

Décisionnel, Performance, Qualité

Dans le cadre des rencontres « Autour d'un verre » d'ADELI

Rencontre animée par Dominique Mollard
Compte rendu de Véronique Pelletier

Le 2 avril 2007, Dominique Mollard nous a présenté son expérience « Décisionnel, Performance, Qualité » lors d'une rencontre « Autour d'un verre ».

La mise en place de Systèmes d'Information Décisionnels (SID) dédiés au pilotage de la performance facilite la prise de décision et l'alignement stratégique des organisations en recherche d'efficacité et d'efficience.

Ces systèmes assurent la restitution d'informations fiables, précises et pertinentes au moyen d'indicateurs structurés en tableaux de bord. Ils s'appuient sur les méthodes du contrôle de gestion et du pilotage de la performance : benchmarking, balanced scorecard, etc.

Ils trouvent également une application dans le domaine de la Gouvernance des Systèmes d'Information où leurs fonctionnalités sont étendues à la mesure et au pilotage de la qualité. C'est en combinant des approches fonctionnelles bien établies, telles que le contrôle de gestion, la mesure de la qualité et de la performance, avec des technologies décisionnelles qui s'appuient sur des portails d'entreprise que les maîtres d'œuvre des systèmes décisionnels peuvent apporter une vraie valeur aux organisations, publiques ou privées. Les principes de construction d'un SID pour le pilotage de la performance et de la qualité doivent respecter cette logique fonctionnelle en s'appuyant sur un modèle d'architecture (datawarehouse et datamarts modélisés en étoile, Operational Data Store, cubes et tableaux de bord diffusés via le Web) qui associe une économie de réalisation et de mise en œuvre avec une qualité maîtrisée des informations restituées.

La rencontre « Autour d'un verre » a été l'occasion d'échanger sur des thèmes majeurs de la construction de systèmes décisionnels tels que la conception des systèmes d'alimentation, de stockage et de restitution, ainsi que la gestion de la qualité des données, la validation du SID et la gestion de ses évolutions.

Le conférencier

Dominique Mollard est ingénieur, diplômé du CNAM. Il est directeur de projet dans une grande administration française. Il a une expérience de la conduite de projets décisionnels, principalement dans le domaine du contrôle de gestion et de la performance et a conduit également un projet qui visait des indicateurs centrés autour du développement et de la qualité informatique. C'est le thème qui a été choisi pour cette rencontre « Autour d'un verre ».

La présentation était articulée autour de trois axes :

- Le premier concerne la prise de décision et les besoins de pilotage de la performance et de contrôle de gestion et permet de répondre à la question : Pourquoi construire des systèmes décisionnels ?
- Le second axe répond à la question : Comment construire un système décisionnel, en privilégiant une présentation pragmatique plutôt que théorique ?
- Le troisième axe fait un parallèle entre des systèmes décisionnels pour la performance et le contrôle de gestion, et la gouvernance des systèmes d'information.

Pourquoi construit-t-on un système décisionnel ?

La prise de décision suppose trois éléments. Pour prendre une décision il faut :

- avoir des objectifs (savoir ce que l'on veut faire),

- disposer d'informations suffisantes par rapport à ces objectifs (savoir où l'on en est),
- rapprocher ces informations des objectifs pour prendre une décision qui va entraîner une action (savoir quoi faire).

Il faut donc définir des objectifs qui, pour être atteints, vont nécessiter des actions appropriées. Pour pouvoir définir ces objectifs et mesurer l'effet des actions, il faut disposer d'informations. C'est là qu'interviennent les systèmes d'informations décisionnels.

C'est le thème de la présentation de Dominique Mollard : construire un système d'information décisionnel.

Les objectifs peuvent être des objectifs d'amélioration, des objectifs d'alignement de l'informatique sur la stratégie de l'entreprise, des objectifs par rapport à la concurrence. Il faut les meilleures informations possibles pour prendre les décisions appropriées et pour engager les actions pertinentes. Et c'est là une bonne partie de la logique de construction de ces systèmes décisionnels, qu'ils soient pour la performance et le contrôle de gestion ou la qualité. Il est donc possible, indique Dominique Mollard, de construire des systèmes qui permettent à une organisation de se situer par rapport à ses objectifs, de savoir où elle en est par rapport à ses concurrents, et pourquoi pas, de l'appliquer au domaine de la qualité.

Comment construit-on un système décisionnel ?

L'objectif est d'aider la prise de décision, en s'appuyant sur des observations telles que :

- le nombre de défauts dans une chaîne de production ;
- le nombre d'heures potentiellement perturbées pour les utilisateurs, en cas d'arrêts « machine », qu'ils soient intempestifs ou programmés pour cause de maintenance ;
- le nombre d'erreurs selon les phases du processus de développement

Il y a forcément des décisions à prendre ; par exemple, changer de fournisseur, si la fiabilité de tel ou tel matériel apparaît insuffisante ou proposer un complément de formation au chef de projet en ingénierie du logiciel, pour améliorer sa stratégie de tests. C'est l'exemple d'actions que l'on pourra décider à partir d'informations pertinentes et factuelles.

Il est difficile de construire un système décisionnel sans connaître la logique fonctionnelle ou métier qu'il y a derrière. Les contrôleurs de gestion veillent à la maîtrise de l'activité et des coûts de l'entreprise afin de garantir sa pérennité. Les comptables procèdent aux enregistrements nécessaires des opérations, mais les contrôleurs de gestion doivent prendre du recul par rapport à ces données pour voir, par l'analyse et la prospective, où va l'entreprise, et éclairer le décideur.

Il en est de même pour la qualité, la logique du contrôleur qualité n'est pas forcément d'obtenir une certification ISO 9001, bien qu'en terme d'image ce soit important, mais également que l'entreprise puisse appliquer ce référentiel à un coût raisonné. Il faut trouver un mode de production optimum qui garantisse à la fois une production de qualité et des coûts d'obtention et d'assurance de cette qualité, qui soient raisonnables. La mesure des coûts dépasse le domaine du contrôle de gestion, et on se pose la question du coût des actions à engager pour obtenir une certification ISO 9001 ou pour atteindre un niveau donné de CMMI : la comparaison avec le gain qui en découlera pour l'organisation n'est pas évidente. La réponse à cette question nécessitera de collecter des informations avec un niveau de fiabilité et de pertinence approprié. L'analyse des coûts est le thème principal du contrôle de gestion. Son implémentation sera réalisée au moyen des règles de gestion du système décisionnel, qui seront alignées par rapport à cette logique.

La performance

C'est un terme « à la mode » depuis quelques années. Des spécialistes du coaching ou des entraîneurs d'équipes sportives animent des conférences et expliquent aux chefs d'entreprise leur approche de

la performance individuelle et collective. La performance est l'obtention de résultats conformes aux objectifs avec une économie optimale des moyens. Par exemple, les administrations ont élaboré dans un premier temps des contrats d'objectifs et de moyens. Un budget était alloué pour réaliser un plan d'actions sur trois ans. Les actions engagées par les services visaient à atteindre ces objectifs. Un bilan annuel permettait de se situer par rapport à leur atteinte.

Les contrats de performance ont permis d'aller au-delà, en évaluant le rapport entre l'objectif et les moyens, c'est la pertinence. Est-ce que les moyens ont été utilisés de façon pertinente par rapport aux objectifs ? Fixer des objectifs implique de les rapprocher des résultats obtenus, ce qui permettra de démontrer l'efficacité.

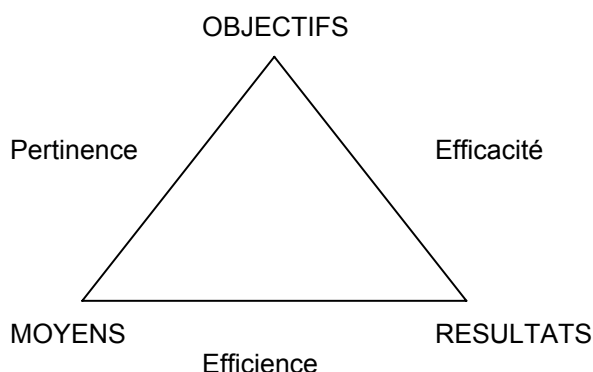
Par exemple, l'objectif à atteindre correspondra peut-être à une diminution du nombre de défauts dans la chaîne de production du logiciel, ou à un essai de stabilisation de la production des spécifications sur un projet. Si l'on reçoit des modifications des spécifications en cours de recette, en cours de qualification, en cours d'intégration, voire même, une fois l'application déployée, ce n'est pas efficace et cela génère probablement des coûts.

Les résultats atteints par rapport aux objectifs ne pourront être discutés qu'à partir du moment où on les mesurera de façon précise et fiable. Pour atteindre cet objectif, Dominique Mollard indique qu'il a utilisé la norme ISO 9126, et en particulier les principes de mise en place d'un système de mesures.

Ce que l'on ne mesure pas, on ne le connaît pas. C'est particulièrement vrai dans le monde du développement logiciel. Il est difficile d'expliquer à un Maître d'Ouvrage, convaincu de la qualité de son travail de spécificateur, qu'il y a des manques, et que des points fonctionnels importants n'ont pas été spécifiés.

Cette situation peut rapidement devenir conflictuelle. L'intérêt de disposer d'indicateurs qui permettent de mesurer la qualité du processus de spécifications réside dans la possibilité d'établir que les spécifications ont bien été livrées dans les délais, donc dans le respect du jalon de livraison, ce qui a permis d'engager la réalisation, mais d'un autre côté, que d'autres versions complémentaires ont été livrées pendant la phase de développement, ou en cours de recette. La livraison en fin de conception a été faite dans les délais mais la qualité de la production n'était pas satisfaisante. La preuve en est que l'indicateur suivant le nombre de livraisons pendant la phase de développement met en évidence ces livraisons « tardives ». Ces informations peuvent être pertinentes lorsque l'objectif est d'améliorer le processus de développement et de déclencher des actions d'amélioration du côté de la Maîtrise d'Ouvrage. Il faut être capable de le mesurer. L'efficacité, dans ce cas là, serait indiquée par la production des spécifications dans les délais. La performance serait analysée par le rapprochement de cette efficacité avec les

moyens. On cherchera alors à produire le même résultat avec des moyens optimaux. C'est le principe de l'efficacité. Les contrôleurs de gestion représentent cela de la façon suivante :



On mesure ainsi la pertinence, l'efficacité et l'efficacité. Ce sont les bases d'un système décisionnel pour la mesure de la performance, pour le contrôle de gestion et pour la démarche qualité.

On peut ajouter une quatrième dimension, proposée par Yvon Pesqueux, qui permet de prendre en compte la satisfaction de l'utilisateur. Yvon Pesqueux introduit la notion de partie prenante ; par exemple un actionnaire de l'entreprise, ou le citoyen contribuable pour une administration.

Le décideur a besoin d'avoir une connaissance de ce qui se passe pour agir. De là, naît le besoin d'un reporting élaboré, quitte, au besoin, à se confecturer ses propres tableaux de bord. Il est également nécessaire de prédire pour orienter l'action. Comment prédire ? Sur le site Internet du CNAM (www.cnam.fr), le LIPSOR (Laboratoire de Prospective) présente des outils permettant de réaliser des études prospectives. Ces outils s'appuient sur une méthode statistique élaborée qui nécessitera des données historiques pour alimenter ces outils et méthodes. Au fur et à mesure de la collecte des données, des chroniques seront disponibles et constitueront, au bout de six mois ou un an, un volume de données suffisant pour faire des prévisions. Il s'agit d'une dimension supplémentaire des systèmes décisionnels, la dimension temporelle.

Logique de construction des systèmes décisionnels

Comment tirer parti de toutes ces données collectées dans les systèmes comptables ? On sait exactement combien on a dépensé sur tel ou tel poste, on sait quand on l'a fait, et on sait à quoi cela sera attaché. Cela se rattache à la version A du logiciel ou au projet x ou y. Dès l'instant où l'on peut rattacher des informations que l'on peut formuler sous forme d'indicateurs, on améliore notre connaissance de ce qui se passe.

La logique de construction des systèmes décisionnels a évolué d'une approche « verticale », autrefois fondée sur un infocentre ou un service de la statistique, vers des architectures plus souples. Cette souplesse est apportée dès la conception du système d'alimentation, appelé à gérer les évolutions des systèmes produisant les données, tels que les logiciels de comptabilité qui ont leur propre cycle de vie.

Il faudra ainsi répercuter vers le système décisionnel chaque évolution de ce logiciel, ce qui soulève la question du coût d'obtention de ces informations. Un système de contrôle de gestion ne peut pas être en contradiction avec lui-même et coûter plus cher que les économies qu'il permettra d'apporter.

Une autre question concerne les utilisateurs des tableaux de bord, et en particulier le « décideur ». On imagine quelqu'un, quelque part vers le haut de la pyramide, qui régulièrement accède au tableau de bord, rapproche les informations dont il dispose des objectifs et prend les décisions nécessaires. Il s'agit d'une approche ascendante qui consiste à faire remonter les données nécessaires à la prise de décision puis à communiquer de façon descendante les actions correspondantes. Les organisations ne sont plus aujourd'hui structurées de cette façon, mais plutôt selon différents niveaux avec la mise en place de mécanismes de délégation ou de déconcentration. Ce type d'organisation induit le besoin de tableaux de bord multi niveaux. Un décideur aura par exemple besoin de n'utiliser à un instant donné qu'une partie de la centaine d'indicateurs disponibles dans le système décisionnel, pour suivre une campagne ou des événements sur une période bien délimitée. Sur un autre axe, l'organisation sera très probablement structurée en domaines fonctionnels, chacun ayant ses propres préoccupations et besoins de suivi. Par exemple, le service marketing aura des préoccupations différentes de celles du service de l'approvisionnement et cherchera à analyser les ventes d'un secteur donné. Ils vont surveiller 7 à 8 indicateurs qui leur permettront de dialoguer avec les responsables des ventes et leur responsable hiérarchique, qui disposera, de son côté, des indicateurs correspondant à son niveau de responsabilité. À l'autre bout de la chaîne de ces systèmes, il faut que la restitution des informations décisionnelles soit adaptée à l'utilisateur selon différents niveaux.

Autre problème, celui du stockage. Si un système décisionnel ne nécessite pas forcément de gros volumes de données au moment de sa mise en œuvre, le stockage des données de détail sur des périodes longues peut faire grossir la base de données dans des proportions importantes. Le volume d'informations potentiellement restituables sera difficile à faire tenir sur un tableau de bord et à traiter pour prendre une décision. Il existe également une différence d'échelle entre le système décisionnel d'une PME, d'une petite structure ou d'une agence et celui d'une organisation plus importante. Dans le

domaine du contrôle de gestion, les volumes dépassent rarement 150 giga-octets. En revanche, pour des domaines d'analyse de population ou de ciblage ou pour suivre des processus tels que la performance ou la qualité, on peut avoir besoin de grosses volumétries. L'idée, c'est d'adopter un modèle d'architecture qui permette, quelle que soit l'échelle, de s'adapter aux besoins.

Un système décisionnel est composé de trois maillons :

- l'**alimentation**, essentielle, que l'on soit alimenté par un système comptable ou un système de management de la qualité ;
- le **stockage**, qui doit rester indépendant des choix techniques faits pour l'alimentation et la restitution ;
- la **restitution**, à laquelle on demandera de la souplesse et des facilités de navigation et d'exploration des données.

Les informations doivent être **fiables, pertinentes et précises**. Beaucoup de systèmes décisionnels ont été construits puis ont disparu, parce que les informations restituées n'étaient pas utilisables, ou encore parce que la règle de gestion, permettant de définir tel ou tel indicateur, était consensuelle, mais n'avait jamais été validée.

Absence de donnée collectée

Lorsqu'on s'appuie sur un système de collecte automatisée, il arrive que le système d'alimentation soit indisponible et ne puisse fournir les données d'une ou plusieurs périodes. Quelle décision prendre ? Par exemple, si les données d'une demi-journée manquent ? Peut-on encore faire une consolidation ? Faut-il renoncer à calculer l'indicateur pour cette période ou seulement indiquer que, pour cette donnée-là, les valeurs n'étaient pas complètes au moment de la collecte, mais que sa valeur révèle quand même une tendance. Un indicateur, calculé à partir de données pour lesquelles manque une demi-journée sur un mois, peut rester exploitable, mais peut introduire une erreur trop importante s'il est calculé sur une période plus courte. Faut-il prendre ou non cette valeur en compte ? Il est primordial de répondre à cette question avant de restituer l'information aux échelons supérieurs, surtout si l'indicateur vient en appui d'une décision.

Application à la gouvernance des systèmes d'information

Le CIGREF avait, il y a quelques années, soulevé la question suivante : « Comment le contrôleur de gestion, peut-il appuyer le DSI ? ». Le DSI rencontre dans son environnement des problématiques de décision sur des investissements, des budgets, des coûts, un plan de production, des objectifs, des moyens. Il a besoin d'informations pour décider des actions à entreprendre. Les processus informatiques

sont complexes et risqués et représentent une part importante du budget d'une entreprise. L'idée, c'était de réutiliser tous les principes qui font qu'une entreprise arrive à maîtriser son activité, et de les appliquer à la maîtrise de l'activité informatique. Le DSI a besoin d'avoir une idée de l'activité, de ce qui a été produit sur une année, en termes de fonctionnalités livrées. Cela représente une bonne métrique, si on lui associe le niveau de qualité des livraisons. On sait en principe mesurer les défauts, on a donc là un champ pour appliquer les techniques du décisionnel et du contrôle de gestion.

Il s'agit de définir des indicateurs qui seront cohérents avec les objectifs, de les structurer en un système de mesures et d'appuyer ce système, théorique, par un système technique composé d'un processus d'alimentation, d'une base de données pour le stockage et d'outils de restitution.

Les indicateurs

Si on regarde « ce que mesure » une organisation, on remarque que les indicateurs permettent principalement de mesurer la « chose produite », sa qualité et son efficacité, et ont pour but de donner de la visibilité sur l'activité. L'efficacité est mesurée par le rapport entre les résultats atteints et les moyens engagés. Des plans de production vont prévoir les livraisons, projet par projet. Un certain nombre de livraisons sont attendues dans l'année et font l'objet d'une planification, d'un diagramme de Gantt, outillé par exemple par MS-Project. Si on constate que 99% des livraisons sont intervenues dans le délai du planning, on peut envisager de resserrer un peu les délais et d'utiliser une marge de progression.

Ce qui est particulièrement intéressant, c'est le dialogue qui prend appui sur les indicateurs. Par exemple, l'examen du nombre de livraisons intervenues hors délai peut révéler un problème dans l'estimation de la charge ou révéler les conséquences d'un aléa pendant la phase de développement. La valeur mesurée de l'indicateur va permettre d'amorcer le dialogue avec les chefs de projet et de voir ce qui s'est passé. C'est peut-être un problème de marché notifié trop tard, d'assistance externe arrivée trop tard, de disponibilité du matériel ou encore un bug sur un progiciel que l'on n'a pas pu contourner tout de suite.

Sur la base de ces informations, on pourra prendre les actions correctives nécessaires, recalculer les délais, surveiller la notification des bons de commandes et la passation des marchés.

On pourra avoir une action vis-à-vis des éditeurs de progiciels, leur faire constater, statistiques à l'appui, que les dernières versions livrées présentaient beaucoup de défauts, qui ont demandé un appel au support. On dispose alors d'une base factuelle pour discuter des problèmes rencontrés et envisager les actions permettant leur résolution.

Gestion des évolutions

Une des applications des systèmes décisionnels porte sur la gestion des évolutions du système d'information, via la définition d'indicateurs tels que le nombre de spécifications livrées dans les délais. On peut utiliser un outil de gestion de bugs, comme Bugzilla¹ pour détecter les demandes d'évolutions. C'est un outil qui est intéressant car il enregistre de façon formalisée, déjà paramétrée (et modifiable), des anomalies ou des événements que l'on rencontre pendant la recette. On dispose d'informations telles que la version du logiciel, la date de création du rapport d'anomalie, le nom de son créateur, la sévérité de l'anomalie (majeure, mineure, bloquante), l'environnement dans lequel elle est intervenue (Recette, Qualification, Production). Les informations sont enregistrées, si possible de façon complète, lorsqu'un bug est déclaré. L'outil permet une discussion, par messagerie, et va produire une trace des échanges, qui seront journalisés. Certaines anomalies sont à comptabiliser comme anomalie en cours de recette, on peut ainsi établir un score sur une version, permettant de mesurer sa qualité. Cela permet également d'analyser le nombre et la nature des demandes d'évolution pendant les phases de développement, de recette ou de qualification d'une version donnée, obtenant ainsi un bon indicateur de maturité des spécifications. Entre anomalie et évolution, il existe cependant un cas qui pose problème : lorsque les spécifications ont bien été respectées à la lettre et qu'il ne s'agit pas d'une anomalie, mais que l'état actuel des spécifications ne permet pas au logiciel de fonctionner. Il faut dans ce cas rectifier ou compléter le dossier pour que l'application puisse fonctionner de façon satisfaisante pour l'utilisateur.

De ce fait, trois cas peuvent être rencontrés :

- on corrige l'anomalie car il s'agit bien d'une non conformité vis-à-vis des spécifications,
- on demande un nouveau cahier des charges afin de prendre en compte l'évolution,
- on examine l'opportunité de prendre en charge le complément selon son caractère : bloquant, majeur ou mineur.

Il est important de disposer, par version, du nombre de cas découverts, classés selon cette typologie et de les classer selon le niveau de sévérité.

Ces constatations vont déboucher sur des actions correctives ou d'amélioration qui ne sont possibles que si l'on dispose des bonnes informations en entrée.

Prenons un exemple : si l'on a 15 demandes d'évolution au moment de la recette, à cheval entre la recette et le développement, cela peut signifier que les spécifications ont été faites trop rapidement, ou que le spécificateur n'était pas vraiment aligné sur la

logique de l'entreprise et les attentes de l'utilisateur. L'utilisateur suivra en effet une logique « métier » et recettera l'application selon cette logique. Si la recette fait apparaître un nombre important de demandes d'évolution, c'est probablement que l'application n'est pas alignée par rapport au métier. On peut se poser la question de la valeur ajoutée par cette version, qui s'éloigne alors de son objectif premier : livrer une application qui soit conforme aux normes et standards tout en étant alignée avec les processus de l'entreprise.

Conclusion du conférencier

Dominique Mollard indique, en conclusion, que des efforts ont été faits pour formaliser le domaine d'application du pilotage de la performance à la gouvernance des systèmes d'information. Ce sujet touche aux moyens et à leur utilisation pour atteindre des objectifs. Il y a cependant un aspect psychologique, lié aux pratiques, qui impose la conduite du changement.

Un système décisionnel pour le suivi des projets, en appui de la gouvernance des systèmes d'information, aura une échelle beaucoup moins importante qu'un système décisionnel pour le pilotage de la performance qui soit institutionnel. Par contre, il est probable que l'un sera imbriqué dans l'autre. Les indicateurs de qualité informatique que l'on mesure avec l'outil appliqué à la gouvernance des systèmes d'information, reflétant par exemple le nombre d'heures potentiellement perturbées par application sur une année, seront remontés dans le système décisionnel institutionnel. L'informatique n'est finalement qu'une des fonctions de l'entreprise, une fonction de support des processus métiers, tout comme la vente, ou les achats.

Il est donc nécessaire de mettre en place un outil pour le décideur, au plus haut niveau, qui lui permettra de connaître l'état de santé de la production informatique, via les indicateurs sur la qualité ou le fonctionnement des services d'assistance. Il est également nécessaire de mettre en place un système un peu plus précis, centré sur l'informatique, qui permette de mesurer à un niveau de détail plus fin les mêmes processus et de donner au DSI une image de ce qui se passe.

Et puis, dernier point, il y a une convergence entre la logique de construction de ces systèmes décisionnels et des référentiels tels que CMMI. Au niveau des domaines « Assurance Qualité » et « Mesures et Analyses », le fait d'avoir défini des indicateurs, d'être outillé pour en tracer la valeur, et de les restituer constitue un bon point de départ par rapport au niveau de maturité que l'on vise. Le système décisionnel n'est pas en contradiction avec cette logique-là, bien au contraire. Il fournit déjà des éléments qui permettent de démontrer de façon précise, fiable, pertinente et le moins possible discutable que l'on a atteint un niveau donné.

¹ Bugzilla est un logiciel libre, développé et utilisé par la fondation Mozilla, pour le suivi des bugs.

Quelques questions

Comment les données peuvent-elles être fiables, pertinentes, précises ?

La réponse de Dominique Mollard semble évidente : en faisant les bons choix de conception dès le démarrage du projet. Il y a différentes façons d'alimenter un système décisionnel : on peut prendre contact avec l'équipe responsable du système et lui demander de produire un fichier ou un flux par l'intermédiaire d'une interface. Deux façons de procéder sont envisageables :

- aller chercher les données directement dans l'application, en développant un collecteur, interroger sa base de données et stocker les données visées ;
- ou alors, négocier un contrat en précisant les données nécessaires en définissant un lot de collecte, par exemple des enregistrements qualité et le format d'échange.

L'objectif est de rechercher un format de données qui facilite la collecte et l'intégration pour le client et qui minimise la charge de production pour le fournisseur. La contractualisation permet d'atteindre un niveau élevé de qualité dans la fourniture des données. Pour un fichier d'échange au format texte délimité, par exemple, des contrôles sur les zones pourront être spécifiés et mis en place dans l'application productrice. L'utilisation du langage XML permet d'étendre ces contrôles au typage des données, à leur caractère obligatoire ou non et de mieux contrôler leur structure. En complément, des procédures spécifiques permettront de détecter les erreurs lors de l'alimentation et d'identifier les applications sources de données pour lesquelles des erreurs récurrentes sont constatées. Par exemple, la collecte de l'application Digital¹ s'appuie sur 33 applications d'environnements très différents produisant des données au moyen de fichiers dont le format est en « texte délimité ». Ce choix technique est bien adapté aux applications traditionnelles et a fait l'objet d'un très large consensus. La mise en place de contrôles qualité lors de l'alimentation a également un effet préventif car le fournisseur sera incité à assurer la qualité de ses productions pour éviter les réfections. Des points de contrôle qualité peuvent être également définis aux différentes étapes de l'alimentation. Par exemple, sur la présence des lots de collecte, le format des données qu'ils contiennent, lors de l'insertion des données dans la base de données et au niveau des outils de restitution.

Pour le contrôle de gestion, cela paraît simple. Comment, dans une DSI, quand on n'a pas d'indicateurs, avoir cette fiabilité des données sans la lourdeur ?

En tant que responsable d'équipe « projet », Dominique Mollard nous indique qu'il alimentait ces

indicateurs, tout en étant également maître d'œuvre de l'outil qui permettait de les restituer au Directeur Informatique. Le système de collecte doit être défini, décrit, contractualisé et explicable. Les utilisateurs savent que les données qu'ils fournissent seront transformées, consolidées et agrégées dans un tableau de bord, et devront s'intéresser aux règles de gestion, pour valider les données fournies. La validation est essentielle. Elle est en même temps très compliquée parce qu'elle s'effectue sur le tableau de bord, qui est le produit fini et suppose de maîtriser toutes les règles d'agrégation, tous les niveaux. Une étape de pré-validation permet de simplifier ces travaux et de gagner du temps en intervenant au niveau de la base de données, puisque les outils de restitution vont aller chercher l'information dans l'espace de stockage. Sur le projet Digital, 150 indicateurs sont restitués, dont 15 de très haut niveau et très sensibles. Dans la base, on retrouve 650 données élémentaires collectées auprès de 800 unités opérationnelles. Pour la recette, on procède domaine par domaine selon deux niveaux :

- sur la base de données, à l'aide de requêtes, dont on compare le résultat avec celui prévu pour les calculs d'agrégation ;
- lorsque les processus d'alimentation (ce sont eux les plus importants) sont validés dans la base de données, les travaux s'étendent alors aux restitutions. On teste élément par élément. Pour afficher un tableau de bord prédéfini, l'outil fait l'interface avec une structure méta-données posée sur la base de données. La procédure est assez simple : on peut reconstituer le tableau de bord avec quelques requêtes, récupérer le résultat, puis vérifier ce qui est affiché par l'outil de restitution. Ensuite, les objets multidimensionnels, cubes ou structures OLAP, peuvent être validés avec toutefois une difficulté plus grande du fait du nombre de croisements possibles. Ces croisements supposent des calculs intermédiaires pour lesquels on utilise aussi des fonctions de la base de données, cube Oracle, par exemple, pour avoir des éléments de comparaison. Ces informations sont comparées avec ce que l'outil décisionnel a restitué. Les anomalies les plus importantes sont rarement constatées sur le système décisionnel lui-même, mais plus fréquemment sur le logiciel qui alimentait les données. Un écart de quelques pour-cents par rapport à la valeur attendue n'est pas facilement détectable dans les restitutions : il est plus facile de le détecter lors de l'alimentation.

Assurer la qualité d'un système décisionnel implique la définition de métriques sur l'alimentation. Les indicateurs essentiels, qui appuient des décisions stratégiques, doivent être sécurisés, car le processus d'alimentation de ces indicateurs sera critique. Les indicateurs ne sont pas tous au même niveau. Certains sont « saillants » du fait des enjeux associés et nécessiteront de définir des métriques et de faire un peu de métrologie sur les sources de données :

¹ Digital Pilote est un éditeur français de solutions décisionnelles, racheté par Sage en 2006.

- Ainsi, il est possible de caractériser la **criticité** d'une l'information, par exemple lorsqu'elle entre dans la composition d'un indicateur qui est essentiel pour le pilotage de l'entreprise à un niveau stratégique.
- Une autre métrique qui peut être intéressante, c'est la **variabilité** de la source qui permettra d'anticiper les évolutions.
- Ensuite la **complexité** est intéressante à mesurer. L'information recherchée pourra être directement mobilisable ou résulter d'un calcul ou d'un algorithme.
- Enfin, la **priorité** de l'indicateur pour l'organisation, pour l'alignement de la stratégie, permet de rendre compte du gain attendu par sa mise en place. S'agit-il d'une information qui manque vraiment pour la prise de décisions, ou bien est-elle disponible par ailleurs ?

On a des indicateurs qui reflètent le passé pour prendre des décisions pour l'avenir. Qu'en est-il du modèle prédictif ? Qui le définit et comment contrôle-t-on sa qualité ?

Les prévisions peuvent être réalisées par des outils de data mining utilisant les données du système décisionnel. Ces outils implémentent des algorithmes traditionnels, tels que les modèles de régression, ou heuristiques, tels que les réseaux neuronaux. La difficulté est le niveau de compétence requis pour définir les règles de gestion pour le calcul des agrégats et le paramétrage du modèle. Des compétences spécifiques sont également nécessaires pour l'interprétation des résultats. Les données présentes dans les entrepôts de données doivent être transformées pour correspondre à la logique des outils de data mining. En effet, pour un sujet donné, plusieurs lignes peuvent être présentes dans la base, chacune correspondant à un événement. Or le modèle utilisé est souvent fondé sur une observation rendant compte de façon synthétique du comportement d'une variable. Il est possible d'utiliser un outil OLAP pour préparer un tableau de données à partir du système décisionnel puis d'utiliser les données obtenues pour alimenter l'outil de data mining.

Ces outils nécessitent trois jeux de données : un pour l'apprentissage du modèle, un pour sa validation lors des tests et le dernier correspondant aux données de production à traiter.

Mais c'est l'éditeur du logiciel qui devrait faire ces vérifications de l'algorithme ! Et ensuite c'est un problème de confiance en l'éditeur. Si les données de départ sont correctes, vous savez ce que vous attendez comme résultat...

En effet, les réseaux neuronaux sont implémentés dans l'outil et les procédures de tests ont permis de les valider. Cependant, lors de l'interprétation des résultats, il est essentiel de savoir si le processus a bien été respecté. Et quelle conclusion peut-on en tirer ? On ne peut pas utiliser directement les données de l'entrepôt de données car on enregistre

des chroniques pour faire des suites temporelles alors que les outils de datamining s'appuient sur des observations. On construit donc un datamart (un magasin de données) à côté du datawarehouse (entrepôt de données) de façon particulière pour que l'outil puisse utiliser les données sous forme tabulaire, c'est-à-dire avec une ligne par observation. Le data mining a un côté exploratoire, qui est très difficile à coder dans les procédures d'alimentation. Inversement on peut avoir des masses de données et ne pas savoir quelle décision prendre. Les statisticiens font du requêtage à partir d'une masse d'informations disponibles. Les bases se mesurent alors en téraoctets.

L'autre logique, c'est la logique de l'information choisie car l'information a un coût d'acquisition. Dans l'outil décisionnel, on peut calculer l'écart entre deux valeurs, par exemple entre l'objectif et le résultat, ou calculer un ratio, ou appliquer un algorithme. Quand les données nécessaires dépendent d'une source d'information dont on ne dispose pas, l'ajout de cette source constituera le coût le plus important et nécessitera de prévoir un contrat applicatif et une structure de flux qu'il faudra ensuite intégrer. Une analyse de la valeur apportée par l'indicateur et sa comparaison avec les coûts d'obtention des données nécessaires permettra d'en apprécier la pertinence.

On peut avoir des données utiles à différents services. Dans ce cas-là, est-ce que tout garder n'est pas la meilleure solution ? Une grande banque refait tous ses systèmes de synthèse. Ils ont décidé de mettre dans un « pot » la collecte des données pour les besoins de la comptabilité, pour le contrôle de gestion, les risques, les besoins réglementaires, les besoins du pilotage de l'activité avec des datamarts sectorisés.

On crée une donnée parce qu'elle entre dans la composition d'un indicateur qui correspond à un besoin par rapport à un objectif. On peut donc en déduire les données élémentaires qu'il faut collecter. On construit un dictionnaire de données, on liste les indicateurs, les données nécessaires, puis on applique les métriques. On peut donc en déduire qu'une donnée est utilisée par plusieurs indicateurs différents et de ce fait a une valeur qui justifie sa collecte et son intégration dans le calcul d'indicateurs. ▲

Véronique Pelletier
home.pelletier@club-internet.fr

Dominique Mollard
dominique_mollard@yahoo.fr

Quelques précisions

Diagramme de Gantt : outil permettant de modéliser la planification des tâches nécessaires à la réalisation d'un projet.

OLAP ou cube (d'après Wikipedia) : **OnLine Analytical Processing (OLAP)** désigne les bases de données multidimensionnelles (aussi appelées cubes ou hypercubes) destinées à l'analyse et il s'oppose au terme OLTP qui désigne les systèmes transactionnels. Ce terme a été défini par Ted Codd en 1993 au travers de 12 règles que doit respecter une base de données si elle veut adhérer au concept OLAP.

Les 12 règles de Codd sont les suivantes :

- vue conceptuelle multidimensionnelle ;
- transparence ;
- accessibilité ;
- constance des temps de réponses ;
- architecture client serveur ;
- indépendance des dimensions ;
- gestion des matrices creuses ;
- accès multi-utilisateurs ;
- pas de restrictions sur les opérations inter et intra dimensions ;
- manipulation des données aisée ;
- simplicité des rapports ;
- nombre illimité de dimensions et nombre illimité d'éléments sur les dimensions.

Ce concept est appliqué à un modèle virtuel de représentation de données appelé cube ou hypercube OLAP.

Sigles, Termes, abréviations

CIGREF	Club Informatique des Grandes Entreprises Françaises
CMMI	Capability Maturity Model Integration
CNAM	Conservatoire National des Arts et Métiers
Data mining	Exploration de données (littéralement forage de données)
Datamart	Magasin de données
Datawarehouse	Entrepôt de données
DSI	Directeur du Système d'Information
MOE	Maîtrise d'Œuvre
OLAP	On Line Analytical Processing
PME	Petite et Moyenne Entreprise
SID	Système d'Information Décisionnel
XML	eXtensible Markup Language : langage de balisage extensible

Bibliographie

Dominique Mollard - Systèmes décisionnels et pilotage de la performance - Hermès-Lavoisier