

LES RÉFÉRENTIELS DU SYSTÈME D'INFORMATION

**Données de référence
et architecture
d'entreprise**

Joël Bizingre
Joseph Paumier
Pascal Rivière

Préface de Dominique Vauquier

DUNOD

Illustration de couverture :
South American Panoramics © digitalvision

Tout le catalogue sur
www.dunod.com



Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.

Le Code de la propriété intellectuelle du 1^{er} juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements

d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour

les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée. Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du

Centre français d'exploitation du droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).



© Dunod, Paris, 2013
ISBN 978-2-10-059874-8

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

Préface

La notion de référentiel véhicule l'idée de bien commun. Cette expression, « bien commun », est prise ici au sens d'actif partagé, de patrimoine mis en commun. Elle peut être comprise aussi au sens moral, tant il est vrai que le partage renforce et fonde la communauté. À partir de là, se profile toute la dimension que peut prendre le thème des référentiels : au sein de l'entreprise, dans la chaîne de valeur complexe associant de multiples partenaires et, bientôt, à l'échelle de la société. Le livre que vous avez entre les mains est une enquête dans cette dimension, enquête qui nous mène des besoins de l'entreprise jusqu'aux perspectives des *big data* et *open data*, en passant par l'interopérabilité.

Ce n'est pas trop de dire qu'avec les référentiels, se joue une dynamique qui, bien conduite, peut changer le destin d'une entreprise. Tout d'abord, penser en termes de référentiels revient à se hisser au niveau du système, à prendre l'entreprise comme un tout et à exhausser la visée de la conception. Ensuite, on aperçoit vite que le partage des informations n'est possible qu'à la condition d'en maîtriser le sens. Là encore, la conception doit se grandir et se montrer capable d'embrasser non seulement l'informatique mais également la connaissance du métier.

Parmi les multiples dénnotations du terme « référentiel »¹, ce livre s'attache à analyser les référentiels de données. Toutefois, on observe une curieuse dépendance entre les différents types de référentiels :

Pour bâtir les « bons » référentiels de données, il est indispensable de commencer par les « bons » modèles, non pas des modèles de portée locale (au niveau du projet), mais des modèles de portée globale, qui se préoccupent du bien commun.

De tels modèles, obéissant à des règles précises et se conformant à une vision esquissée par l'architecture, viennent enrichir le Référentiel de description de l'entreprise.

Or, pour réussir ces modèles et ces architectures, il est nécessaire d'asseoir les compétences sur un référentiel de pratiques.

1. La définition qu'en donne le Thesaurus de la méthode Praxeme est : « ensemble d'objets ou d'informations partagé par une communauté » (cf. <http://www.praxeme.org>, rubrique Thesaurus).

« Référentiel de pratiques » est un autre nom pour « méthode ». La méthode reste le seul moyen pour unifier les pratiques et pour maintenir le niveau d'exigence requis. Cette conviction motive l'initiative pour une méthode publique, qui produit la méthodologie d'entreprise Praxeme.

Il est un autre sens du terme « référentiel », emprunté à la linguistique et que l'on trouve, par exemple, dans l'expression « fonction référentielle du langage »¹ : le fait, pour un signe, de désigner les objets du réel ou de l'univers du discours (lequel peut être imaginaire). Il en va de même avec les référentiels informatiques et, de façon générale, avec toute donnée informatique. Les données n'ont de valeur que pour autant qu'elles « désignent », dénotent les objets du réel et les notions manipulées par les acteurs. C'est encore plus vrai dans le cas des référentiels, de par leur position dans le système d'information. Dès lors, se pose la question de la qualité des données, question critique à laquelle nos auteurs consacrent un plein chapitre, à juste titre.

Le lecteur trouvera, dans cet ouvrage, des distinctions très utiles, issues d'une longue pratique, ainsi qu'un examen systématique des éléments liés à la problématique du partage des données dans l'entreprise et au-delà de l'entreprise. Il bénéficiera de l'effort de classification qui traduit l'exigence scientifique et satisfait la nécessité pratique. Les auteurs, tous praticiens, conjuguent leur expérience dans ce livre nourri d'exemples précis et traversé du souci des enjeux évidents pour la transformation des entreprises comme pour la modernisation de l'action publique.

Quelle que soit la motivation de la transformation – stratégique, métier, organisationnelle ou informatique –, les référentiels sont au cœur de la démarche : ils cristallisent la vision et témoignent d'un profond changement d'approche, à rebours des approches fonctionnalistes. Leur simple présence ouvre la voie vers une simplification radicale du système et offre des possibilités nouvelles d'action. On peut dire que, d'une certaine façon, l'ambition d'une transformation se mesure à la qualité et à la réalité des référentiels qui forment le socle du système informatique. Dans une démarche pratique, progressive et *bottom-up* de transformation, les référentiels de données préparent aux référentiels d'objets qui deviendront le cœur palpitant du système entreprise, le noyau essentiel à partir duquel articuler de nouvelles offres et de nouvelles pratiques.

Si, à l'inverse, on opte pour une démarche radicale, *top-down*, de transformation, on retrouvera également les référentiels, par cet autre chemin : d'abord, isoler les fondamentaux du métier et dégager l'aspect sémantique de l'entreprise (ce qui demande un gros effort d'abstraction, effort payé de retour par l'innovation conceptuelle) ; puis, structurer cette connaissance, c'est-à-dire identifier les domaines d'objets ; enfin, architecturer le système informatique en l'alignant sur cette structure de la connaissance métier.

1. Voir, par exemple, le *Dictionnaire encyclopédique des sciences du langage*, Oswald Ducrot, Tzvetan Todorov, Éditions du Seuil, 1972.

En effet, les domaines d'objets¹ fournissent le point de départ pour bâtir les référentiels dans le système informatique, référentiels de données et, plus encore, fournisseurs de services au sens de SOA (*Services Oriented Architecture*).

Ainsi, ces deux chemins, *bottom-up* ou *top-down*, aboutissent au même constat : la nécessité des référentiels pour construire les fondations du système d'information.

Cette idée simple, comme toutes les idées simples, a contre elle toute la tradition :

- les habitudes de travail caractérisées par le mode projet,
- les schémas de pensée dominés par l'approche fonctionnaliste,
- le poids des systèmes existants qui écrase toute velléité de reconception et décourage l'imagination architecturale.

Face à ces obstacles, ce livre a le mérite d'affirmer avec force et clarté une vision nécessaire des systèmes d'information... en vue du bien commun.

Dominique Vauquier
Architecte d'entreprise
Fondateur du Praxeme Institute

1. La méthode Praxeme distingue les domaines d'objets et les domaines fonctionnels. Ceux-ci composent l'aspect pragmatique ou organisationnel de l'entreprise (l'activité métier) ; ceux-là l'aspect sémantique (les objets métier). Cette distinction a des conséquences gigantesques en termes d'architecture logique et modifie totalement la physionomie des systèmes informatiques pour mieux les aligner sur le métier et pour plus d'agilité.

Table des matières

Préface	III
Avant-propos	XV
Introduction	1
Chapitre 1 – La notion de référentiel	3
1.1 Origines et usages	4
1.1.1 Référentiel : ce à quoi l'on se réfère	4
1.1.2 Usages courants	4
1.1.3 Un précurseur des référentiels – Paul Otlet	6
1.1.4 Référentiels et systèmes d'information	8
1.2 Qu'est-ce qu'un Référentiel de Système d'information ?	12
1.2.1 Contexte	12
1.2.2 Définition	13
1.2.3 Propriétés caractéristiques d'un référentiel	14
1.3 Le référentiel, une base vivante au cœur du S.I.	19
1.3.1 Logique d'ensemble	20
1.3.2 Le paradoxe stabilité - évolutivité	21
1.3.3 Coûts induits par la « vie » du référentiel	22
1.3.4 Référentiel et photographie du référentiel	23
1.3.5 Contre-exemples	24

Pour conclure	25
Chapitre 2 – Répertoires et nomenclatures	27
Introduction	27
2.1 La notion d'objet métier	27
2.1.1 <i>Des données informatives aux objets métier</i>	28
2.1.2 <i>Définition possible</i>	28
2.1.3 <i>Objet métier référentiel et objet métier opérationnel</i>	29
2.1.4 <i>Cycle de vie</i>	29
2.2 Les répertoires, ensembles d'objets métier	30
2.2.1 <i>Définition</i>	30
2.2.2 <i>Champ couvert</i>	30
2.2.3 <i>Contenu</i>	31
2.2.4 <i>Répertoire et système d'information</i>	32
2.3 Les nomenclatures, ensembles de catégories	33
2.3.1 <i>Définition</i>	33
2.3.2 <i>Contenu</i>	34
2.3.3 <i>Sémantique</i>	35
2.3.4 <i>Nomenclature et système d'information</i>	36
2.4 Les situations mixtes	37
2.4.1 <i>Une question de point de vue</i>	37
2.4.2 <i>Répertoires articulés</i>	38
2.4.3 <i>Nomenclatures enrichies</i>	39
2.5 Le cas des référentiels externes	40
2.5.1 <i>Répertoires externes</i>	40
2.5.2 <i>Nomenclatures externes</i>	41
Pour conclure	43
Chapitre 3 – L'identification au cœur des référentiels	44
Introduction	44
3.1 Identifier un objet métier	45
3.1.1 <i>Les données individuelles</i>	45
3.1.2 <i>La notion de référentiel-socle</i>	49
3.1.3 <i>Les fonctions de base</i>	50

3.2	Identifier une catégorie, i.e. classer un objet métier	57
3.2.1	Classer plutôt qu'identifier	58
3.2.2	Situation de départ : l'objet métier à classer	62
3.2.3	Classer : quand, qui, quoi, comment	67
3.2.4	Modalités d'automatisation et qualité	71
	Pour conclure	74
	Chapitre 4 – Les enjeux des données de référence	76
	Introduction	76
4.1	Le contexte informationnel de l'entreprise	77
4.1.1	Le pouvoir de l'information	77
4.1.2	L'objet métier	78
4.2	Les enjeux opérationnels	81
4.2.1	Classer le monde	81
4.2.2	Servir les processus métier	83
4.2.3	Faire collaborer les métiers	89
4.2.4	Porter la politique d'entreprise	90
4.2.5	Organiser le reporting	93
4.2.6	Les enjeux spécifiques des nomenclatures	95
4.2.7	Synthèse	97
4.3	Les freins à la mise en place de systèmes référentiels	98
4.4	Les enjeux S.I.	101
4.4.1	Les systèmes référentiels et l'architecture du S.I.	101
4.4.2	Les systèmes référentiels et les projets	102
	Pour conclure	104
	Chapitre 5 – Situations à éviter : études de cas	105
	Introduction	105
5.1	La fausse simplicité des référentiels, structure et organisation	106
5.1.1	Une unité de sens difficile à capter	106
5.1.2	Un périmètre souvent ambigu	107
5.1.3	Un « maintien à jour » qui relève de la haute voltige	107
5.1.4	Un point d'entrée vers d'autres référentiels	108
5.1.5	Un référentiel sans propriétaire	108

5.2	Un référentiel négligé mais qui porte des enjeux essentiels	108
5.2.1	<i>L'enjeu par rapport au pilotage de l'entreprise</i>	110
5.2.2	<i>L'enjeu vis-à-vis de l'exécution des processus de l'entreprise</i>	110
5.2.3	<i>Au final un enjeu clé pour l'agilité de l'entreprise</i>	111
5.3	La double peine des référentiels structure et organisation.....	111
5.3.1	<i>Point 1 : un système, trois référentiels</i>	112
5.3.2	<i>Point 2 : un dispositif d'alimentation compliqué</i>	112
5.3.3	<i>Point 3 : des copies sans asservissements</i>	115
5.3.4	<i>Point 4 : des consommations par transitivités</i>	115
5.3.5	<i>Point 5 : l'utilisation couplée de deux représentations</i>	116
5.3.6	<i>Point 6 : une mise à jour par duplication systématique</i>	116
5.3.7	<i>Point 7 : l'accès direct aux tables internes du référentiel</i>	116
5.3.8	<i>Point 8 : la duplication vers un autre S.I.</i>	116
5.4	Processus de l'entreprise et référentiels : dysfonctionnements.....	117
5.4.1	<i>Opérer le processus : une exécution mise à mal</i>	117
5.4.2	<i>Production du processus : des résultats non partagés voire en défaut</i>	119
5.4.3	<i>Piloter à partir des résultats des processus : des tableaux de bord erronés</i>	120
5.4.4	<i>Optimiser le déroulement du processus : des indicateurs en défaut</i>	121
	Pour conclure.....	122
	Chapitre 6 – De la qualité des référentiels	123
	Introduction.....	123
6.1	La notion de qualité	124
6.1.1	<i>Définitions</i>	124
6.1.2	<i>La démarche qualité</i>	124
6.1.3	<i>Évolutions récentes</i>	125
6.2	La qualité des données.....	126
6.2.1	<i>Les données : un produit particulier</i>	127
6.2.2	<i>Critères de qualité des données</i>	128
6.2.3	<i>Les limites de la notion de justesse</i>	130
6.2.4	<i>Vérification de données et fiabilité</i>	132
6.2.5	<i>Démarches classiques de maîtrise de la qualité des données</i>	135
6.3	Qualité d'un référentiel.....	137
6.3.1	<i>Qualité d'un répertoire</i>	138

6.3.2	<i>Les paradoxes de l'exactitude référentielle</i>	140
6.3.3	<i>Qualité d'une nomenclature</i>	141
6.4	Maîtriser la qualité d'un référentiel	143
6.4.1	<i>Le niveau préventif</i>	143
6.4.2	<i>La maîtrise des flux</i>	146
6.4.3	<i>Le niveau curatif : l'analyse du stock</i>	152
	Pour conclure	154
	Chapitre 7 – Architecture d'un système référentiel	155
	Introduction	155
7.1	L'architecture métier d'un référentiel	156
7.1.1	<i>Les objets d'un système référentiel</i>	157
7.1.2	<i>Les utilisateurs du référentiel</i>	159
7.1.3	<i>Les processus et activités métier</i>	159
7.1.4	<i>Les exigences non fonctionnelles</i>	163
7.1.5	<i>La qualification des contenus</i>	163
7.2	L'architecture fonctionnelle	166
7.2.1	<i>Services de gestion du méta-modèle</i>	168
7.2.2	<i>Services d'administration et de pilotage des contenus</i>	169
7.2.3	<i>Services de collecte et d'enrichissement</i>	170
7.2.4	<i>Services internes de gestion et de support à la qualité</i>	171
7.2.5	<i>Services offerts aux consommateurs</i>	174
7.2.6	<i>Services attachés à la sécurité</i>	176
7.2.7	<i>Services communs d'échange</i>	178
7.2.8	<i>Services spécifiques</i>	180
7.2.9	<i>Vision d'ensemble de l'architecture fonctionnelle</i>	181
7.3	L'implémentation des services fonctionnels	182
7.3.1	<i>Les scénarios de déclinaison de l'architecture fonctionnelle</i>	182
7.3.2	<i>Les familles d'architecture applicative</i>	182
7.3.3	<i>Les systèmes référentiels basés sur des structures d'accueil</i>	188
7.4	Les référentiels dans l'architecture du S.I.	190
7.4.1	<i>La zone des référentiels et l'architecture fonctionnelle du S.I.</i>	192
7.4.2	<i>L'intégration dans le S.I.</i>	193
7.4.3	<i>Le rôle de la couche d'échange</i>	199

7.4.4	Référentiels et normes d'échanges	201
7.4.5	Les référentiels dans la cartographie du S.I.	203
	Pour conclure	207
Chapitre 8 – Construire un référentiel : le projet		208
	Introduction	208
8.1	Qu'est-ce qu'un projet référentiel ?	208
8.1.1	Sa définition et son origine	209
8.1.2	Sa raison d'être	210
8.1.3	Projet référentiel et schéma directeur	211
8.1.4	Le contenu d'un projet référentiel	213
8.2	L'organisation d'un projet référentiel	216
8.2.1	L'organisation classique d'un projet	216
8.2.2	Les acteurs d'un projet référentiel	218
8.3	Le déroulement d'un projet référentiel	222
8.3.1	La phase de cadrage	222
8.3.2	La réalisation : ce qu'il faut faire	229
8.4	Combien coûte un système référentiel	241
8.4.1	Les trois coûts d'un système référentiel	241
8.4.2	Les coûts de l'absence de système référentiel	243
8.4.3	Le piège de la complexité	244
	Pour conclure	246
Chapitre 9 – Opérer un référentiel		247
	Introduction	247
9.1	Que signifie opérer un référentiel ?	248
9.2	Opérer un référentiel vu du D.S.I.	250
9.2.1	Le contexte : la montée en puissance du « run »	250
9.2.2	Assurer le service	252
9.2.3	Gérer la relation avec les clients utilisateurs du référentiel	253
9.2.4	Piloter le niveau de service	257
9.3	Opérer le référentiel vu du métier	258
9.3.1	L'administration des données de référence	258

9.3.2	<i>La gouvernance des données de référence</i>	259
9.3.3	<i>Niveaux de maturité en gouvernance des données</i>	262
9.3.4	<i>Le pilotage des référentiels</i>	263
9.3.5	<i>Les trois niveaux d'intervention pour les métiers</i>	270
	Pour conclure	271
	Chapitre 10 – Et demain ?	272
	Introduction	272
10.1	Les référentiels : phares dans un océan de données ?	273
10.1.1	<i>DaaS : le business des données de référence</i>	274
10.1.2	<i>Big Data : la maîtrise de l'infini des informations</i>	276
10.1.3	<i>Open Data : l'interopérabilité pour les données publiques</i>	280
10.1.4	<i>Les chaînes de valeur de l'exploitation des données</i>	280
10.1.5	<i>Référentiels puissance Web !</i>	281
10.2	Un référentiel universel des personnes	284
10.2.1	<i>L'étendue de nos ombres numériques</i>	284
10.2.2	<i>Une gestion défaillante</i>	286
10.2.3	<i>Des solutions bien réelles</i>	287
10.2.4	<i>Les enjeux liés aux données personnelles</i>	290
10.2.5	<i>Le pire et le meilleur</i>	292
	Épilogue	295
	Bibliographie	297
	Index	299

Avant-propos

Notre ambition est d'écrire un ouvrage de référence... sur les référentiels, ou données de référence au sein des entreprises. En effet, il nous semble que le sujet, bien que régulièrement abordé dans des publications¹, bien que jugé essentiel par nombre d'auteurs, n'est abordé à chaque fois que de façon incidente ou imprécise, et toujours parcellaire. Chacun voit bien ce qu'est, dans l'esprit, un référentiel d'entreprise, mais s'il s'agit de le caractériser pratiquement dans toutes ses dimensions, cela ne va pas de soi.

Formaliser et approfondir la notion de référentiel d'entreprise présente en outre un intérêt supplémentaire : cela permet d'apporter un repère « dur » dans le monde parfois « mou » et abstrait de l'urbanisation des systèmes d'information.

À qui s'adresse ce livre ?

Cet ouvrage s'adresse à tous ceux qui souhaitent comprendre comment s'articulent et s'organisent les données de référence au sein du système d'information : informaticiens (chefs de projet, architectes, urbanistes, analystes fonctionnels, maîtrises d'ouvrages) mais aussi aux non techniciens de l'informatique qui veulent mieux voir comment et en quoi les données informatiques contribuent à la performance de l'entreprise, à sa différenciation concurrentielle et à son pilotage.

Comment est organisé l'ouvrage ?

L'ouvrage est articulé en quatre temps (cf. figure 1).

Le premier temps pose les fondamentaux, avec les chapitres 1 à 4, qui explorent les définitions, les principaux contextes d'usage et les enjeux des données de référence.

- Le chapitre 1 permet de définir les différents concepts qui seront utilisés par la suite, afin de caler la perspective. Et principalement le concept de référentiel.

1. Les ouvrages anglo-saxons parlent de « Master Data », ou « Données Maîtres », terme que l'on retrouve dans l'acronyme MDM, ou Master Data management, qui qualifie les outils et les systèmes informatiques spécialisés dans la gestion des données de référence.

- Le chapitre 2 explore les répertoires et les nomenclatures, les deux catégories de référentiels d'entreprise que nous trouvons au sein des systèmes d'information. Il s'agit d'explicitier leur structure, leur contenu, en s'appuyant sur la notion d'objet métier et en mettant en évidence les situations intermédiaires.
- Le chapitre 3 complète notre compréhension avec la notion d'identification, au cœur des référentiels. On verra que dans le cas des répertoires, cela requiert des algorithmes spécifiques et une attention particulière aux identifiants ; alors que les nomenclatures renvoient à la notion de classement et soulèvent des questions de sémantiques.
- Le chapitre 4 nous donne l'occasion de nous interroger sur les enjeux adressés par les référentiels. Car il faut bien réaliser que les référentiels ne constituent pas un but en soi, mais qu'ils contribuent à servir une stratégie. Ainsi dans certains cas les référentiels se révèlent cruciaux, et à ce titre ils doivent faire l'objet d'une attention particulière.

Le deuxième temps, charnière, traite du concept de qualité pour les référentiels, et des dysfonctionnements qui peuvent découler d'une médiocre qualité. Nous proposons tout d'abord une illustration par l'exemple, avec le chapitre 5 qui décrit les causes et conséquences d'un système référentiel mal organisé, situation qu'il convient d'éviter. Le chapitre 6 nous présente une vision globale de ce qu'est la qualité pour un référentiel. On distinguera notamment la notion de qualité en général, puis la qualité de données, puis ce qu'un référentiel présente de spécifique.

Le troisième temps aborde l'ingénierie des systèmes référentiels, traitée de A à Z au travers des chapitres 7, 8 et 9.

- Le chapitre 7 est consacré à l'architecture d'un système référentiel. Il nous explique comment les référentiels s'inscrivent dans le système d'information et interagissent avec lui, comment ils s'y intègrent et selon quelles architectures. C'est, finalement, cet objet en mouvement, dans son milieu naturel, qu'il s'agira de décrire. Car un référentiel n'a pas de sens pris isolément : il ne prend sa signification que vis-à-vis du système qu'il irrigue. Le chapitre se présente, à ce titre, comme un guide de conception.
- Au chapitre 8 nous découvrirons ce que peut être un projet de mise en place d'un système référentiel : comment et à quelle occasion nous pouvons et nous devons le construire, ce que nous devons faire pour le construire, combien cela coûte de le faire... ou de ne pas le faire...
- Enfin, au chapitre 9, nous verrons comment faire fonctionner un système référentiel, et qui doit le faire fonctionner. Nous précisons ainsi le rôle de l'informatique mais aussi celui, crucial, des métiers.

Le quatrième et dernier temps clôture cet ouvrage, au chapitre 10, par un regard prospectif. Ce regard nous proposera deux voyages vers le futur : le premier voyage sera une plongée dans un océan de données, le second nous emmènera à la découverte d'un hypothétique référentiel universel des personnes.

Les référentiels du système d'information

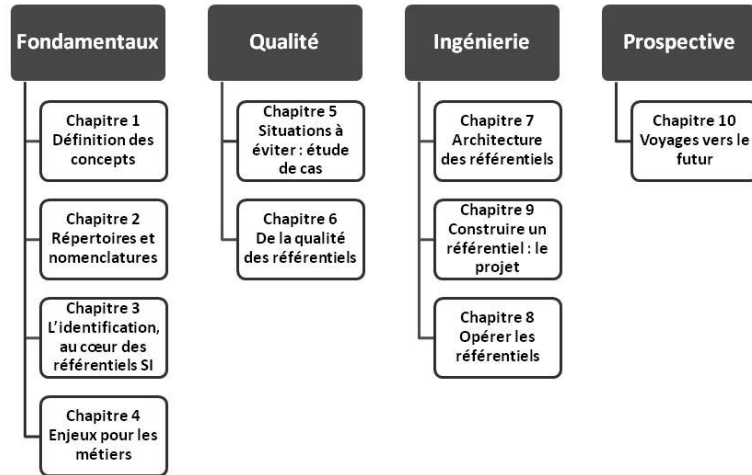


Figure 1 – Articulation de l'ouvrage

Introduction

Les référentiels d'entreprise dans les systèmes d'information

C'est désormais un lieu commun : les systèmes d'information, tout en occupant une place de plus en plus stratégique dans l'entreprise, se complexifient. Cela devient un problème majeur pour les entreprises, sachant que les coûts du S.I. sont croissants pour cette même raison. Nombre d'articles, de méthodes, d'ouvrages et bien sûr d'outils ont été produits pour tenter d'y apporter, fût-ce partiellement, une réponse.

On sait par exemple que pour y remédier, il est important de structurer au mieux le système d'information, pour en avoir la maîtrise. On cherche donc notamment à le subdiviser en « blocs » (fonctionnels, applicatifs) homogènes et suffisamment indépendants les uns des autres, puis à assurer la cohérence de cette vision macroscopique avec les processus métier d'une part, l'architecture applicative d'autre part. C'est la démarche dite « d'urbanisation » du système d'information : si elle est intellectuellement séduisante, elle présente le risque de demeurer parfois abstraite.

Le présent ouvrage s'attaque à un sujet essentiel pour une bonne maîtrise du système d'information : les *référentiels d'entreprise*, qui jouent un rôle pivot dans le S.I. Ce sujet présente la particularité d'être central sur le plan conceptuel tout en étant très opérationnel. Peu à peu, on en réalise l'importance, et les entreprises le prennent de plus en plus en considération. Ainsi, en 2011, on peut citer au moins trois grands groupes français qui ont créé une direction dédiée aux référentiels : on entend par là non pas une cellule d'urbanistes au sein d'une direction des systèmes d'information, mais une véritable direction, rattachée à la direction générale au même titre qu'une direction logistique, par exemple.

Un récent tour d'horizon dans des secteurs d'activité très distincts (banque, assurance, administration, grande distribution, énergie, sécurité sociale, industrie, télécommunications...) a confirmé le pressentiment selon lequel la problématique référentielle prenait de l'ampleur depuis un ou deux ans, de façon tout à fait significative.

Mais utiliser le mot « référentiel » ne suffit pas, et ce d'autant moins qu'il est utilisé de façon massive, avec des significations différentes. On va donc s'attacher à le définir. Ce qui signifie, au passage, que l'on va en restreindre le champ, en ne prenant en compte qu'une partie de ses sémantiques possibles.

En effet, on ne considérera ici que les référentiels gérés dans le système d'information, c'est-à-dire capables d'irriguer ce dernier en *données de référence* (« *Master Data* » pour les Anglo-Saxons) et d'influer, à travers lui, sur les processus métier de l'entreprise.

Ces données qui dirigent nos entreprises

Les données qui irriguent un système d'information n'ont pas toutes le même rôle, et, les architectes S.I. le savent bien, l'urbanisme du système d'information passe par un urbanisme des données. Dans ce contexte les *données de référence* ont une place à part, par le rôle qu'elles jouent dans le S.I. et par les enjeux qu'elles prennent en compte au sein de l'entreprise. Elles constituent l'ossature informationnelle sur laquelle les applications informatiques, de plus en plus complexes, étendues et multiples, pourront se construire, demain, avec le maximum d'agilité.

À condition que ces *données de référence* soient correctement qualifiées, isolées et gérées, au sein de systèmes dédiés.

C'est aujourd'hui loin d'être le cas : les données de référence sont le plus souvent *disséminées dans de multiples bases de données*, au milieu d'autres informations. Elles sont aussi fréquemment *dupliquées à de multiples endroits du S.I.* et leur qualité est loin d'être sous contrôle.

Et pourtant, comme nous allons le voir, pour une large part, *elles dirigent nos entreprises !*

1

La notion de référentiel

DE QUOI PARLE-T-ON ?

« Référentiel de bonnes pratiques », « référentiel de métiers », « référentiel CMMI »... tout est référentiel. Le terme est galvaudé. Et si l'on approfondit chacun de ces exemples, on peut s'apercevoir qu'ils recouvrent des objets radicalement différents, ayant peu en commun.

La question est donc de savoir ce que l'on peut entendre par ce terme : d'abord dans une acception très générale, en partant des origines pour aboutir à la vaste palette d'usages actuels de ce mot, et en esquissant une typologie. Puis en se centrant sur la notion de *référentiel du système d'information*, et plus précisément de système porteur de données de référence d'une entreprise. On définira alors ce concept par un ensemble de propriétés caractéristiques, et l'on présentera les deux grands types de référentiels qui structurent le système d'information, les répertoires et les nomenclatures. Par ailleurs, et ce sera là un leitmotiv de l'ouvrage, on verra qu'un référentiel est fondamentalement un objet vivant : stable, certes, mais certainement pas figé.

Ainsi, l'objectif de ce chapitre est, avant tout, de chercher à mieux cerner l'objet *référentiel*, à en formaliser la définition, et à analyser tout ce qui découle de cette définition. Cela est nécessaire, compte tenu du flou artistique qui entoure généralement cette notion. Mais l'on se restreindra délibérément à une vision générale : non seulement on n'entrera pas dans des questions d'architecture technique ou d'organisation, mais on ne précisera pas non plus le contenu, la structure du référentiel, toutes choses qui seront étudiées dans les chapitres suivants.

1.1 ORIGINES ET USAGES

1.1.1 Référentiel : ce à quoi l'on se réfère

Si l'on consulte les dictionnaires d'étymologie, le terme lui-même est assez récent, et date du XIX^e siècle. Il provient du latin *referre*, littéralement rapporter, voire attribuer¹. Ce terme appartient à une famille de mots émanant autant du langage administratif que de l'action : référé, référence, référencer, se référer, etc., « dont le point commun est de rapporter une chose à autre chose, avec une certaine force conférée à l'acte ».

En physique, un « référentiel galiléen » désigne un repère spatial servant de référence pour décrire les coordonnées de n'importe quel point. Il définit bien *ce par rapport à quoi* on se positionne : sans référentiel, pas de coordonnées. Les référentiels galiléens jouent ainsi un rôle central dans la mécanique newtonienne. Un référentiel temporel est également toujours nécessaire.

Lors de la révolution industrielle, les référentiels répondaient au besoin d'uniformisation croissant de l'économie, et relevaient d'un usage plus concret, opérationnel, généralisé. À nouveau, il s'agissait en premier lieu de temps et d'espace. La mise en place d'un référentiel temporel unique ne fut pas une mince affaire et souleva des défis nouveaux². Le besoin de référentiels spatiaux communs était un besoin de cartographies, en vue de faciliter les transports, notamment. À la même période, on pourrait citer aussi les référentiels monétaires, indispensables pour fluidifier l'échange de marchandises.

1.1.2 Usages courants

Le spectre d'utilisation du mot « Référentiel » s'est peu à peu étendu, s'appliquant progressivement à tous les domaines d'activité : médecine, industrie, physique, linguistique, psychologie, formation...

Cette généralisation de l'usage du terme a mis en évidence un besoin de standardisation, une inflation normative : d'un côté, toute la complexité et la diversité du monde, son irréductibilité à une quelconque normalisation, et, de l'autre, le besoin de repères bien définis.

On peut donner de multiples exemples d'utilisation, qui en illustrent différents aspects, en commençant par le domaine des sciences sociales.

En sciences du langage, le référentiel est « ce qui est censé se rapporter au monde réel (objectif) par opposition aux sentiments (subjectifs) des locuteurs. Le référentiel

1. CROS F., RAISKY C. (2010), *Autour des mots de la formation « Référentiel »*, Recherche et Formation, n°64-2010, pp. 105-116. Cet article porte sur les référentiels dans le domaine de la formation mais il traite aussi des référentiels en général. Certains éléments sont repris dans cette section et au début de la section 1.1.2..

2. ROSA H. (2010), *Accélération*, La Découverte, p. 127.

est relatif à la référence, c'est-à-dire à la fonction d'un signe linguistique en tant qu'il désigne une chose. »

En psychologie du développement, le référentiel spatial, par exemple, est « l'ensemble des activités formant système par lesquelles l'enfant localise les objets les uns par rapport aux autres, ainsi que par rapport à lui-même ».

Dans le champ des sciences sociales, le référentiel est un construit social qui clarifie les normes d'une activité ou d'un sens attribué à des systèmes sociaux. Il est *ce par rapport à quoi* un jugement ou un sens est donné à un objet ou une action. Autrement dit, c'est un outil de médiation à caractère normatif permettant aux activités humaines de s'y référer (de s'y rapporter) pour étudier un écart ou des différences. Mais il ne prétend pas particulièrement correspondre à une réalité. C'est à travers les comparaisons que l'on peut faire qu'il produit des savoirs. « Il permet de repérer la position d'un objet par rapport à un ensemble d'autres objets qui jouent le rôle de système de référence ». Cette idée de point de repère, base de calcul d'écart, dépasse largement le champ des sciences sociales.

Mais un référentiel est aussi une *source d'information officielle* dans laquelle on vient puiser. C'est le cas, par exemple, en gestion des ressources humaines, avec les référentiels de compétences, de fonctions, de métiers, de postes. L'optique n'est alors pas nécessairement d'analyser des différences, mais, plus prosaïquement, de classer, de catégoriser des notions complexes, en offrant un langage commun et reconnu, ce qui peut faciliter des processus opérationnels, ici le processus de recrutement.

On retrouve l'importance du classement des objets avec l'exemple des référentiels comptables, qui permettent notamment de classer tel ou tel enregistrement comptable dans un *poste* de la comptabilité. L'intérêt est ensuite de faciliter l'agrégation de ces données par poste, et donc de rendre possible l'élaboration d'états synthétiques et normalisés comme le compte de résultat ou le bilan.

Le référentiel des *produits* d'une entreprise offre un langage commun, un cadre pour le vendeur. Il permet d'assurer la médiation entre l'activité de vente, l'activité de production et les activités support. D'autre part, ce référentiel embarque également avec lui des informations indispensables comme les caractéristiques du produit, son prix, les conditions de réduction pendant une période donnée, etc.

Dès lors qu'ils s'inscrivent dans un système d'information, les référentiels présentent des caractéristiques particulières : ils ne constituent pas seulement un point de référence pour calculer un écart, ou une documentation officielle dans laquelle puiser, mais une source d'information capable d'alimenter *automatiquement* le S.I., d'y jouer un rôle actif. Cette dimension d'automatisation n'est pas sans impact sur la structure même de l'objet.

Aujourd'hui, le terme « référentiel » s'inscrit dans le développement massif des échanges et de la communication induit par la société de l'information, et la nécessité d'un langage commun pour ces échanges. Ainsi l'informatisation de la société en général génère-t-elle de nouveaux besoins, susceptibles d'alimenter la demande de nouveaux référentiels.

Par ailleurs, on constate qu'il est difficile de dissocier la notion de référentiel des usages qui en sont faits : il ne faut pas l'assimiler à un vieux dictionnaire, au fond d'une bibliothèque, qu'on irait simplement consulter. Le référentiel, on l'a vu, déclenche une action (comparer, classer, mettre à jour...) ; il s'inscrit ainsi dans une perspective dynamique, dans un entrelacs de processus plus ou moins automatisés du système d'information.

1.1.3 Un précurseur des référentiels – Paul Otlet

En 1894, bien avant l'ère du numérique, deux Belges, Paul Otlet et Henri Lafontaine, ont eu l'idée de créer l'un des référentiels les plus ambitieux de tous les temps. De cette réflexion va naître l'épopée du plus grand répertoire papier jamais créé par l'homme : le *Mundaneum*. Quarante ans après, et avant que le désintérêt des politiques et la seconde guerre mondiale ne mettent indirectement un terme à cette aventure, le répertoire comportera plus de 16 millions de fiches restées célèbres par leur format (7,5 x 12,5 cm).

L'objet de ce répertoire papier est de référencer exhaustivement tous les documents, articles et ouvrages publiés dans le monde, quelle que soit l'époque ou la langue employée, autrement dit de réaliser un *répertoire bibliographique universel*.

L'idée n'est pas neuve pour l'époque. Au début du XIX^e siècle, d'autres acteurs comme le français Gabriel Peignot ont eu la même ambition. Mais, avec Paul Otlet, c'est la première fois que la problématique de mise en œuvre d'un tel monument est formalisée dans toutes ses dimensions : système d'indexation, standardisation du modèle de fiches de référence, système de production et évaluation de la productivité de composition des fiches, contrats de sous-traitance de classements (30 centimes la fiche), moyens matériels et système de stockage (armoires, fiches), systèmes de consultation (quels ouvrages publiés sur telles matières, par tels auteurs, à telle époque...) et de distribution (copie et publication des fiches).

En bon concepteur du système, Paul Otlet en fixe l'ambition et les qualités visées : « *Les buts de la documentation consistent à pouvoir offrir sur tout ordre de fait et de connaissance, des informations documentées 1° universelles quant à leur objet, 2° sûres et vraies, 3° complètes, 4° rapides, 5° à jour, 6° faciles à obtenir, 7° réunies d'avance et prêtes à être communiquées, 8° mises à la disposition du plus grand nombre* »¹. Autant de qualités que tout système référentiel actuel devrait reprendre à son compte !

Dès le départ, Paul Otlet et Henri Lafontaine portent leurs efforts sur le mode d'indexation du contenu du répertoire. Ils reprennent à leur compte et font évoluer le système de classification décimale de l'américain Melvin Dewey. Ce système propose un arbre de classification et de subdivision décimal des connaissances humaines (0 : Ouvrages généraux, 2 : Religion, 3 : Sociologie, 4 : Philologie, 5 : Sciences, 6 : Sciences appliquées, 7 : Beaux-arts, 8 : Littérature, 9 : Histoire. Puis des subdivisions par classe sont définies. Ainsi, par exemple, la comptabilité y est classée sous le code

1. OTLET P. *Traité de documentation : le livre sur le livre : Théorie et pratique*. Brussels, éditions Mundaneum, 1934. Accessible également en ligne.

657, sous division des techniques de commerces 65, elle-même sous-division des sciences appliquées 6).

Les évolutions qu'ils y apportent permettent d'indiquer qu'une même œuvre peut appartenir à plusieurs catégories différentes, qu'une catégorie est en relation avec une autre catégorie, qu'une œuvre est issue d'un lieu donné...

L'idée d'internationalisation du système apparaît rapidement. En 1910, Paul Otlet et Henri Lafontaine s'intéressent aux initiatives prises dans les autres pays. Des standards émergent dans les pays de langue anglaise et allemande, mais les deux acteurs visent une uniformisation mondiale dans laquelle ils vont s'investir au-delà de la simple problématique du répertoire bibliographique.

Cette uniformisation, en particulier autour des évolutions sur le système de classification décimale, ne se réalisera pas, mais la structure du répertoire démontre son interopérabilité. 600 000 fiches de la prestigieuse bibliothèque du congrès américain sont échangées avec le répertoire universel et intercalées directement dans le système sans transformation.

Dans son *Traité de la documentation*, 40 ans après le début du projet, Paul Otlet rédige un ouvrage monumental, à la fois dans sa couverture technique (au passage il étend les fonctions du répertoire avec l'idée d'écriture secrète – cryptographie, ou encore d'accessibilité aux aveugles via le braille), et également dans sa dimension visionnaire.

En effet, il s'intéresse aux nouvelles technologies de communication de l'époque avec ce qu'il appelle les « substituts du livre » autour du cinéma, de l'image, de la radio, de la télévision qui vient de naître. Il va en dégager une vision qui préfigure Internet et les usages que l'on peut en faire aujourd'hui.

Parmi les avancées qu'il imagine, on peut citer : l'acquisition photographique d'ouvrages (filmer les pages d'un livre pour le reproduire), mais aussi la machine à lire qui restituerait les ouvrages à partir des copies du film, la microphotothèque (microfilms), les centrales d'édition de livres épuisés, l'association des autres sens aux documents (l'odorat, le goût, le toucher). La plus audacieuse de ces avancées est la consultation d'ouvrage à distance où le serveur est une salle spéciale de bibliothèques dans lesquelles les livres sont exposés à la demande et consultés à distance grâce à un dispositif qui rappelle l'Internet et nos tablettes numériques !

« Ici, la Table de Travail n'est plus chargée d'aucun livre. À leur place se dresse un écran et à portée un téléphone. Et là-bas, au loin, dans un édifice immense, sont tous les livres et tous les renseignements. [...] De là, on fait apparaître sur l'écran la page à lire pour connaître la question posée par téléphone, avec ou sans fil. Un écran serait double, quadruple ou décuple s'il s'agissait de multiplier les textes et les documents à confronter simultanément ; il y aurait un haut-parleur si la vue devait être aidée par une donnée ouïe [...] Et ce perfectionnement pourrait aller peut-être jusqu'à rendre automatique l'appel des données sur l'écran ». Extrait « *Traité de documentation* » Paul Otlet – p. 275 « L'homme qui voulait classer le monde » - F. Levie.