

9. Inventaire des données

Après avoir, dans les chapitres précédents, établi les bases pour l'activité de modélisation des données, nous entrons à présent (enfin!) dans le vif du sujet: comment procéder pour établir un modèle des données correct, qui corresponde aux besoins de l'organisation ou de l'application, qui soit capable d'évoluer en cas de modifications futures de ces besoins.

La première étape de cette modélisation est en général la réalisation d'un **inventaire des données**. On entend par là le fait de récolter l'ensemble des informations qui devront être gérées dans la base. Par contre on ne s'occupe pas encore, lors de cette étape, de la manière dont ces informations seront organisées dans une base de données.

Cet inventaire des données doit-il être établi de manière formelle, complète, documentée et validée avant de commencer la conception de la base proprement dite? Cela dépendra avant tout de l'étendue du projet, du nombre et du genre de personnes impliquées et de leur expérience.

Nous ne sommes pas du tout opposés à l'idée d'établir, si les circonstances l'exigent, un inventaire des données formel, documenté et mis à jour en permanence par la suite. Notre souci ne porte toutefois pas sur la méthodologie et l'organisation du projet, mais sur le résultat à atteindre: à savoir une structure de base de données correcte. À notre avis, un inventaire formel s'impose surtout s'il doit être obligatoirement validé par les utilisateurs. Se pose pourtant ensuite la question de maintenir cet inventaire constamment à jour par la suite en parallèle avec le modèle des données qui comprendra non seulement tout ce que contient l'inventaire, mais bien davantage encore, à savoir la structure de ces données dans la base.

Nous considérons ici l'inventaire des données non pas comme un document, mais comme une étape durant laquelle les personnes impliquées, qu'ils s'agisse de représentants des futurs utilisateurs ou d'informaticiens qui n'interviendront que pendant la durée du projet, rassemblent tout ce qui devra figurer dans la future base. Pour des personnes qui ne connaissent pas le domaine couvert par le projet, ce sera également un moyen de se mettre au courant. Pour les personnes qui connaissent ce domaine comme leur poche, un moyen de rassembler leur savoir et de le mettre au clair.

Nous avons déjà parlé au chapitre 4 des rôles joués par différents types de participants à un projet de modélisation. Le responsable de l'établissement de l'inventaire des données est évidemment, à ce stade déjà, l'**administrateur des données**. Il est vraiment la personne qui doit avoir, dès le départ, une vision totale sur l'information contenue dans la base de données. Il est donc important que ce travail, malheureusement souvent cléricale, ne soit pas relégué à une personne n'ayant pas la vision du projet ou les compétences nécessaires.

Contenu de l'inventaire des données

Les types d'objets à décrire dans la base de données. Ils deviendront, quelques fois après transformation, les types d'objets contenus dans le modèle de la future base de données. Nous verrons toutefois qu'il n'existe pas forcément dans tous les cas une identité parfaite entre les types d'objets de l'inventaire et ceux du modèle terminé.

Exemples de types d'objets: le compte, le client, l'employé, la commande, l'article, etc.

Les informations rattachées à ces types d'objets. Chacun des types d'objets découverts sera complété par une liste d'éléments de données qui le décrivent. Ces éléments de données deviendront les attributs des types d'objets dans le futur modèle. Ici encore, il n'existe toutefois pas toujours une identité parfaite entre les éléments de données découverts lors de l'inventaire et les attributs du modèle définitif.

Exemples d'éléments de données liées aux types d'objets découverts:

Employé: nom, prénom, date de naissance, no d'employé, adresse, date d'entrée, etc.

Article: code article, désignation en différentes langues (4 aujourd'hui, davantage peut-être dans le futur), prix de vente, catégorie article, etc.

Important:

Ne pas faire d'inventaire purement atomique des données! Nous entendons par inventaire atomique, une liste des éléments de données détachés des types d'objets qu'ils décrivent. À éviter absolument donc un fichier global de tous les attributs, triés alphabétiquement par exemple, sans lien avec le type d'objets auquel ils se rattachent. On appelait par le passé un tel inventaire *dictionnaire des données*. Procédant de cette façon, on perd une partie importante de l'information récoltée. À notre avis, ce serait faire le travail deux fois et nous ramener aux exercices de normalisation d'antan.

Un dictionnaire des données peut bien sûr être produit pour des besoins de documentation, mais ceci après coup et automatiquement à partir du modèle terminé. Mais pas durant la phase de modélisation! L'expérience nous a montré que dans les projets où un tel dictionnaire est créé dès le départ, son entretien était souvent confié à un collaborateur junior rapidement démotivé par la tâche. Du fait des changements qui interviennent continuellement et qui ne se reflètent que tardivement dans le dictionnaire, ce dernier ne pourra pas servir de base de travail fiable aux personnes chargées de la modélisation.

Faut-il se soucier de la normalisation?

Lorsqu'on établit l'inventaire des données, on ne s'occupe absolument pas de telles questions. À ce stade, on n'est pas en train de définir les tables finales de la future base de données, mais de découvrir les concepts avec lesquels les utilisateurs travaillent aujourd'hui, ou travailleront demain avec le futur système. **Le souci principal, pour le moment, est d'avoir compris ce dont les utilisateurs ont besoin et, surtout, d'être complet.**

L'exemple ci-dessus (éléments de données liés au type d'objets *Article*) illustre cet aspect: on y parle de *désignation en différentes langues*. L'expérience nous apprend qu'il faudra probablement créer une table des désignations d'articles en plusieurs langues, mais nous traiterons ce problème plus tard.

Mais ceci n'empêche évidemment nullement d'anticiper le futur, particulièrement si l'on possède déjà une certaine expérience du domaine couvert. Un connaisseur des systèmes de gestion commerciale saura par exemple qu'une facture se décompose toujours en données d'entête, différentes lignes de détail et un pied de facture. Rien n'empêche de faire figurer cette structure dès le départ dans l'inventaire.

À ce stade du projet, on ne se soucie pas non plus des performances du futur système ou du volume de données impliqué.

Comment documenter l'inventaire?

L'inventaire ne constituant qu'une étape préliminaire, sa documentation est sans importance.

Nous proposons des fiches, sur papier ou dans l'ordinateur. Créer par exemple une fiche par type d'objet découvert avec la liste des éléments de données, futurs attributs, qui s'y rattachent.

Au fur et à mesure que le modèle sera conçu, on pourra détruire les fiches des types d'objets déjà traités, à condition bien sûr que tous les éléments de données aient trouvé leur place dans le modèle.

Lorsque le modèle sera terminé, l'inventaire des données aura fini de jouer son rôle et aura perdu toute utilité. Il ne sera d'ailleurs probablement plus à jour et continuer de le gérer constituerait un travail fastidieux et inutile. Si une documentation sur papier de la base de données est requise, elle sera établie automatiquement à partir du modèle enregistré dans l'ordinateur et donc constamment à jour.

Comment effectuer l'inventaire des données?

Nous voyons deux sources principales de sources pour l'inventaire des données:

- Analyser tous les documents et fichiers existants dans l'entreprise
- Rencontrer les futurs utilisateurs et enregistrer leurs besoins.

Analyse des documents et fichiers existants

L'analyse des documents et fichiers existants est particulièrement importante lorsqu'il s'agit de moderniser un ancien système ou d'automatiser une procédure effectuée manuellement jusqu'à ce jour.

Il est évident que les documents que produira le nouveau système ne seront pas identiques à ceux de l'ancien, sinon le projet n'existerait pas. Par contre, les informations qu'ils contiennent constituent une bonne base de départ. Il en est de même des fichiers existants, informatisés ou non.

On procédera donc à un inventaire détaillé, dans l'esprit de ce qui est décrit plus haut, de tout ce qui existe sous une forme ou une autre.

Fichiers sur papier ou informatisés, listes, rapports, statistiques, documents contractuels, etc.

On "épluchera" ainsi, dans un système commercial, le fichier existant des articles, les confirmations de commandes, bulletins de livraison et factures envoyés aux clients, les statistiques produites. Dans une application pour une caisse maladie, le fichier des assurés, les polices, les factures de primes, les décomptes de prestations. Dans un système de gestion du personnel, un contrat de travail, un certificat de salaire, une attestation de revenu, etc. Dans un système de fabrication, les gammes opératoires, les nomenclatures, les bons de travail, les commandes d'atelier, etc.

Souvent, il faudra se faire remettre différentes versions de ces documents. Une facture à un client final ne se présentera par exemple peut être pas de la même manière qu'une facture à un revendeur.

Entretiens avec les futurs utilisateurs

Les entretiens avec les futurs utilisateurs (généralement leurs représentants) permettront:

- De comprendre le contenu des documents et fichiers analysés
- D'enregistrer leurs vœux au sujet de ce qui n'était pas géré par le passé, mais devra l'être dans le futur et manque donc dans le système actuel.

L'art de l'entretien avec les utilisateurs est très difficile et n'est franchement pas donné à tout le monde. Il est très rare que ces discussions réunissent, tant du côté des informaticiens que des utilisateurs, exclusivement ou principalement des personnes extrêmement douées pour la communication et la compréhension des idées d'autrui. Souvent il est difficile de se comprendre, tant parce que celui qui possède la connaissance du domaine ne sait pas l'exprimer, que parce que celui qui n'en est pas familier doit s'y plonger.

Durant de tels entretiens, les spécialistes de la construction de systèmes (informaticiens) écouteront en premier lieu les spécialistes du domaine (les utilisateurs). L'informaticien doit vraiment se mettre sur écoute, enregistrer, et absolument éviter à ce stade déjà de construire des solutions dans sa tête ou, pire encore, essayer d'en imposer aux utilisateurs. Des remarques telles que: "je vois ce que vous voulez dire, mais c'est ainsi qu'il faut faire..." ou encore "c'est malheureusement impossible, la normalisation nous oblige à..." ou "ce sera beaucoup trop lent, nous ferons donc ainsi..." sont totalement inappropriées à ce stade.

Un entretien ne suffit en général pas, sous risque de créer ensuite un système ne correspondant pas aux besoins réels. On impliquera au contraire les futurs utilisateurs à tous les stades du développement, profitant de ces rencontres pour vérifier les informations reçues précédemment. Un bon concepteur de système doit être aussi méfiant qu'un détective: on ne lui révèle pas forcément l'entière vérité dès le départ, ici sans mauvaise volonté. Cette vérification se fera au plus tard dès qu'une première version du modèle des données sera établie. La structure de ce modèle sera validée en commun, fournissant aux utilisateurs toutes les explications nécessaires afin qu'ils puissent porter un jugement fondé sur la structure proposée. On créera peut être des masques de saisie ou des exemples de listes pour vérifier que tous les éléments requis y figurent.

Peut-on mélanger analyse des données et analyse des fonctions?

Les années 1980 ont été marquées par l'avènement des bases de données en informatique. Ce fut une révolution totale. Auparavant, chaque programme recevait des fichiers en entrée et en produisait de nouveaux qu'il transmettait à d'autres programmes. Ce système fut alors remplacé par une nouvelle architecture comprenant une base de données centrale unique, exploitée par les programmes et alimentée par eux. Une nouvelle science était née: concevoir une base de données. Rapidement, les spécialistes plaidèrent pour que la base de données, recueil de toutes les informations gérées par l'entreprise, soit conçue **indépendamment** des programmes qui l'alimentent et l'exploitent, programmes qui sont réduits au rôle d'interface entre l'utilisateur et la base. La majorité des méthodes de développement, Merise n'y fait pas exception, prônaient donc de séparer nettement **analyse des données** et **analyse des traitements**. Ces deux tâches étaient même quelques fois confiées à des équipes différentes.

Cette approche n'a rien perdu de sa validité. Mais, avec la disponibilité sur le marché de progiciels performants dans tous les domaines, il est rare que l'on procède aujourd'hui encore à la conception d'une base de données d'entreprise unique et complète. En général, on acquiert plutôt un progiciel livré avec sa base de données. La conception d'une base de données se fait par contre encore pour de petites applications spécialisées, ainsi que pour des systèmes d'aide à la décision.

Dans le cadre de petites applications, il est difficile d'expliquer aux utilisateurs que l'on ne doit pour le moment pas s'intéresser aux fonctionnalités du futur système: ce sera finalement leur outil de travail et c'est cela qui les intéresse avant tout.

Tout en cherchant à orienter au maximum la discussion sur les types d'objets à gérer, on n'empêchera donc pas les utilisateurs d'exprimer leurs désirs concernant les fonctionnalités du futur système. Il suffira de les noter et de compléter un cahier des charges qui aura probablement déjà existant.

Travail en équipe

La modélisation des données se fait toujours en équipe, nous l'avons déjà dit. C'est particulièrement important pour l'inventaire des données. L'idéal serait de disposer de deux personnes du côté des informaticiens et deux du côté des utilisateurs. Il est important que tout ce qui est exprimé enregistré soit confirmé par quelqu'un d'autre, que plusieurs interlocuteurs se trouvent sur la même longueur d'onde. Il faut particulièrement se méfier d'une paire unique informaticien – utilisateur: complices, ils risquent à deux trop aisément de créer une solution qui sera rejetée par les autres utilisateurs.

En ce qui concerne les utilisateurs, faire attention aux personnes qui ont déjà une idée (trop) forte du futur système et veulent l'imposer à leurs collègues. S'il n'est pas possible de les neutraliser, peut-être parce qu'ils ont la meilleure connaissance du domaine ou parce que leur participation est imposée par la direction, faire au moins participer au débat une autre personne capable de servir de contrepoids.


N'accepter parmi les informaticiens que des personnes possédant une grande capacité d'écoute et de compréhension d'autrui. L'informaticien qui croit posséder la solution au départ et ne cherche qu'à convaincre l'utilisateur de son point de vue n'est pas utilisable ici, nous l'avons déjà dit.

Un travail itératif

Tout développement se fait aujourd'hui de manière **itérative**, impliquant constamment les utilisateurs.

En aucun cas, les développeurs ne remettront aux utilisateurs une base de données ou un système qui ne correspondent pas à leurs besoins. Des remarques telles que "C'est ce que vous avez demandé" ou "C'est ce que vous avez signé" ne sont jamais admissibles. Si le système final ne correspond pas aux besoins, la faute en incombe **toujours** à l'informaticien qui n'a pas su ou voulu assez tôt régler les différends avec l'utilisateur ou client, ou qui n'a pas été suffisamment à son écoute. Que l'utilisateur change d'avis durant le développement est normal et non pas l'exception. Le processus d'itération est justement là pour tenir compte de cette réalité.

Exemple d'inventaire des données

					
Happy Biker Sàrl Route de la Venoge 32 1026 Denges 021 777 33 45 www.happybiker.ch					
Atelier Vélo Henri Cartona Chemin du Crêt 33b 1066 Epalinges					
FACTURE 2010_0027					
Denges, 20 janvier 2010					
Votre commande: 22326		Réf: arbf		Du: 13.1.2010	
Notre livraison: 2010_L0017				Du: 19.1.2010	
No art	Description	Quantité	Prix unitaire	Rabais	Montant
223564	Pignon 6"	2	8.35	20%	13.35
342993	Selle cuir Wabadu	5	50.25	10%	226.15
185663	Guidon Happy	3	27.30	10%	73.70
997995	T-shirt Happy Biker	10	25.00	0%	250.00
Total marchandise					563.20
Soumis TVA					563.20
TVA 7.6%					42.80
Total Facture					CHF 606.00 =====
Conditions de paiement: net à 30 jours TVA no 345 879					

Nous enregistrons sur cette facture les éléments (types d'objets) suivants avec leurs composants:

Facture avec: client (rouge), entête de facture (noir), plusieurs lignes de facture (vert), bas de facture (violet)

Entête facture (noir) avec: no facture, date facture, no commande client, référence client, data commande client, no livraison, date de livraison.

Ligne facture (vert) avec: no article, libellé article, quantité, prix unitaire, rabais, montant

Article (orange) avec: no article, libellé, prix unitaire

Bas de facture (violet) avec: total marchandise, total soumis TVA, taux TVA, montant TVA, total facture, conditions de paiement avec escompte éventuel, délai de paiement

Remarques concernant l'exemple d'inventaire des données:

En effectuant l'inventaire de ce document:

- Nous ne nous posons pas de questions au sujet de la normalisation
- Nous ne nous posons pas de questions concernant le fait de répéter des éléments déjà découverts précédemment
- Nous ne nous soucions pas de la description des liens et associations entre objets
- C'est un pur travail de découverte et de consignation
- C'est un travail "intelligent". Nous avons le droit de "lire entre les lignes", d'extrapoler et de faire entrer en jeu notre expérience ou notre connaissance du métier. Exemple: nous avons consigné ici la ligne de facture avec tous ses éléments, puis nous avons également consigné l'article avec tous les éléments qui ne sont pas propres à la ligne de facture
- Rien ne nous empêche d'ajouter des éléments qui ne figurent pas sur le document, mais viennent à l'esprit lors de l'analyse. Exemple: dans cette facture les conditions de paiement stipulent "paiement net", mais il pourrait y avoir d'autres modes avec des escomptes. En discuter avec les utilisateurs responsables ou exiger d'autres exemples de factures.