans les faits, elles les gèrent de façon très déstructurée. Résultat : elles ne monétisent jamais cet actif. Les entreprises génèrent des volumes de données toujours plus importants et ne peuvent plus se permettre d'être aussi négligentes que par le passé. Elles doivent à présent s'assurer que la gestion des données est correctement définie. C'est précisément le rôle du modèle de quadrant de données ». ■



**Caroline Denil**

CHEF DE PROJET  
SPF ECONOMIE

**Vincent  
Vanesse**

ANALYSTE MÉTIER  
SPF ECONOMIE

Un accès immédiat à des données aisément compréhensibles

## AMÉLIORER L'INFORMATION SUR LA CONJONCTURE ÉCONOMIQUE DE LA BELGIQUE

Le Service Public Fédéral (SPF) Economie belge s'est engagé à créer une présentation plus efficace et transparente de la situation économique de la Belgique à l'adresse du grand public, des statisticiens et des étudiants, entre autres. En partenariat avec SAS, il a constitué un portail web unique offrant un accès direct aux principaux indicateurs conjoncturels belges.

Tous les indicateurs sont présentés sous forme de graphiques pour une meilleure compréhension et sont entièrement personnalisables pour permettre aux utilisateurs de consulter immédiatement ceux qui les intéressent. Le portail a non seulement créé un environnement statistique convivial, mais aussi ouvert de nouvelles opportunités économiques avec d'autres directions générales du gouvernement fédéral belge.

### La dissémination des informations est synonyme de recherches fastidieuses

L'une des principales missions du SPF Economie est la génération et la publication de statistiques et de chiffres représentatifs de la conjoncture belge. Il y a peu de temps encore, ces informations étaient accessibles via différentes sources : Statbel, be.Stat, Belgostat et la Banque Nationale de Belgique. Dans ce type de situation, il est difficile pour les étudiants, chercheurs, journalistes et autres nombreux utilisateurs de trouver les informations dont ils ont besoin pour répondre à leurs questions et en tirer des conclusions précises. Le SPF Economie a donc lancé un projet pour améliorer la convivialité des données économiques.

### La collaboration entre services améliore les statistiques

Le premier objectif du projet était la valorisation des informations. La réunion des analystes métier et statisticiens du SPF Economie fut un processus intense, mais réellement indispensable. Elle a conduit au développement de graphiques illustrant des informations économiques, mais aussi de métadonnées que les utilisateurs peuvent consulter pour mieux comprendre ces informations. « Une vingtaine de graphiques ont ainsi été choisis, puis répartis en huit catégories, dont l'énergie, le produit intérieur brut et l'indice des prix à la consommation », explique Vincent Vanesse, analyste métier au SPF Economie.

« Il est aujourd'hui facile de trouver des informations sur la conjoncture économique de la Belgique. L'outil Ecozoom du site web du SPF Economie permet d'accéder immédiatement aux vingt principaux graphiques économiques »

Caroline Denil

### Un portail unique pour tous les indicateurs économiques

Ensuite, le SPF Economie s'est associé à SAS pour rendre les indicateurs économiques accessibles via un outil convivial. « Nous travaillons avec SAS depuis un certain temps, et ses compétences n'ont plus de secret pour nous. Pour ce projet particulier, les fonctions statistiques, la robustesse et l'extensibilité exceptionnelles des solutions SAS ont rendu notre choix évident », affirme Caroline Denil, chef de projet au SPF Economie.

Cette collaboration s'est traduite par le lancement d'un portail web unique (Ecozoom), où les utilisateurs de tous horizons peuvent trouver les indicateurs économiques dont ils ont besoin en quelques clics. « Il est aujourd'hui facile de trouver des informations sur la situation économique de la Belgique »,

« Les utilisateurs peuvent choisir les indicateurs qui les intéressent le plus et enregistrer ces informations. Par la suite, chaque fois qu'ils consulteront l'outil, les données souhaitées apparaîtront immédiatement »

Vincent Vanesse

observe Caroline Denil. « L'outil Ecozoom du site web du SPF Economie permet d'accéder immédiatement aux vingt principales courbes économiques. Ceux qui ont besoin d'informations plus détaillées peuvent toujours accéder en un clic aux sites web Statbel, be.Stat et Belgostat traditionnels ».

### La visualisation facilite la compréhension

Le portail en ligne présente les indicateurs conjoncturels sous forme de graphiques qui favorisent l'interprétation rapide et précise des informations. Caroline Denil souligne qu'il est bien plus facile de déduire des tendances d'un graphique que d'utiliser un tableau ou une longue suite de chiffres. Par ailleurs, l'outil permet de visualiser quatre graphiques en simultané, ce qui facilite les comparaisons entre différents types de données – pour mesurer les effets du nombre de défaillances d'entreprise sur le taux de chômage, par exemple.

« Le vieil adage selon lequel "une image vaut mieux qu'un long discours" s'applique parfaitement au SPF Economie », confirme Caroline Denil. « Nos graphiques sont souvent bien plus parlants qu'un long texte ou une série de tableaux. Dans notre cas précis, ils aident les utilisateurs à évaluer plus facilement et précisément la conjoncture de la Belgique ».

### La personnalisation augmente la convivialité

Vincent Vanesse s'empresse de préciser que les quatre graphiques illustrant la page d'accueil Ecozoom sont entièrement personnalisables. « Les utilisateurs peuvent sélectionner les indicateurs qui les intéressent le plus et enregistrer ces informations. Par la suite, chaque fois qu'ils consulteront l'outil, les données souhaitées apparaîtront immédiatement ».

### L'ouverture de nouvelles opportunités

Même si l'outil Ecozoom a considérablement augmenté la convivialité des données économiques, le SPF Economie envisage déjà de nouvelles possibilités pour le rendre encore plus ergonomique. « Nous réalisons actuellement des tests de géovisualisation pour visualiser les données de régions belges précises », note Caroline Denil. « Nous prévoyons en outre de rendre l'outil accessible sur les smartphones et les tablettes pour une utilisation mobile ».

L'outil Ecozoom pourrait bien ouvrir de nouvelles opportunités économiques. « Il a éveillé l'intérêt d'autres directions générales, y compris au plus haut niveau, ce qui pourrait intensifier la collaboration entre les différents Services Publics Fédéraux, voire donner naissance à un nouveau type de service », conclut Caroline Denil. ■



**DSM lance une solution de gestion des données maîtresses, aboutissement de ses succès en matière de qualité des données**

# DSM PREND LE CONTRÔLE DES DONNÉES MAÎTRESSES

DSM est convaincu de la valeur de données de bonne qualité. La multinationale scientifique, présente dans les domaines de la santé, de la nutrition et des matériaux, a déjà mis en œuvre avec succès une stratégie de qualité des données, dont la solution SAS® Master Data Management (MDM) est aujourd'hui l'aboutissement.

La gestion des données maîtresses (MDM) est une méthode de gestion centralisée des données stratégiques destinées à une utilisation décentralisée. Elle vise en particulier les erreurs et divergences contenues dans les données maîtresses – noms de clients, types de matières et fournisseurs, entre autres informations utilisées dans l'entreprise, tous systèmes informatiques et divisions confondus. La cohérence de ces données stratégiques joue un rôle déterminant dans l'efficacité des opérations. Grâce à la gestion des données maîtresses, DSM prend le contrôle des systèmes ERP absorbés dans le cadre des acquisitions majeures opérées ces dernières années.

### Des mines d'État aux produits chimiques et à l'environnement

« Nous employons 25 000 personnes, chiffre qui s'est considérablement accru ces cinq dernières années en raison d'acquisitions », explique Bart Geurts, directeur du pôle Master Data Shared Services de DSM. Bart Geurts cite ainsi Roche Vitamins, l'une des principales acquisitions du groupe, réalisée en 2003. DSM est désormais le premier fabricant mondial de vitamines, ce qui implique des besoins différents en matière de données. « La qualité des données est extrêmement importante pour la sécurité alimentaire et la santé, mais aussi pour les arômes. Elle l'est moins pour les produits chimiques de base. » Bart Geurts fait ici allusion aux origines de DSM dans les mines d'État des Pays-Bas. « Les vieilles entreprises savent qu'elles doivent se réinventer pour survivre. »

DSM s'est réinventé plusieurs fois, de l'exploitation minière à la pétrochimie et ces dernières années, des produits de chimie fine aux sciences de la vie et des matériaux. Dans sa forme actuelle, DSM se concentre sur les marchés émergents, ainsi que sur le climat et l'énergie. Bart Geurts évoque ainsi des matériaux plus légers, destinés à remplacer l'acier des voitures pour en réduire le poids et les rendre plus économiques. Le groupe développe également des produits fabriqués à partir d'enzymes plutôt qu'à base de pétrole. Avec des activités aussi variées, tournées vers des marchés différents, les besoins de l'entreprise en matière de données sont multiples.

« La qualité des données est extrêmement importante, que ce soit pour la santé ou la sécurité »

Bart Geurts

### Un organigramme plus complet

Les nombreuses entités impliquées dans cette mutation ont non seulement apporté de nouveaux collaborateurs et activités, mais aussi de nouveaux systèmes informatiques. Bart Geurts explique que de nombreux systèmes ERP figuraient parmi ces derniers. Mais les différents systèmes informatiques de la nouvelle organisation contenaient des erreurs. Il ne s'agissait pas d'erreurs graves, mais plutôt de divergences que l'utilisation conjointe des données et des systèmes a permis de révéler.

Bart Geurts cite en exemple une fête réunissant le personnel à l'occasion du lancement du nouveau logo de la société. Lors de l'envoi des invitations, 800 personnes ont été « oubliées », ce qui était dû à un organigramme incomplet de l'environnement RH. Il signale également d'autres incidents : une

**Bart Geurts**

DIRECTEUR DU PÔLE  
MASTER DATA SHARED  
SERVICES DSM

contamination des données fournisseur, par exemple. Un fournisseur peut en effet avoir des appellations différentes d'un pays à l'autre, au risque d'être considéré comme plusieurs entreprises dans les différents systèmes utilisés dans une multinationale.

### Lier la gestion des données maîtresses aux processus métier

Fort de ses expériences précédentes et de ses résultats en matière de qualité des données, DSM est passé à une approche MDM centralisée. Selon Bart Geurts, la qualité des données de l'entreprise est suffisante pour les transactions réalisées au sein d'un silo particulier – pays ou division. « Mais dès que les activités concernent plusieurs divisions, des problèmes risquent de se poser. » D'après lui, les principaux fournisseurs de systèmes ERP proposent des solutions de gestion des données maîtresses, mais attachent trop d'importance aux silos individuels, ce qui explique pourquoi DSM a choisi celle de SAS.

Bart Geurts insiste sur l'importance du lien entre la gestion des données maîtresses et les processus métier. Il en souligne les avantages pour l'ensemble de l'entreprise et chacune de ses divisions, qui fonctionnent efficacement au sein de leurs silos respectifs. Les questions essentielles concernent le responsable du processus MDM au sein de l'entreprise, le rôle attribué à chacun et les indicateurs de performance clés utilisés. À titre d'exemple, l'un de ces indicateurs peut mesurer le délai de traitement, de livraison et de facturation des commandes client.

### Voir grand mais commencer petit

Selon B. Geurts, le plus facile fut de mettre en place le processus MDM et de résoudre les problèmes correspondants. Il qualifie d'ailleurs cette étape de « promenade de santé ». Puis vint la phase d'implémentation, avec le choix délibéré d'un démarrage à relativement petite échelle. « Nous avons mené un projet pilote au niveau du service Achats, selon le principe « voir grand mais commencer petit ». Il convient toutefois de replacer le terme « petit » dans son contexte. À l'international, DSM possède six services Achats « front-office » et deux « back-office ». Dans ce projet pilote à petite échelle, les incohérences observées dans les données fournisseur ont été traitées en premier. Ces données, qui comprenaient des doublons, ont notamment été nettoyées au moyen d'algorithmes en plusieurs langues dans la solution SAS MDM. « La complexité réside dans les détails », affirme Bart Geurts d'expérience.

### Quelles données sont stratégiques pour votre entreprise?

Outre la « décontamination » des données fournisseur, des procédures ont été mises en place pour traiter d'autres éléments de données maîtresses. Nous avons demandé aux agences concernées de signaler les données stratégiques pour leur activité. « Car nous ne pouvions pas analyser toutes les données de l'entreprise. » L'opération serait trop vaste et solliciterait beaucoup trop les systèmes. L'identification des données stratégiques a permis de faire le lien entre l'initiative MDM et les services impliqués. Ils ont finalement sélectionné eux-mêmes les données primordiales pour leurs processus. Une telle sélection est nécessaire face à l'ampleur des données maîtresses. Pour DSM, ces dernières englobent tout ce qui sous-tend les processus et les transactions. À première vue, n'importe quelle erreur peut être une source d'inefficacité, mais sa portée réelle dépend du type de données. « Si le numéro de téléphone du commercial d'un fournisseur est incorrect, vous pouvez toujours lui adresser un e-mail », explique Bart Geurts. En revanche, il n'existe aucune alternative en cas d'adresse ou de numéro de compte bancaire erroné.



### Éviter les erreurs

Des règles ont été établies sur la base des données désignées comme critiques par les services. L'opération a pris environ six mois, suivie de la mise en œuvre, finalisée en l'espace de trois semaines. La gestion des données maîtresses a notamment permis à DSM d'éviter les erreurs. Bart Geurts cite l'exemple d'une commande saisie au nom d'un service incorrect. Le groupe s'efforce également d'améliorer la saisie des données fournisseur, les erreurs étant fréquentes lors de la recherche d'un fournisseur existant ou de l'enregistrement d'un nouveau venu. Si la recherche échoue, une nouvelle entrée est créée ; or il s'agit essentiellement d'un doublon. Un algorithme est désormais lié à la saisie des données pour effectuer une vérification et demander à l'utilisateur : « Ce fournisseur est-il celui que vous recherchez ? »

### La gestion des données maîtresses est un processus continu

Les avantages de la gestion des données maîtresses cités précédemment sont liés à des questions internes, comme la vue d'ensemble des effectifs, la déduplication des données fournisseur et la prévention des erreurs. Mais la gestion des données maîtresses offre également à DSM des avantages extérieurs. « Supposons qu'un produit ou un matériau présente une erreur. Nous voulons connaître immédiatement le produit concerné. » Dans ce type de cas, la rapidité est essentielle. Il est également important de continuer à vérifier les données après la mise en œuvre initiale de la gestion des données maîtresses. « Maintenez les contrôles ! Sinon, vous aurez de nouveaux problèmes deux ou trois mois plus tard », prévient Bart Geurts. La gestion des données maîtresses est en effet un processus continu, qui reste actif pour prévenir de nouvelles erreurs qu'il faudrait sans cela résoudre ultérieurement. « C'est peu souhaitable, car cela perturberait votre processus métier. » Veiller à ce que toutes les personnes concernées dans l'entreprise comprennent le rôle déterminant de la gestion des données maîtresses est un gage de réussite. ■



**Comment la gouvernance des données peut favoriser le data mining du futur**

# QUI GOUVERNE VOS DONNÉES ?

Le monde des données dépasse largement le cadre de leur simple collecte et utilisation à des fins lucratives. Seules des données de qualité peuvent rapporter de l'argent, et cela suppose qu'elles soient bien gérées. Et il ne s'agit pas de gestion informatique classique, mais plutôt de gouvernance des données. Avez-vous mis en place une stratégie de gouvernance des données? Qui gouverne vos données?

Aux yeux du profane, la qualité et la gouvernance des données reviennent au même. C'est pourtant loin d'être le cas, même si ces deux disciplines sont étroitement liées. « La qualité des données implique la mise en œuvre de règles, tandis que la gouvernance va beaucoup plus loin », explique Bas Dudink, expert en gestion des données chez SAS Pays-Bas. « À qui incombe la qualité des données ? Quelles responsabilités implique-t-elle ? Quels accords ont été conclus dans ce domaine ? »

## D'où proviennent vos données ?

La qualité des données concerne l'exactitude des adresses postales et des bases de données, par exemple. Pour garantir une amélioration durable dans ce domaine, vous devrez passer des accords et les appliquer. La qualité des données peut à ce titre être mise en œuvre comme une partie de la gouvernance, mais l'une et l'autre ne sont pas inextricablement liées. La gouvernance des données peut venir de plusieurs directions.

Elle peut être motivée par une augmentation des besoins de l'entreprise, ou imposée par la législation et les réglementations. Bas Dudink cite à titre d'exemple les accords de Bâle et les réglementations qui en ont découlé pour les établissements financiers. Les banques sont aujourd'hui tenues de répondre à la question suivante : « D'où proviennent vos données ? » En pratique, les usines – qui appliquent de gré ou de force certaines normes concernant les matériaux qu'elles utilisent – sont soumises aux mêmes règles.

## Métadonnées, formats de fichier et normes

La gouvernance ne s'arrête pas à l'origine des données. Elle couvre également les métadonnées et les normes applicables aux formats dans lesquels les données sont fournies, traitées et transmises à d'autres organisations – partenaires extérieurs, services internes, futurs clients, etc. Elle pourrait en outre s'étendre à des applications d'information jusqu'à présent inconnues ou inédites.

La gouvernance dépasse ainsi le cadre des services et processus métier. En général, la gestion des données est aujourd'hui encore cantonnée à une activité ou division particulière. L'utilisation des données dépend de la conjoncture. « Les processus quotidiens essentiels fonctionnent globalement bien », explique Bas Dudink. Des changements doivent toutefois être apportés pour intégrer de nouvelles activités ou fonctions d'entreprise, comme la gestion des risques.

## Des conseils en gestion plutôt que des technologies

Une bonne gouvernance des données traite principalement les problèmes non informatiques : processus métier, procédures organisationnelles, accords collectifs et mise en application de ces derniers. Elle implique également un volet « sécurité » : qui a accès à quelles données et quelles sont les mesures de protection ? « Cela relève davantage du conseil en gestion, puisqu'il s'agit d'élaborer des procédures », précise l'expert SAS. Selon lui, l'aspect technologique ne dépasse pas 10 à 20 % du total d'un projet de gouvernance des données.

La question clé concerne la manière dont l'entreprise traite ses données (pour gérer les risques, par exemple). Cette question doit être résolue à l'échelle de l'entreprise, et non au niveau d'une activité



**« Si vous investissez dans des améliorations en amont, vous récupérerez plus que votre mise en aval. Mais si c'est une évidence, pourquoi construire des entrepôts de données aussi complexes que des stations d'épuration, plutôt que de purifier l'eau à la source? »**

Bas Dudink

ou d'une division unique. Une gouvernance des données efficace tient compte de l'avenir, afin que les futurs clients puissent être aisément satisfaits au moyen de données cohérentes – pour de nouveaux partenariats, de nouvelles activités métier, etc.

Il paraîtrait logique de lier cet aspect à la qualité des données, ce qui pourrait offrir des avantages à plus long terme. Toutefois, se contenter d'améliorer la qualité des données sans impliquer la gouvernance peut réduire le projet à des avancées secondaires à court terme. Il se résumera alors à une opération dépourvue d'effets durables ou à répéter indéfiniment pour obtenir un résultat. C'est ce type de situation qu'un gouverneur des données permet d'éviter.

### Emprunter des raccourcis pour parvenir aux indicateurs de performance clés

Si un employé, un responsable ou un service est chargé d'atteindre certains objectifs, c'est bien évidemment sur ces derniers qu'il se concentrera. Si ces objectifs impliquent la responsabilité d'une activité, mais aucune traçabilité, cette activité ne sera certainement pas prioritaire en cas de ressources occupées ou limitées. De telles décisions commerciales et professionnelles, aussi logiques soient-elles, ne sont pas toujours bénéfiques pour l'entreprise.

Prenons l'exemple d'un centre d'appels. Son indicateur de performance clé (KPI) était le temps d'attente des clients. Pour maintenir des scores élevés lors de la période particulièrement chargée de Noël, le centre a décidé d'ignorer plusieurs champs du système CRM, « en attendant le retour à la normale ». S'agit-il d'un raccourci astucieux pour obtenir des résultats, ou d'une attitude négligente vis-à-vis des processus métier de l'entreprise?

Si ce comportement est sans conséquences, il risque tôt ou tard de devenir permanent. Après tout, il permet de gagner du temps et améliore les performances du service concerné et de son responsable. Même si ce scénario pessimiste ne se réalise pas, le mal est déjà fait. En effet, qui vérifiera s'il manque des données et éventuellement, les saisira par la suite ? Les lacunes sont bel et bien là ; il manque des pépites dans le coffre au trésor des données. Une gouvernance des données efficace aurait empêché l'entreprise de commettre cette erreur dans son propre centre d'appels.

### Corriger les erreurs en amont

La procédure de facturation et de règlement d'un établissement de santé est une autre illustration pratique. Le délai entre la réception des factures et leur règlement était long et augmentait constamment. Une enquête en a révélé la cause : les factures contenaient des transactions erronées. L'assureur santé a signalé ces erreurs et l'établissement a dû s'atteler à une lourde tâche de rectification. Il a fallu vérifier chaque facture, ce qui a considérablement retardé le paiement.

L'établissement a décidé de s'attaquer au problème en amont plutôt que de reprendre les factures erronées en aval, après qu'elles ont été générées. Aujourd'hui, chaque procédure médicale liée à un patient est soumise à plusieurs tests de qualité des données, et le professionnel de santé qui saisit les données est directement informé en cas d'erreur, ce qui lui permet de réagir immédiatement. Résultat : les délais de paiement ont été considérablement écourtés. Les rapports et analyses de gestion et les relations avec l'assureur santé ont en outre été améliorés.

### Récupérer davantage que sa mise

Le principe est universel et d'une simplicité enfantine : si vous investissez dans des améliorations en amont, vous récupérerez plus que votre mise en aval. Mais si ce principe est évident, pourquoi est-il si rarement appliqué à la gestion des données ? Pourquoi construire des entrepôts de données aussi complexes que des stations d'épuration, plutôt que de purifier l'eau à la source ?

La gouvernance des données permet d'éviter l'apparition de brèches dans le coffre au trésor des données et la fastidieuse tâche de correction des erreurs qui en découle. Mais comment la mettre en œuvre ? Selon Bas Dudink, la procédure est comparable à un numéro de claquettes. Le premier pas concerne généralement la qualité des données, à l'origine du problème. Mais avant même ce premier pas, il convient de sensibiliser l'ensemble de l'entreprise. Néanmoins, même s'il y a prise de conscience et si les données sont de meilleure qualité, vous avez besoin d'une vision au sens large. « La qualité des données est un processus pas-à-pas », poursuit le consultant de SAS. Pour la gouvernance des données, qui ne concerne pas « simplement » l'administration informatique et la gestion des données, il est en revanche conseillé d'adopter une approche globale.

### L'appréciation de la valeur

La gouvernance des données est donc difficile à mettre en œuvre. « C'est un processus complexe », ajoute Bas Dudink. La théorie s'oppose à la pratique. En théorie, les données sont toujours importantes, tant en termes de qualité que de contrôle. En pratique, cette importance ne trouve pas toujours un écho dans l'entreprise. Le « gouverneur » des données et la nature de ces dernières ne sont pas systématiquement identifiés. Ce problème concerne non seulement la prise en charge des données, mais aussi leur valeur.

Autrement dit : la valeur des données est-elle reconnue à temps ? « Imaginez que votre toit présente une fuite alors que le soleil brille », illustre Bas Dudink. La fuite devient un problème lorsqu'il commence à pleuvoir, mais il est alors trop tard. La gouvernance des données est un sujet complexe, une opération à l'échelle de l'entreprise qui dépasse le simple cadre informatique et couvre plusieurs domaines – opérations, culture d'entreprise, gestion des ressources humaines, etc. Mais à terme, une bonne gouvernance des données offre d'énormes avantages dans la société orientée données actuelle. ■



# LA NOUVELLE DSI

**Jill Dyché**

VICE-PRÉSIDENTE  
SAS BEST PRACTICES

Un jour, quelqu'un a classé les individus en deux catégories : ceux qui aiment classer les gens selon deux catégories, et ceux qui ne le font pas. Je faisais alors partie de ceux qui voient deux sortes de dirigeants : ceux focalisés sur leur métier et les autres sur la technologie.

Je remarque que d'autres ont la même tendance. Que je parle aux clients d'analyse, de gestion de la relation client, de données ou de numérique, la question revient toujours : « Qui doit porter le projet ? Les métiers ou l'informatique ? »

La question de l'appartenance s'immisce dans le quotidien des entreprises du monde entier. Chacune semble avoir les yeux rivés sur les autres – de quel domaine cela relève-t-il ? Qui doit en assurer la gestion ? Qui doit s'en attribuer le mérite ? Qui doit le financer ?

**Mais en observant les entreprises ayant surmonté le traditionnel clivage entre métier et DSI, j'ai relevé trois points communs :**

- » Les dirigeants de ces entreprises ont pris conscience qu'il ne suffisait pas de nommer des personnes ou de modifier des structures organisationnelles. De nouveaux modes de gestion des entreprises sont essentiels et de nouveaux processus doivent être mis en pratique pour assurer l'adoption des changements.



- » Ces entreprises ont créé leur culture en interne, en travaillant sur leur propre structure de manière descendante et ascendante, en veillant à ce que les nouveaux processus et règles de fonctionnement soient suffisamment novateurs pour être compétitifs, sans être disruptifs au risque de favoriser l'inertie, voire le sabotage.
- » Les dirigeants de ces entreprises n'ont pas engagé ces changements dans leur propre intérêt. Ils ont plutôt tenu – et tiennent encore – compte du fait que la transformation numérique impose des approches innovantes à des fonctions qui existent depuis longtemps.

Je me suis inspirée du rôle déterminant de l'économie numérique et de l'innovation comme éléments moteurs du changement de l'IT, pour rédiger un rapport sur les pratiques mises en place par les dirigeants les plus clairvoyants, non seulement pour transformer la DSI, mais aussi pour tirer pleinement parti de la technologie afin d'en tirer des avantages business. « The New IT: How Business Leaders are Enabling Strategy in the Digital Age » (La nouvelle DSI : comment les décideurs actionnent la stratégie à l'ère du numérique) donne ainsi la parole aux acteurs de ce changement qui nous font part de leur expérience.

« L'informatique ne doit plus être opposée aux métiers, mais valorisée par les métiers »

**Dans cet ouvrage, j'évoque ce que j'ai appris auprès de ces dirigeants :**

- » Si votre service informatique n'a que deux vitesses, vous vous exposez à de graves problèmes.
- » La question « Quel type de PDG êtes-vous ? » est hors de propos. La vraie question est la suivante : « Quel type d'entreprise dirigez-vous et à quoi devrait-elle ressembler ? »
- » Réunir tout le monde dans une pièce n'est plus un gage de collaboration. (En fait, c'est dangereux.)
- » La stratégie d'entreprise et la stratégie informatique doivent être alignées.
- » La hiérarchie est en passe d'être remplacée par l'holocratie et l'uniformité par la diversité.
- » L'innovation ne doit pas être l'apanage d'une équipe de choc friande de sushis, se balançant dans des fauteuils ergonomiques et installée dans des locaux séparés. Chacun doit être encouragé à innover !
- » Ceux qui parlent de numérique sont plus nombreux que ceux qui œuvrent en sa faveur. À part vous, peut-être, s'il est possible de limiter la diffusion numérique.
- » Nul besoin d'être à la Silicon Valley pour prendre part à la révolution. Mais vous n'en avez probablement pas envie !

Les dirigeants dépeints dans « The New IT » – dont ceux de Medtronic, Union Bank, Men's Wearhouse, Swedish Health, Principal Financial et Brooks Brothers, pour n'en citer que quelques-uns – démontrent qu'il n'est plus question de clivage entre métier et service informatique. Il s'agit plutôt des métiers qui boostent l'informatique. Et inversement. ■

Une compagnie d'assurance-crédit intègre des ratios de qualité dans l'évaluation des risques

## CRÉDITO Y CAUCIÓN AJOUTE LA QUALITÉ DES DONNÉES À SON MODÈLE DE GESTION

Savoir déterminer le risque d'insolvabilité est vital pour les compagnies d'assurance-crédit. Qu'en est-il si elles ne disposent pas d'informations précises ? Chez Crédito y Caución, la qualité des données trace la voie pour des stratégies à long-terme.

Depuis sa création en 1929, Crédito y Caución est la première compagnie d'assurance-crédit espagnole au niveau national et à l'export. Avec une part de marché proche de 60 % en Espagne, la compagnie contribue depuis plus de 80 ans à la croissance des entreprises, qu'elle protège contre les risques de paiement inhérents à la vente à crédit de biens et services. Depuis 2008, Crédito y Caución est l'opérateur du groupe Atradius en Espagne, au Portugal et au Brésil.

Les clients couverts par les polices d'assurance Crédito y Caución ont la garantie que les factures émises dans le cadre de leurs opérations commerciales seront payées. Le système d'analyse des risques d'entreprise proposé par Crédito y Caución traite plus de 100 millions de dossiers régulièrement mis à jour. Il contrôle en permanence la solvabilité du portefeuille clients des assurés. Par ailleurs, les différents systèmes de la compagnie étudient chaque jour plus de 10 000 opérations de crédit, auxquelles ils appliquent une limite de solvabilité qui les engage vis-à-vis des clients. Pour déterminer ce risque, Crédito y Caución a besoin d'informations complètes et précises sur les clients de ses clients. L'efficacité du service offert par l'entreprise dépend en grande partie de la qualité des données contenues dans ces informations.

### S'adapter aux nouvelles réglementations

Comme tous les assureurs européens, Crédito y Caución doit se conformer au système de surveillance financière axé sur les risques du secteur de l'assurance. Ce cadre comprend la directive Solvabilité II, avec son lot de normes techniques et autres mesures de mise en œuvre. Outre des conditions financières, cette directive prévoit des exigences de qualité des données traitées par les compagnies d'assurance. L'exactitude des données n'est plus une option, mais une obligation. La qualité des données est essentielle pour la prise de décision et pour certifier la conformité aux exigences du nouveau cadre réglementaire.

Crédito y Caución s'est rapproché de la conformité Solvabilité II en développant une vision stratégique qui dépasse largement le cadre de la directive européenne. « L'information est notre bien le plus précieux », affirme Miguel Ángel Serantes, responsable du développement informatique chez Crédito y Caución. « Nous sommes passés maîtres dans la localisation, le stockage et l'analyse des informations, et nous savons en dégager des renseignements stratégiques pour nos activités. Le défi posé par la directive Solvabilité II nous a donné l'opportunité d'intégrer des ratios de qualité dans la gestion des informations et dans nos procédures. Nous ne nous contentons pas de répondre aux exigences, nous nous engageons à instiller le degré de qualité maximal dans nos opérations de gestion des données. Nous avons transformé Solvabilité II en processus d'entreprise en l'intégrant dans notre environnement décisionnel »

### Première étape : évaluer la qualité des données

Pour relever ce défi, la première étape consistait à évaluer la qualité des données traitées par Crédito y Caución. « Nous avons tiré parti de l'option d'évaluation de la solution SAS pour analyser la qualité de nos propres données », ajoute M. Serantes. « Les résultats ont démontré que nous avions encore des progrès à accomplir. Il faut savoir que la plupart de nos données – noms de société, numéros de téléphone, codes fiscaux, etc. – proviennent de sources publiques comportant des degrés d'inexactitude divers. Nous avons alors décidé de concevoir et mettre en œuvre un modèle de qualité des données et avons opté pour SAS, que nous avons intégré à notre système de gestion. »

Miguel Ángel Serantes et son équipe ont développé les bases de la stratégie de gestion des données de la compagnie en définissant les critères essentiels auxquels les données gérées doivent répondre : elles doivent être précises, complètes et adaptées aux activités de l'entreprise. Ils ont ainsi déterminé les différents niveaux de détenteurs des données, qui seraient responsables de leurs contenus, définition, utilisation et administration. Enfin, ils ont établi des taux de conformité pour chaque catégorie de données, afin qu'il soit possible – via un système d'indicateurs – d'obtenir un aperçu immédiat du niveau de qualité de chaque élément de données.

### Un processus en constante évolution

« Nous avons choisi SAS pour plusieurs raisons », explique Miguel Ángel Serantes. « SAS est notre fournisseur de solutions de gestion des informations depuis des années. Nos entretiens d'excellentes relations. SAS a d'ailleurs détaché une équipe sur site chez Crédito y Caución, qui travaille en étroite collaboration avec notre service informatique. Tout ceci a contribué à l'intégration efficace de SAS Data Management dans notre système de gestion des informations. Cette solution est parfaitement adaptée à nos besoins. Elle permet d'établir des critères et des attributs pour définir la qualité des données, pos-

**Miguel Angel  
Serantes**

RESPONSABLE DU  
DÉVELOPPEMENT  
INFORMATIQUE  
CRÉDITO Y CAUCIÓN

« L'information  
est notre  
bien le plus  
précieux »

Miguel Angel Serantes

sède des options d'évaluation, identifie les problèmes de qualité et aide à résoudre les inexactitudes. Elle facilite la mise en œuvre de stratégies à long terme et assure un contrôle permanent de la qualité des données. »

Le déploiement du système de contrôle de la qualité des données chez Crédito y Caución a duré près d'un an. C'est pourtant un processus sans fin, comme le souligne Miguel Ángel Serantes. Le contrôle de la qualité des données est en constante évolution. Les avantages de cette stratégie et de la solution technologique implémentée ont été immédiats. « Pour les débutants, nous disposons de règles de gestion des données bien définies et connues de toute l'entreprise. Nous savons comment traiter les données et qui est responsable de chaque domaine. Nous savons à quel niveau se situent les lacunes et comment les combler. D'autre part, SAS nous fournit des informations sur la cause des inexactitudes. Nous prévoyons d'autres avantages qualitatifs, comme la définition d'objectifs de qualité pour chaque élément de données, ce qui nous permet de nous concentrer sur les contrôles propres à notre activité. »

Crédito y Caución vise 100 % de qualité dans la gestion des données générées par la compagnie elle-même, qu'elle peut soumettre à des contrôles rigoureux. Pour les données de sources extérieures, dont Crédito y Caución n'a pas la maîtrise directe, l'objectif est d'établir des critères de normalisation afin d'obtenir une qualité maximale. ■

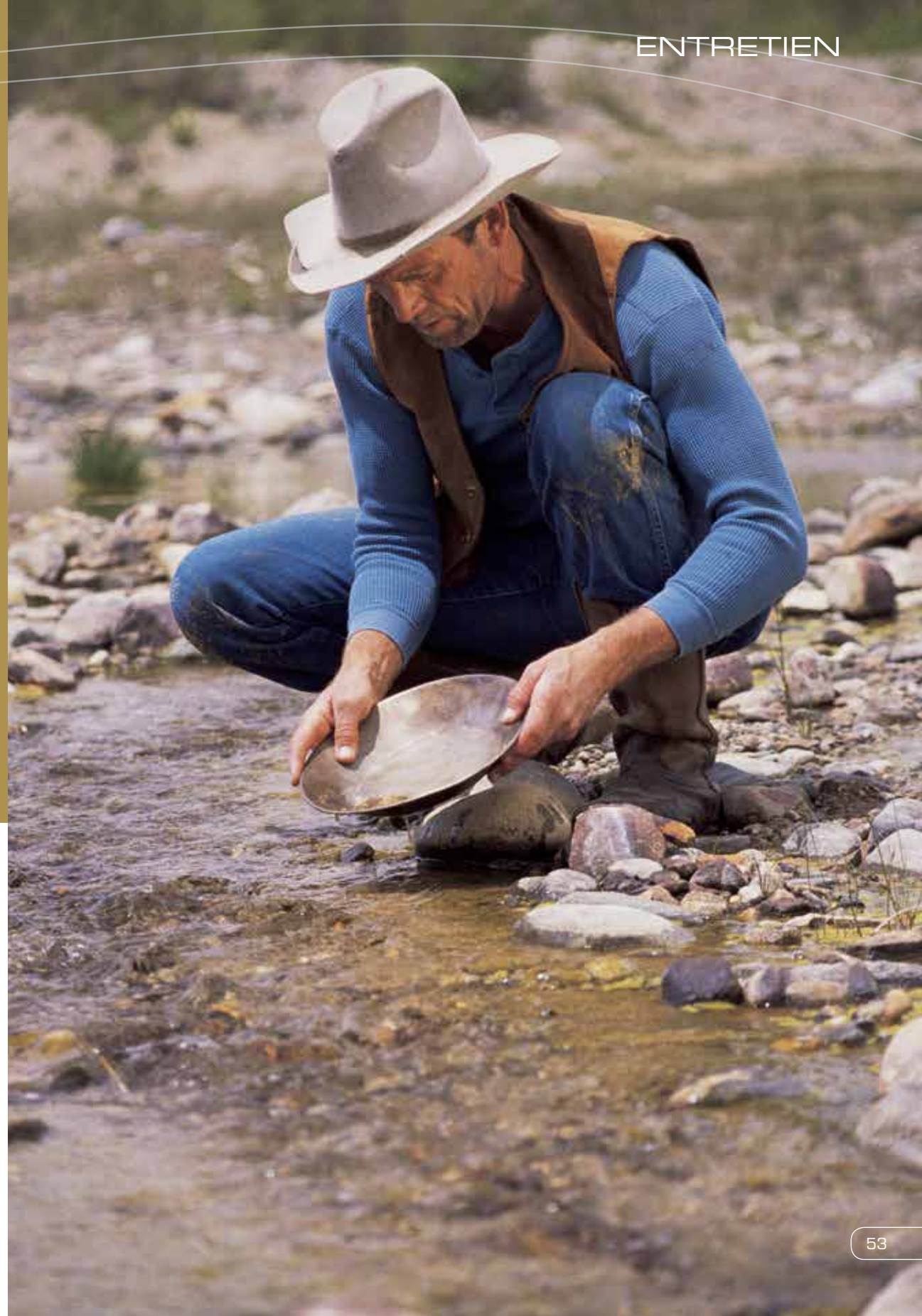
L'utilité des données dépend de leur qualité

## DES DONNÉES TRIBUTAIRES DE L'AMÉLIORATION DE LA QUALITÉ

Le temps, c'est de l'argent. De nos jours, ce vieil adage s'applique également aux données. Ces dernières valent leur pesant d'or, ce qui rend d'autant plus étonnant le manque d'intérêt apporté à leur qualité. Face à l'importance actuelle des données, notamment les métadonnées et les Big Data, la qualité devrait aller de soi. Le sujet se passerait-il d'explications ? « C'est ce que tout porte à croire », répond Bas Dudink, expert en gestion des données chez SAS Pays-Bas. « Mais la qualité des données est par définition médiocre » .

Bas Dudink sait parfaitement pourquoi les données collectées et utilisées par les entreprises sont souvent d'aussi médiocre qualité : parce que tout le monde utilise les mêmes. Cela pourrait sembler contradictoire, et pourtant ça ne l'est pas. Au contraire, cela signifie que des données de piètre qualité sont réutilisées à l'infini, avec leurs erreurs, leurs omissions et leur manque de contexte. Bas Dudink est catégorique sur ce dernier point : « Les données dépourvues de contexte sont inutiles. »

D'un autre côté, l'utilisation de données identiques suppose que la même source de données soit employée pour des utilisations diverses et à des fins différentes. Cette approche d'administration et de saisie de données peut paraître judicieuse, mais le niveau de qualité et les champs de données varient selon les applications. Ce problème concerne les applications et systèmes informatiques, mais aussi l'utilisation des données dans les différentes activités d'une entreprise.



### Le piège se cache dans les détails

Pour adresser une commande à un client, vous avez besoin de son adresse complète, mais pas de ses coordonnées bancaires. C'est l'inverse pour régler des factures. Cet exemple simple illustre des champs de données relativement identifiables et gérés avec attention – ou censés l'être. La situation se complique lorsque des détails mineurs susceptibles d'avoir de lourdes conséquences sont impliqués. C'est notamment le cas des livraisons entre entreprises comportant plusieurs divisions, où les différentes unités s'achètent et se vendent mutuellement des services. De minuscules écarts au niveau des ventes et/ou des achats peuvent entraîner des différences majeures. Telle ou telle division ou filiale offre-t-elle une remise spéciale, ou ses conditions de livraison et de paiement sont-elles différentes ? Les données relatives aux tarifs et conditions générales peuvent varier et par là même, induire une qualité médiocre.

### Le danger de l'ignorance

Bas Dudink remet en perspective son jugement sévère sur la médiocrité inhérente aux données : « Tout est affaire de personne. Ce qui est bien pour une personne ne l'est pas forcément pour une autre. C'est une question de sensibilisation. » Le pire n'est pas le désordre des données, mais l'absence de conscience du désordre. Bas Dudink reprend la célèbre citation selon laquelle « l'ignorance est un gage de bonheur ». Le plus grand danger consiste à penser que les données sont exactes et à jour, et de prendre des décisions et agir selon cette conviction. Autre problème : la plupart des entreprises estiment à tort ne pas être des sociétés de données, alors qu'aujourd'hui, chaque entreprise est une société de données. Cela s'applique même au cas apparemment simple d'un fabricant qui transforme des biens physiques en nouveaux produits. Son processus de fabrication et la conception de sa chaîne de production reposent l'un et l'autre sur des données !

### Un problème économique plutôt qu'informatique

La conception est généralement suivie d'une phase d'optimisation basée sur l'expérience pratique. Les données accumulées au fil du temps seront importantes si la société souhaite ouvrir une nouvelle usine ou optimiser une unité de production existante. La reproductibilité et l'amélioration sont des activités centrales qui dépendent d'une bonne qualité des données, résume Bas Dudink. Et une bonne qualité des données engendre de bons processus métier. C'est seulement à ce stade que l'automatisation est possible à l'échelle de l'entreprise. Une prise de conscience de la qualité des données au sein de l'entreprise est alors indispensable. « Ils doivent prendre conscience qu'ils ne savent pas tout. » Les fournisseurs informatiques qui proposent des solutions de données sont ainsi plus impliqués dans le domaine du conseil que dans le logiciel. « Nous ne nous contentons pas de proposer des logiciels, nous offrons également un éclairage. »

### Ne pas se contenter de gommer les erreurs

Les premières étapes d'amélioration des données ont pourtant peu de choses à voir avec les outils et les systèmes. Il y est principalement question d'analyse : quelles sont les données concernées, comment l'entreprise se les est-elle procurées, qui les a fournies et de quelle manière ? Ces questions relativement basiques sont toutefois susceptibles d'affecter les processus et procédures établis, voire les partenaires. En effet, qui ose se plaindre ouvertement de la mauvaise qualité des données d'un fournisseur ou d'un client ? C'est pourtant assez souvent le cas. « Si les données sont erronées,

« Ce qui est bien pour une personne ne l'est pas forcément pour une autre. C'est une question de sensibilisation »

Bas Dudink

se contenter de corriger les erreurs n'est pas suffisant. Il faut passer le processus au peigne fin et se demander d'où provient l'erreur. » La qualité des données est trop souvent considérée comme un moyen de gommer les erreurs. C'est parfois égocentrique de la part d'une entreprise désireuse de se faire passer pour une machine bien huilée. C'est ce que Bas Dudink appelle « sauver les apparences », mais la bonne approche de la qualité des données va bien plus loin et de manière plus approfondie.

### La nécessité d'un spécialiste

La qualité des données ne remporte pas une large adhésion dans les faits. Les brillants exemples d'amélioration dans ce domaine qui viennent à l'esprit de Bas Dudink ont été réalisés sous la contrainte. « Nous pouvons remercier les réglementations. Si la qualité des données n'était pas obligatoire, la plupart des gens laisseraient les choses en l'état. » Les règles et les lois qui imposent des données de meilleure qualité avancent un argument concret : la gestion des risques. La présence des banques et des établissements financiers – y compris les assureurs – dans le peloton de tête n'est guère surprenante.

Sans le bâton réglementaire, une entreprise a besoin d'un spécialiste de la qualité des données : « Une personne capable d'identifier les informations susceptibles d'être perdues et qui perçoit l'intérêt qu'il pourrait y avoir à les traiter. » La difficulté réside dans le fait que cette personne doit avoir une vue d'ensemble de l'entreprise, mais également le recul et le pouvoir nécessaires pour faire changer les choses. « Un DSI de préférence », résume brièvement Bas Dudink.

### Conseil élémentaire : commencez petit

Même si une personne occupant un poste de direction doit plaider en faveur de l'amélioration de la qualité des données, il est conseillé de débiter par un petit projet, le terme « petit » étant bien évidemment relatif et proportionnel à la taille de l'entreprise. Bas Dudink recommande d'utiliser le premier projet pour faire germer l'idée de l'amélioration de la qualité des données, comme un terrain propice à l'apprentissage. Le responsable haut placé à l'origine du projet devra associer une vision descendante à une approche ascendante. Cette combinaison est indispensable pour obtenir l'amélioration de la qualité des données de l'entreprise souhaitée. « Je suis persuadé que les DSI seront du même avis. » ■

# AMÉLIORER LA FIABILITÉ ET LA QUALITÉ DES DONNÉES

Feu Vert utilise la technologie  
de Data Management de SAS pour  
valoriser ses données clients



Feu Vert, premier réseau de centres automobiles en Europe, disposant de 400 points de vente en Europe, dont plus de 300 en France, a déployé la solution SAS® Master Data Management pour améliorer la fiabilité et la qualité des données de ses 10 millions de clients.



Avec 40 ans d'existence, Feu Vert est aujourd'hui le leader européen des accessoires et services automobiles. Avec la volonté de devenir la meilleure alternative aux réseaux constructeurs en devenant l'enseigne de référence des centres automobiles, Feu Vert a mis en place une stratégie de croissance globale sur 5 ans, intégrant un plan d'action centré sur ses données marketing. « Nous avons choisi la technologie de Master Data Management de SAS pour supporter cette stratégie et ré-internaliser la gestion de nos données marketing », explique Marie-Pierre Dussolliet-Berthod, Responsable Gestion de la Relation Client, Feu Vert. « Cela nous permettra également d'accélérer notre croissance grâce à une meilleure connaissance de nos clients et un meilleur adressage de nos offres ».

### Un référentiel clients unique

Au cours des six derniers mois, Feu Vert et SAS ont collaboré pour la mise en place d'un référentiel unique des clients en s'appuyant sur la plateforme SAS Data Management et la solution SAS Master Data Management. Ce projet a permis également l'initialisation d'un programme de gouvernance des données, l'amélioration des informations clients et de leur qualification, ainsi que la diminution des coûts par contact.

« Grâce à une meilleure connaissance de nos clients, nos offres sont mieux adressées »

Marie-Pierre Dussolliet-Berthod

Le référentiel de données ainsi créé constitue la pierre angulaire du système d'information marketing de Feu Vert en intégrant l'ensemble des applications opérationnelles qui impliquent les données client (ERP, Encaissement, Fidélité, eCommerce, Service Client ...). La solution de Master Data Management de SAS va en effet permettre de consolider et d'améliorer la qualité des données clients de Feu Vert, de capitaliser des informations client provenant des systèmes opérationnels et de fournir les données de référence au système décisionnel marketing.

### Un marketing personnalisé

« Les entreprises doivent désormais aborder le client de manière ciblée et personnalisée si elles souhaitent mettre en œuvre une communication et un marketing efficaces et profitables », commente Serge Boulet, Directeur Marketing de SAS France. « SAS aide les organisations à développer un référentiel de données maîtres qui fournit des informations pertinentes sur les clients et sur leurs besoins, améliorant ainsi la fidélité de ses clients et la profitabilité ».

La stratégie de gouvernance des données de Feu Vert va permettre à l'entreprise d'établir une analyse approfondie de ses bases de données marketing et de mener des campagnes marketing fiables et pertinentes, pour un meilleur ciblage des prospects et une augmentation de la fidélité de ses clients. ■



L'interview de Marie Pierre Dussolliet-Berthod est disponible en vidéo sur Youtube: [tinyurl.com/nwn57x4](https://tinyurl.com/nwn57x4)

## Analyse des flux de données pour la prise de décision opérationnelle

# LES DÉCISIONS ÉTAYÉES PAR DES DONNÉES FONT TOUTE LA DIFFÉRENCE

**Vite, plus vite, toujours plus vite ; ce pourrait être le slogan de l'époque actuelle. L'informatique a accéléré les choses, mais le flot des données menace aujourd'hui de nous inonder. Même si une analyse attentive des données peut s'avérer fastidieuse, le traitement des flux d'événements et la gestion des décisions permettent de l'accélérer et d'en augmenter l'efficacité.**

Nous générons, stockons et combinons de plus en plus de données, que nous souhaitons soumettre à des analyses toujours plus complexes. « Dans certains environnements, les énormes volumes de données peuvent être écrasants », affirme Andrew Pease, Principal Business Solutions Manager au Centre d'excellence Plate-forme analytique de SAS Belgique. « Il n'est pas toujours possible de les stocker intégralement, malgré la baisse des coûts de stockage des données. » D'autant plus que les coûts de stockage sont loin d'être les seuls en jeu. Les coûts d'analyse peuvent être tout aussi prohibitifs, notamment en cas d'analyse rétrospective. Dans certains cas, l'analyse est du reste trop tardive pour être d'une quelconque utilité.

La solution consiste à filtrer les données en temps réel via le traitement des flux d'événements, puis d'appliquer un processus décisionnel automatisé à l'aide d'une fonction de gestion des décisions. Cette automatisation a lieu sur la base des processus métier intrinsèquement liés aux niveaux décisionnels tactiques et stratégiques. Les décisions concernant la production, voire les interfaces directes avec le client, peuvent à présent être prises automatiquement sans perdre de vue les objectifs stratégiques de l'entreprise.

## Tendances et anomalies des plates-formes de forage

Andrew Pease évoque un exemple où le traitement des flux d'événements peut être utile : la saisie des données de capteurs sur les plates-formes de forage. Certains de ces capteurs détectent les vibrations, ce qui permet d'identifier les anomalies. Les anomalies pouvant renvoyer à des problèmes, il est important d'obtenir rapidement les résultats des analyses de données.

La priorité et le champ d'application sont ici déterminants. « Vous devez vous concentrer en priorité sur ce que vous voulez savoir. » L'expert SAS explique qu'une analyse des vibrations chaque seconde ou chaque minute serait sans intérêt, un intervalle fixe par heure pouvant être suffisant. L'essentiel est de repérer les tendances et anomalies pour effectuer des pronostics. Le traitement des flux d'événements peut traiter aussi bien le code de scoring généré par de puissantes analyses prédictives que les règles définies par des experts pour permettre ce type d'analyse et déterminer si la maintenance préventive est garantie.

## Offres mobiles et antennes-relais supplémentaires

Autre exemple : l'utilisation des données d'appel par les opérateurs téléphoniques. « Les opérateurs sont assis sur d'énormes piles de données », relève Andrew Pease. Chaque échange téléphonique implique 20 à 30 enregistrements de détails. « Mais ils ne valent pas tous la peine d'être analysés. » En revanche, l'analyse des éléments de données utiles peut révéler nombre d'informations intéressantes. À titre d'exemple, les données clairement utiles concernant l'heure où un client appelle permettent de lui proposer une offre plus adaptée. Le comportement téléphonique, qui permet souvent de dégager des tendances générales, doit faire l'objet de toutes les attentions. « Si la facture du client est trop élevée pendant trop longtemps, il risque de changer d'opérateur. »

Une telle analyse peut en outre indiquer des lacunes dans la couverture réseau. L'opérateur peut ainsi décider d'installer des antennes-relais supplémentaires sur la base du nombre d'appels interrompus. En analysant ces derniers, il peut en effet identifier précisément les problèmes de couverture ou de capacité, ce qui lui permet ensuite d'installer une antenne-relais supplémentaire à l'endroit exact où le nombre d'appels interrompus est le plus élevé.

## Prendre les fraudeurs en bourse en flagrant délit

Le monde financier travaille plus vite que la plupart des autres secteurs. Les échanges boursiers sont en effet automatisés et ultra-rapides. Dans cet univers soumis à un rythme effréné, le traitement des

## Event Stream Processing

Le traitement des flux d'événements permet de contrôler et analyser en temps réel de grandes quantités de données, issues par exemple de transactions ou de capteurs. Le système peut immédiatement et automatiquement réagir à des situations spécifiques ou aux interférences concernées.



« La rapidité est devenue aussi importante que l'était la qualité de l'analyse »

Andrew Pease

flux d'événements permet de détecter rapidement la fraude boursière, affirme Andrew Pease. La rapidité des systèmes d'échange rend impossible le stockage de toutes les données qui transitent par ces systèmes, sans parler de les analyser.

« Reste à éliminer les anomalies, autrement dit, les transactions suspectes. » Andrew Pease cite un ouvrage paru cette année, intitulé *Flash Boys*, qui évoque les « flash traders », ou spéculateurs à haute fréquence. L'auteur, Michael Lewis, dresse un tableau extrêmement sombre des marchés monétaires internationaux, où la rapidité des données est essentielle. Plus elles sont rapides, plus le prix d'achat ou de vente est intéressant. Même si cet ouvrage – qui est tout sauf un livre de fiction – a fait l'objet de critiques, il n'en contient pas moins des leçons et vérités instructives.

Andrew Pease relate l'escroquerie imaginée par un trader canadien. Il a séparé ses activités d'échange et a trouvé un second serveur plus proche de la Bourse de deux kilomètres. Il s'est alors fixé un délai d'attente extrêmement précis afin que ses ordres parviennent exactement à la même heure à différentes chambres de compensation, ce qui lui a permis d'échanger des actions en nombre tout en s'assurant que ses ventes n'entraînent pas une baisse du prix des actions et par là même, de ses bénéfices. Malin et quasiment indétectable.

### Suivre le rythme rapide des flux de données

« De nos jours, certains flux de données sont si rapides qu'il est impossible d'en analyser entièrement le contenu », souligne Andrew Pease. Ceci s'explique par la lenteur du traitement des données. De même que la baisse des coûts de stockage ne compense pas entièrement l'explosion des données, les progrès de la puissance informatique ne parviennent pas toujours à suivre le rythme des impératifs d'analyse. Le tour de passe-passe spéculatif basé sur la temporisation du serveur est pratiquement impossible à détecter sans traitement des flux d'événements. Ce traitement transforme en effet des flux de données rapides en blocs de taille raisonnable.

La demande d'analyse rapide et en temps réel des données ne cesse d'augmenter. Ceci est dû aux nouvelles applications, au nombre croissant de capteurs qui collectent des données et à la multiplication des modèles économiques, tous tributaires de la fiabilité de l'analyse des données. La rapidité

est devenue aussi importante que l'était la qualité de l'analyse. Les développements concernent au premier chef l'industrie, le secteur financier et celui des télécommunications, précise Andrew Pease.

### Brassage de bière et conseils gastronomiques

Les industriels installent eux aussi de plus en plus de capteurs dans leurs environnements de production. C'est bien évidemment un avantage pour la fabrication de produits technologiquement complexes, mais également pour d'autres applications. Les brasseurs, par exemple, utilisent des capteurs pour détecter le taux d'humidité des houblons et programmer la transformation des matières premières en conséquence. Cela leur permet en outre de planifier plus efficacement la maintenance des chaudières.

Le traitement des flux d'événements est également implémenté dans d'autres secteurs, comme la grande distribution. Il en va de même pour la gestion des décisions. Andrew Pease en souligne l'intérêt pour l'utilisation de scanners par les clients des grandes surfaces. Si un client pose une bouteille de vin dans son chariot, pourquoi ne pas lui suggérer un fromage pour l'accompagner ? Le supermarché pourrait même proposer un rabais pour rendre l'offre plus attrayante. Le système informatique sait qu'il existe actuellement un surplus de fromage proche de la date de péremption et que sa vente à prix réduit reste valable. Un système de gestion des stocks fiable, comprenant des informations sur les dates limites d'utilisation, est ici nécessaire. Ce ne sont pas tant les téraoctets qui comptent que la rapidité : tout ceci doit avoir lieu avant que le client n'arrive à la caisse. « Vous devez faire la bonne offre au bon moment. »

### L'importance de la stratégie commerciale

L'offre adressée au client doit également s'harmoniser avec la stratégie commerciale de l'entreprise. Le principe de la gestion des décisions consiste à centraliser et normaliser cette stratégie. Chaque décision doit concorder avec la stratégie commerciale. Une remise proposée par le service marketing cadre-t-elle avec la politique du service de gestion des risques ? Si un client achète en gros, la remise pourrait être plus élevée. La stratégie concerne également la valeur à plus long terme que ce client représente pour l'entreprise. Des règles clairement définies sont nécessaires pour la gérer correctement. « Les règles doivent être saisies et administrées depuis un emplacement central. » Elles doivent en outre être appliquées uniformément à tous les systèmes et activités. C'est la seule manière de s'assurer que les différents silos de l'entreprise interagissent en vue d'une gestion efficace des décisions. ■

## Decision Management

Il a pour but d'améliorer les organisations en permettant une prise de décision plus rapide et avisée (opérationnelle).

Sa mise en œuvre passe par l'application automatique de modèles analytiques et de déclinaisons de règles métiers à l'ensemble des données disponibles, afin de mesurer les délais requis et les risques potentiels.

## Prendre le contrôle de vos données

# LA GESTION DES DONNÉES MAÎTRESSES, PIERRE ANGULAIRE DE VOTRE ACTIVITÉ

Il semble logique que les données fournisseur soient exactes et à jour. Après tout, les fournisseurs sont des partenaires commerciaux importants ; ils nous procurent des biens et services pour lesquels nous les payons. Pourtant, les données les concernant laissent souvent énormément à désirer. Mis à part les simples erreurs de saisie et la corruption des données, il existe une autre cause de disparités : les différences de saisie, parfois d'un système à l'autre.

## Fournisseur A = Fournisseur 1 (= B2)

Un système métier peut parfaitement renvoyer au « Fournisseur A » tandis qu'une application back-end le désigne sous l'appellation de « Fournisseur 1 ». La recherche et le remplacement automatiques de « A » par « 1 », ou inversement, semble la solution la plus évidente. Le problème, c'est que cela peut avoir des effets indésirables, comme des références inutilisables dans les systèmes concernés ou d'autres applications.

Ensuite, cela crée un risque supplémentaire d'écarts d'écritures. Par exemple, une facture envoyée par Fournisseur A a disparu. Elle a en fait été renommée. En théorie, il pourrait exister une couche intermédiaire ayant supervisé le processus de changement de nom et capable de remonter jusqu'au nom d'origine. Reste la question de l'impact sur les performances système, car cet exemple est excessivement simple.

Les données sont moins simples qu'il n'y paraît. Prenez l'exemple d'une entité de données élémentaire : le nom d'un fournisseur. Il peut varier d'un pays, d'un processus et/ou d'une activité à l'autre. La gestion des données maîtresses (Master Data Management ou MDM) permet de résoudre ces différences et d'éviter bien du travail et des rappels inutiles.



Bas Dudink

EXPERT EN  
GESTION DES  
DONNÉES  
SAS PAYS-BAS

« Certaines entreprises possèdent jusqu'à 50 systèmes ERP, ce qui signifie que les données maîtresses peuvent être stockées et utilisées de 50 manières différentes »

Bas Dudink

### Des divergences dues à une croissance incontrôlée

Comme le sait d'expérience Bas Dudink, expert en gestion des données chez SAS Pays-Bas, les divergences observées dans les données maîtresses sont généralement dues à une croissance incontrôlée des systèmes informatiques. Cette croissance est souvent liée aux systèmes ERP, les environnements professionnels stratégiques de planification des ressources. Ces systèmes complexes et étendus ne peuvent jouer leur rôle inestimable qu'après un long projet d'implémentation, ce qui signifie qu'ils sont étendus, personnalisés ou abandonnés uniquement en cas d'absolue nécessité.

Si la croissance explosive des systèmes informatiques peut s'inscrire dans le cadre d'un processus organique, elle peut également être la conséquence d'acquisitions. La société X achète la société Y et les fournisseurs A, 1, B2, etc. se retrouvent dans l'environnement informatique commun de la nouvelle division obtenue. Il est souvent impossible d'intégrer complètement la totalité des systèmes. L'opération est parfois inscrite dans un mégaprojet censé se dérouler « ultérieurement », une fois l'impact de l'acquisition absorbé et les opportunités commerciales saisies.

### Jusqu'à 50 systèmes ERP

« Certaines entreprises possèdent jusqu'à 50 systèmes ERP », affirme Bas Dudink. « Pour différents pays, différents processus métier, etc. » Cela signifie que les données maîtresses peuvent être stockées et utilisées de 50 manières différentes, différences qui n'ont rien de théorique, comme il le constate trop souvent.

### Premières étapes : centralisation et déduplication

Bas Dudink conseille d'impliquer très tôt les personnes et les services amenés à bénéficier des avantages du MDM, afin de rendre la valeur du projet tangible. Le MDM peut avoir un énorme impact sur l'incroyable inefficacité causée par ces 50 formes différentes de stockage des données dans 50 systèmes ERP. La première étape est la mise en œuvre centralisée, à l'origine d'une réduction considérable des frais administratifs. En toute logique, l'étape suivante devrait consister à rendre les données disponibles dans l'entreprise élargie, mais ce n'est pas encore le moment idéal. La deuxième étape essentielle du MDM doit en effet avoir lieu au préalable.

Il s'agit de la déduplication des données. Cela suppose, par exemple, d'identifier un fournisseur figurant sous deux appellations différentes dans l'environnement informatique et d'en supprimer une. C'est la seule manière d'être certain que les divergences des données maîtresses n'entraînent pas la disparition d'une facture pendant trois mois. Cela vous semble inconcevable ? C'est pourtant réellement arrivé dans une grande entreprise. Une facture d'un montant considérable s'est simplement retrouvée dans la mauvaise corbeille – à ceci près que personne ne savait dans quelle « corbeille » de l'immense environnement numérique elle se trouvait.

### Qui est autorisé à modifier quoi ?

Si l'erreur est humaine, il n'en est pas moins nécessaire de l'éviter le plus possible. C'est ce que l'on pourrait considérer comme la troisième étape de mise en œuvre d'un projet de MDM : prévenir les nouvelles erreurs de données. La question clé porte sur la désignation des personnes ayant le contrôle des données maîtresses. Qui a le droit de modifier les coordonnées bancaires d'un fournisseur ? La distinction entre appartenance et utilisation des données est importante. À qui appartiennent des données saisies aux Pays-Bas mais principalement utilisées aux États-Unis ?

L'identification de l'appartenance et de l'utilisation des données doit conduire à un examen logique de cette appartenance. Une fois nettoyées, les données seront plus facilement exploitables par les divisions et les processus métier qui en tirent le plus parti. Les données seront initialement déployées à partir du lieu de stockage central (désormais nettoyé) et seront ainsi mises à la disposition des niveaux « inférieurs » de l'organisation.

### La gestion des données maîtresses ne se limite pas à la détection des erreurs

Une gestion efficace des données maîtresses doit aller bien au-delà de la détection et de la résolution des erreurs. Elle peut également servir à détecter et prévenir les fraudes telles que les « fausses factures ». Ces dernières sont envoyées par des escrocs et ressemblent à s'y méprendre à celles d'honnêtes fournisseurs. Le MDM empêche ces fraudeurs de se glisser à travers les failles du système.

Les divergences, les erreurs et la fraude peuvent être dues à la complexité croissante des entreprises et de leurs systèmes. Lorsqu'un fournisseur ou un client est enregistré la première fois, l'opération se déroule normalement sans problème. Cela change après quelques années et des milliers de transactions, traitées par des centaines de personnes de différents services, dont la structure organisationnelle risque d'avoir changé elle aussi. L'enregistrement de ce fournisseur ou client a depuis fait l'objet d'innombrables modifications, actions et exceptions. « Veiller à ce que chaque transaction soit correctement traitée est un véritable défi », affirme Bas Dudink. Ce défi peut être relevé grâce à une gestion efficace des données. ■

## A PROPOS DE SAS

SAS est pleinement conscient du rôle majeur exercé par les données dans le monde d'aujourd'hui. Notre ambition est de vous aider à appréhender efficacement ce phénomène. Vos données sont-elles aisément accessibles, nettoyées, intégrées et correctement stockées ? Savez-vous quels types de données sont utilisés dans votre entreprise et qui s'en sert ? Pour finir, disposez-vous d'une méthode de validation automatique des données entrantes avant qu'elles soient stockées dans vos bases de données ?

### **Prendre de meilleures décisions**

Des milliers de décisions, peut-être même des centaines de milliers, sont prises chaque jour dans votre entreprise. Des décisions quotidiennes prises dans le cadre d'un processus : Pouvons-nous accorder un prêt à ce client ? Quelle offre dois-je proposer à un client qui appelle notre centre de contact ? Des décisions tactiques, telles que : Quel est l'équilibre optimal entre entretien préventif et maintenance réactive du matériel ? Si nous devons abandonner cinq des quinze parfums que nous proposons, lesquels serait-ce ? Mais aussi des décisions stratégiques sur l'orientation de l'entreprise. Par exemple : dans quelles combinaisons de produit et de marché voulons-nous être présents ? Les informations jouent un rôle dans toutes ces décisions. Plus la qualité des données sous-jacentes est élevée, plus les décisions que vous prenez sont pertinentes.

### **Prendre le contrôle des données**

Veiller à ce que les données soient complètes, exactes et à jour peut être une lourde tâche, qu'il est heureusement possible d'automatiser en grande partie. Passez moins de temps à recueillir et actualiser des informations et davantage à diriger votre entreprise grâce à SAS Data Management. Cette solution a été intégrée à une plate-forme unifiée et conçue dans une optique à la fois économique et informatique. Elle vous permet de prendre le contrôle de vos données de la manière la plus rapide, facile et globale. SAS Data Management optimise les performances en mémoire et celles des bases de données, pour offrir un accès en temps réel à des informations fiables.

### **Définir correctement les processus**

Pour réussir dans le monde actuel orienté données, il vous faudra bien plus qu'une solide plate-forme de gestion des données. Les processus et le comportement jouent également un rôle important en matière de MDM, sans oublier l'intégration, la qualité, la gouvernance et la fédération des données. Nous pouvons aussi vous aider à définir correctement ces éléments, car la mise à jour des données n'est pas une activité ponctuelle, mais un processus continu. Quel que soit le volume de données dont vous disposez, vous pouvez les transformer en valeur précieuse et en formidables opportunités.

*Vous voulez en savoir plus ? Visionnez la vidéo: [tinyurl.com/ocbynpu](http://tinyurl.com/ocbynpu)*

**THE  
POWER  
TO KNOW®**

# COLOPHON

**Réalisation :**

SAS Netherlands

**Rédactrice en chef :**

Ilanite Keijsers

**Photographie :**

Eric Fecken

**Rédacteurs :**

Jasper Bakker

Mirjam Hulsebos

Chantal Schepers

**Couverture :**

Philip van Tol

**Conception réalisation :**

Alain Cohen

**Direction de la publication :**

SAS Netherlands

La version originale de cet ouvrage, Future Bright – A Data Driven Reality, a été réalisée sous la supervision de SAS Netherlands. Sa traduction française intitulée Future Bright – Laissez-vous guider dans un monde de données, a été réalisée pour SAS France. Le contenu de l'ouvrage peut être utilisé sous forme d'impression, de photocopie, de vidéo ou tout autre support multi média, publié sur Internet ou tout autre média, sous condition expresse de l'autorisation de SAS Netherlands et de la mention de la source. SAS Netherlands comme SAS France ne peuvent être tenus responsables des déclarations faites par les personnes interviewées dans cet ouvrage.