



HAL
open science

Wiki sémantique en entreprise : un retour d'expérience

Thierry Bénard

► **To cite this version:**

Thierry Bénard. Wiki sémantique en entreprise : un retour d'expérience. [Rapport Technique]
Hutchinson. 2014. hal-01185539

HAL Id: hal-01185539

<https://hal.science/hal-01185539>

Submitted on 21 Aug 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Wiki sémantique en entreprise : un retour d'expérience

Thierry Bénard

24 octobre 2014

thierry-b.benard@hutchinson.fr
Hutchinson
Châlette-sur-Loing, France

Résumé

La donnée technique est l'un des actifs de l'entreprise. Dans un contexte où de plus en plus d'organisations sont mondialisées, il devient urgent d'équiper les entreprises de systèmes informatiques permettant de gérer et de diffuser les données aux équipes. Il a été choisi dans cette entreprise d'utiliser un wiki et d'y associer les principes du web sémantique. Au bout de 18 mois d'utilisation et de 1400 pages créées, l'intérêt d'un tel système ne fait plus aucun doute. Les utilisateurs apprécient la technologie web pour consulter l'information et les liens hypertextes permettent de naviguer et de retrouver une information facilement. La structuration de l'information a aussi permis de faire de l'analyse des données et de créer des tableaux de bord automatisés. Il faut maintenant que les wiki intègrent mieux les éditeurs wysiwyg (*What you see is what you get*) pour déployer plus largement l'écriture des pages dans l'entreprise.

Mots clés : gestion des données, web sémantique, wiki, gestion de la connaissance

INTRODUCTION

L'objectif de cette contribution est de présenter un retour d'expérience sur l'utilisation d'un wiki sémantique pour organiser l'information technique au sein d'une entreprise industrielle. Il s'agit d'un cas concret et réel mis en place au mois de février 2013. Nous détaillerons au long des pages de cet article la solution retenue pour gérer l'information technique de cette entreprise et nous analyserons les succès et échecs de cette mise en place au bout de 18 mois d'utilisation.

Le cadre général de cet article repose sur le constat suivant : de plus en plus, pour développer un produit, les entreprises ont besoin de gérer un nombre croissant d'informations, et celles-ci sont de plus en plus diverses. C'est la conséquence directe de l'augmentation de technicité des objets qui nous entourent et la voiture en est un bel exemple (Tarondeau 1994).

Prenons l'exemple de la protection contre le vol des voitures. Auparavant, pour sécuriser son véhicule, il suffisait d'une clé et d'un verrou. Les entreprises, qui concevaient et fabriquaient ces éléments, manipulaient des données uniquement mécaniques c'est-à-dire des dimensions et les tolérances de fabrication associées. Un plan avec toutes les dimensions et les tolérances suffisait pour décrire complètement le produit. Avec l'arrivée des télécommandes à infra rouge ou à radio fréquence, ces entreprises ont appris à manipuler des données du domaine de l'électronique (tension électrique d'alimentation, consommation électrique des équipements) et de traitement du signal (cryptographie pour sécuriser la communication entre la clé électronique et le véhicule). Un plan de la pièce ne suffit plus pour décrire complètement le produit, il faut aussi le schéma électrique de la clé, le logiciel et l'étude de cryptographie pour décrire techniquement cet objet. Cet exemple montre clairement l'augmentation de la quantité et l'accroissement de diversité des informations techniques.

Les entreprises ont aussi à faire face à un nouveau challenge : la mondialisation de l'économie. Les entreprises doivent réagir en déployant des organisations mondialisées (Montel-Dumont 2011). Les équipes sont dispersées sur tous les continents, proches des clients (les constructeurs de

voitures) et des usines le plus souvent localisées dans des pays à faible coût de main d'œuvre. Pour être efficace, l'entreprise doit avoir ses informations techniques à jour et disponibles n'importe où dans le monde.

Dans cette entreprise, il a été décidé de résoudre le problème de la gestion de l'information avec le support d'un wiki sémantique. Ce principe n'est pas vraiment nouveau et a déjà été expliqué par (Wagner and Bolloju 2004).

Un wiki sémantique est une technologie web donc accessible n'importe où dans le monde à partir d'un accès sécurisé à internet (Caby-Guillet et al. 2009). Le wiki est un outil de travail collaboratif facilitant la maintenance d'un nombre élevé d'informations. En ajoutant au wiki les principes du web sémantique, il devient alors plus facile de gérer la diversité des informations et celles-ci deviennent lisibles par une machine (Völkel et al. 2006).

En résumé :

- wiki = collaboratif → solution aux données nombreuses
- wiki + sémantique → solution aux données diverses
- wiki = web → solution à l'accessibilité des données

CONTEXTE ET SUPPORT THÉORIQUE

Description du cadre : l'entreprise

L'entreprise est une société française spécialisée dans la conception, la production et la vente d'équipements participant à la construction de véhicules automobiles. Le type précis des équipements n'a pas vraiment d'importance dans cet article mais il sera néanmoins utile de garder à l'esprit que ces équipements sont réalisés à partir de la transformation de composants primaires ou de matières premières.

Comme la plupart des entreprises du secteur automobile, cette entre-

prise travaille en mode projet au travers d'une organisation matricielle (Zannad 2009). Les grandes directions de cette entreprise directement impliquées dans la vie des projets sont : Projet, R&D (Recherche et Développement), Achat, Commerce, Industrialisation et Production dont la Logistique et la Qualité. Il faut ajouter à ces directions opérationnelles sur les projets les directions supports suivantes : la direction générale, la direction des ressources humaines et la direction finance et comptabilité. Les directions "métier" sont alors tenues de mettre au service des projets les ressources nécessaires pour garantir les objectifs QCD (Qualité, Coût et Délai) de chaque projet.

Cette entreprise a en outre une dimension internationale. Les équipes de conception, de vente ou de production ne sont pas forcément regroupées sur un même site. Un équipement peut être conçu en France, l'équipe commerciale peut-être italienne si le constructeur est installé en Italie et la production peut être localisée dans un pays à bas coût de main d'œuvre.

Les données

Afin de bien saisir l'importance du sujet de la gestion des informations techniques, passons un peu de temps à expliciter ce que l'on entend par "informations techniques". Dans une entreprise industrielle dont le rôle est de fabriquer un produit, une liste (non exhaustive) d'informations à structurer pour être mieux maîtrisées pourrait être :

- un plan de définition du produit
- une liste de caractéristiques du produit dont certaines peuvent être critiques
- la nomenclature listant les composants nécessaires à la réalisation du produit
- un synoptique de fabrication décrivant les opérations de fabrication et leur enchaînement
- une liste de composants à acheter entrant dans la fabrication du produit
- le prix des composants à acheter et les fournisseurs retenus pour

leur approvisionnement

Pour s'assurer que le produit proposé réponde bien aux exigences du client, il faut être capable de mesurer ses caractéristiques techniques ; on peut donc ajouter à la liste et pour chaque caractéristique :

- le plan de validation listant tous les essais et toutes les mesures à réaliser pour prouver que le produit satisfait l'ensemble des exigences du client
- la référence à la méthode d'essai décrivant comment mesurer la caractéristique
- le rapport d'essai enregistrant les résultats de la mesure
- la liste des outils et appareils de mesure ainsi que la date de leur dernière calibration permettant de s'assurer que la grandeur mesurée par l'appareil est fiable.

Mais il faut aussi que la durée de conception du produit soit compatible avec le planning de développement du véhicule ; il faut donc ajouter à la liste des données projet :

- la date de début de l'étude
- la date de fin de l'étude du produit
- la date de lancement des outillages qui permettront aux usines de fabriquer le produit
- la date de mise à disposition des premières pièces
- la date de mise en production

C'est une liste, largement non exhaustive, et chacun en la lisant aura sûrement à l'esprit une autre information manquante.

Quel problème pose un nombre important et varié de données ?

Dans une entreprise mondialisée, organisée de façon matricielle et travaillant en mode projet, la gestion des données techniques est un point critique. En effet, pour travailler efficacement, l'information doit être

disponible partout dans le monde et constamment à jour. Dans certaines entreprises, un temps considérable est passé par les équipes pour récupérer l'information et s'assurer que celle-ci est à jour.

Ces informations sont parfois mal partagées, mal connues par l'ensemble des membres de l'équipe et mal maintenues. Elles sont stockées de façon précaire au mieux sur un disque réseau et au pire sur des disques durs en local. La diffusion se fait au travers de courriels et il est alors fréquent de voir la même information stockée sur plusieurs disques durs.

Un autre problème est l'organisation même des données dans la documentation. Si nous prenons, par exemple, une dimension caractéristique d'un des composants du produit. Cette dimension est créée dans un logiciel de CAO 3D (Conception Assistée par Ordinateur en trois dimensions). Elle est disponible dans le plan 2D de la pièce en format PDF ou image TIF. On la retrouvera aussi dans le plan de validation au niveau de l'essai qui vérifiera cette caractéristique et dans le rapport d'essai correspondant. Normalement, ces quatre documents doivent faire référence à la même caractéristique. Si cette caractéristique évolue au cours de l'étude, elle doit évoluer dans les trois premiers documents. Nous voyons bien aussi qu'un document peut contenir une ou plusieurs informations techniques et, à chaque fois, dans des formes (plus exactement un format informatique) différentes. Imaginons maintenant un système qui contiendrait cette information technique dans un endroit unique, et ce serait au document qui en a besoin d'aller chercher cette information à cet unique endroit. La modification de l'information dans cet endroit unique mettrait automatiquement à jour l'ensemble de la documentation qui y fait référence. Si de plus ce système est un système web, nous touchons là du doigt le point clé de notre exposé : une information toujours à jour et disponible. Il ne reste plus qu'à donner un sens à l'information en lui attribuant un type et nous arrivons alors au "web sémantique" et au "web des données".

Nous voyons aussi que toutes ces informations sont reliées entre elles. Cet exemple illustre que l'important pour une entreprise réside plus dans les données (et les liens entre elles) que dans les applications qui les créent. L'entreprise devient alors centrée sur ses données et non plus sur ses applications.

Les données deviennent actuellement de véritables actifs virtuels de l'entreprise (Simottel 2014) et les entreprises centrées sur les données réussissent financièrement mieux que les autres (Giles 2013). De la même manière, une documentation technique peu sûre provoque des pertes financières (Avers et al. 2012).

Web sémantique, web des données

Pour diffuser et partager des informations, la technologie web est une bonne solution (Cutkosky et al. 1996). Tout les ordinateurs possèdent un navigateur et les connexions à internet sont disponibles facilement dans le monde entier grâce au wifi gratuit dans les hôtels, dans les cafés ou dans les aéroports. La structuration des données suscite depuis longtemps l'intérêt du W3C (*World Wide Web Consortium* <http://www.w3.org/>). En effet, le passage du "web des documents" au "web des données" fait l'objet d'un groupe travail à part entière et l'un des premiers standards du "web des données", le RDF (*Ressource Description Framework*), a été diffusé dès 1999.

Nous allons nous intéresser à l'un des standards du web sémantique qui est le plus proche de la solution qui a été retenue par l'entreprise et qui permettra de mieux appréhender l'articulation du web des données.

Le RDFa est une syntaxe qui permet de structurer des données en RDF à l'intérieur d'une page HTML (*Hypertext Markup Language*) (Adida and Hausenblas 2007). Ce principe nous intéresse car il n'est alors pas nécessaire d'avoir un fichier RDF associé au wiki, la structuration des données se faisant directement dans la page. Le RDFa étant un peu compliqué, nous étudierons plutôt les Microdata, le principe est sensiblement identique mais la compréhension et la lecture beaucoup plus simples.

Nous aborderons le langage Microdata au travers d'un exemple simple. Le texte brut, ou pour être plus précis le texte HTML, est de la forme :

```
<div itemscope itemtype="http://data-vocabulary.org/Person">
  Je m'appelle <span itemprop="name">Jacques Dupont</span>.
  J'habite à <span itemprop="address_city">Paris</span>.</div>
```



```
Voici ma page d'accueil : <a href="http://www.mapage.com"
itemprop="url"> www.mapage.com </a>. Je suis <span
itemprop="title"> comptable </span> dans la société <span
itemprop="affiliation"> ABC </span>. </div>
```

Les différentes balises DIV et SPAN du langage HTML incluent des propriétés particulières : `itemscope`, `itemtype`, `itemprop`. La balise qui intègre la propriété `ITEMSCOPE` définit le début et la fin de ce que l'on pourrait appeler un "enregistrement" d'une base de données. La balise `ITEMTYPE` définit le type, ou bien le nom de la table d'une base de données dans laquelle serait stocké l'enregistrement. Les balises qui contiennent la propriété `ITEMPROP` définissent les champs et les valeurs associées à chaque champs dans l'enregistrement concerné.

La visualisation dans un navigateur donne :

```
Je m'appelle Jacques Dupont. J'habite à Paris. Voici ma
page d'accueil : http://www.mapage.com. Je suis comptable
dans la société ABC.
```

Et voilà comment, un moteur de recherche pourrait récupérer l'information s'il est programmé pour comprendre le langage de structuration Microdata.

```
itemscope (type : person)
name : Jacques Dupont
address_city : Paris
url : http://www.mapage.com
title : comptable
affiliation : ABC
```

En notion de base de données, ces données s'exprimeraient telles qu'affichées dans le tableau 1.

Table 1: Visualisation dans une table de base de donnée

name	address_city	url	title	affiliation
...				
Jacques Dupont	Paris	http://www.mapage.com	comptable	ABC
...				

Nous voyons donc très nettement que l'évolution des technologies web permet d'associer directement dans du texte libre des notions de base de données. Plus besoin d'avoir des bases de données d'un côté et des documents textuels de l'autre pour mettre en forme et publier les informations : la technologie permet d'associer les deux fonctions en un seul langage.

Les wikis

Le terme wiki définit un type de site web dynamique dont des fonctions particulières permettent à tous les utilisateurs de créer et de modifier des pages facilement sans être un expert de la programmation en technologie web (Leuf and Cunningham 2001). Le premier wiki a été créé en 1995 par Ward Cunningham. L'un des plus célèbres est l'encyclopédie collaborative Wikipedia (<http://fr.wikipedia.org/>) qui se trouve classée en sixième position des sites les plus visités en France en juillet 2014 (chi 2014). Le wiki n'est qu'un concept de site web. Techniquement il repose sur un serveur web qui contient des logiciels qui intègrent eux-mêmes les principes et fonctions du wiki, on appelle ce serveur : un moteur wiki. Il existe plusieurs moteurs wiki. Par exemple, l'encyclopédie Wikipedia repose sur le moteur wiki Wikimedia. Le wiki a été créé au départ pour permettre à tout le monde de rédiger des pages web avec un formalisme simplifié. Le mot wiki vient du mot hawaïen "wikiwiki" qui signifie "vite".

Les principes de base des wikis sont :

- Le contenu est modifiable par tout le monde
- Le lien entre deux pages se fait simplement par ajout du nom de la page B dans la page A ce qui crée automatiquement un lien hypertexte vers la page B
- La syntaxe de saisie du texte est simple (pas besoin de connaître le code HTML)
- Une fonction permet de lister toutes les pages qui mentionnent la page courante (c'est à dire les rétroliens)

Rapidement d'autres fonctionnalités sont apparues :

- L'historique de toutes les modifications
- La gestion des droits d'utilisation des pages par création de compte utilisateur

Pour mettre en forme le texte d'un wiki, il faut utiliser un langage WikiML (*Wiki Markup Language*) (Buffa et al. 2007). Ce langage permet d'intégrer simplement des commandes de mise en forme directement dans la page d'édition. Par exemple, dans le moteur de wiki TWiki, la mise en gras de caractères utilise les astérisques, la mise en italique les tirets bas. Un exemple illustre la mise en forme par WikiML pour TWiki dans le tableau 2.

Table 2: Exemple de langage WikiML pour TWIKI

Langage WikiML de TWiki	Rendu dans le navigateur
essai de gras	essai de gras
essai d'italique	<i>essai d'italique</i>

Certains moteurs de wiki ont intégré nativement les principes du Web sémantique pour faire des wikis sémantiques.

MATÉRIEL ET MÉTHODE

Avant la mise en place d'un système de gestion de l'information qui suit les principes du web sémantique, il nous a fallu répondre à deux questions :

- Quel outil informatique de technologie web choisir ?
- Quelle méthode de structuration des données choisir ?

A la première question, il a été décidé de répondre par l'installation d'un wiki. Le principe du wiki a été retenu pour les raisons suivantes :

- technologie web : pour l'utilisateur un navigateur internet est suffisant. Il n'y a pas à déployer un logiciel spécifique.
- collaboratif : tous les employés peuvent, et doivent contribuer à

faire vivre les données du wiki en les produisant, les complétant, les corrigeant ou à tout le moins en les lisant et en les utilisant.

- gestion des révisions : toutes les versions sont conservées et automatiquement signées. C'est la garantie contre le vandalisme, conscient ou non (Pfaff and Hasan 2006).

La plupart des wikis possède aussi maintenant des fonctionnalités qui ne font pas partie des principes fondateurs des wikis mais sont indispensables dans la plupart des situations d'usage multi-utilisateurs :

- Gestion fine des droits d'accès en lecture et en écriture sur les pages.

Pour construire ce wiki sémantique, nous avons utilisé le moteur de wiki TWiki développé par Peter Thoeny (DaLio 2005). A l'origine, ce moteur a été choisi uniquement parce que le service informatique l'avait déjà installé sur le serveur web sécurisé de l'entreprise. Nous verrons comment l'usage a confirmé ce choix.

A la deuxième question, concernant la méthode de structuration de l'information, nous avons choisi d'utiliser des techniques spécifiques au moteur de wiki TWiki mais qui pourraient facilement être étendues à d'autres moteurs wiki. Deux méthodes complémentaires ont été utilisées.

1) Le moteur de wiki TWiki intègre nativement des fonctions de structuration de l'information. En effet, TWiki offre la possibilité d'ajouter facilement des champs formatés à une page. Ces champs structurés, associés à une page, permettent d'ajouter des fonctions de base de données au wiki. Le wiki n'est plus seulement un ensemble de pages web reliées entre elles par des liens hypertextes mais aussi une immense base de données. De ce fait, TWiki possède une fonction de recherche qui accepte le langage SQL (<http://twiki.org/cgi-bin/view/TWiki/QuerySearch>).

2) L'autre solution pour structurer et donc sémantiser l'information à l'intérieur d'une page wiki a été d'utiliser une astuce décrite par Peter Thoeny (<http://twiki.org/cgi-bin/view/Sandbox/EditActionItems>). Cette solution repose sur l'utilisation d'un tableau dont la première cellule qui compose chaque ligne contient un marqueur de l'information. Ce

marqueur identifie la ligne du tableau comme étant un ensemble d'informations structurées. Ce marqueur est entouré des codes html "<!--" et "-->" qui masquent tous les caractères compris entre ces deux balises. Visuellement, lors de la lecture des pages nous ne voyons rien mais le code source de la page contient bien ce marqueur. La machine peut alors le lire et faire des recherches à partir de ce marqueur. On retrouve là le même principe qui a été à l'origine des langages de description des données *RDFa* ou *Microdata* qui structurent l'information directement dans une page web.

Prenons maintenant un exemple pour illustrer ce concept en décrivant dans le détail comment est structurée une caractéristique technique dans ce wiki.

Voici le code source d'un tableau de caractéristiques techniques :

```

|*<NOP>*|*Symbol*|*Name*|*Nom. Value*|*Unit*
|*Min*|*Max*|*Test Method*|
|<!--CHARACTERISTIC-->|I|Elec Current|10|mA
|-0.5|+0.5|IECxxx-4|
|<!--CHARACTERISTIC-->|W|Weight|145|g|-2
|+10|ISOzzz §3.2|

```

Le rendu tel qu'il s'affiche dans un navigateur est illustré au tableau 3.

Table 3: Vue d'une information structurée dans un tableau

	Symbol	Name	Nom. Value	Unit	Min	Max	Test Method
	I	Elec Current	10	mA	-0.5	+0.5	IECxxx-4
	W	Weight	145	g	-2	+10	ISOzzz §3.2

Ce tableau est contenu dans la page d'un composant, les caractéristiques décrites dans ce tableau sont donc celles du composant de la page. Chaque ligne représente une donnée (ici une caractéristique technique) et chaque colonne une propriété de cette donnée.

La page complète ressemble alors à la figure 1.

Figure 1: Vue complète d'une page TWiki

Nom de la page :	PartOring6022																		
1. Overview																			
This page describes an O-Ring component.																			
2. Product Characteristics																			
<table border="1"><thead><tr><th>Name</th><th>Nom. Value</th><th>Unit</th><th>Min</th><th>Max</th><th>Test Method</th></tr></thead><tbody><tr><td>ID</td><td>7.67</td><td>mm</td><td>-0.02</td><td>+0.05</td><td>ISOxyyz</td></tr><tr><td>Tore</td><td>2</td><td>mm</td><td>-0.01</td><td>+0.01</td><td>ISOxyyz</td></tr></tbody></table>	Name	Nom. Value	Unit	Min	Max	Test Method	ID	7.67	mm	-0.02	+0.05	ISOxyyz	Tore	2	mm	-0.01	+0.01	ISOxyyz	
Name	Nom. Value	Unit	Min	Max	Test Method														
ID	7.67	mm	-0.02	+0.05	ISOxyyz														
Tore	2	mm	-0.01	+0.01	ISOxyyz														
3. Design and Operations																			
This o-ring is used for all fluids except fuel. It could be used with oil only for temperature below 125 degC.																			
Form:	PartForm																		
P/N:	HR64789																		
Date:	25/07/2014																		

Nous voyons alors une page wiki avec du texte libre *information non structurée* et du texte formaté, soit dans un tableau, soit dans les champs attachés à la page *information structurée*. Si un lien hypertexte relie deux pages structurées, alors nous avons une relation entre deux données structurées.

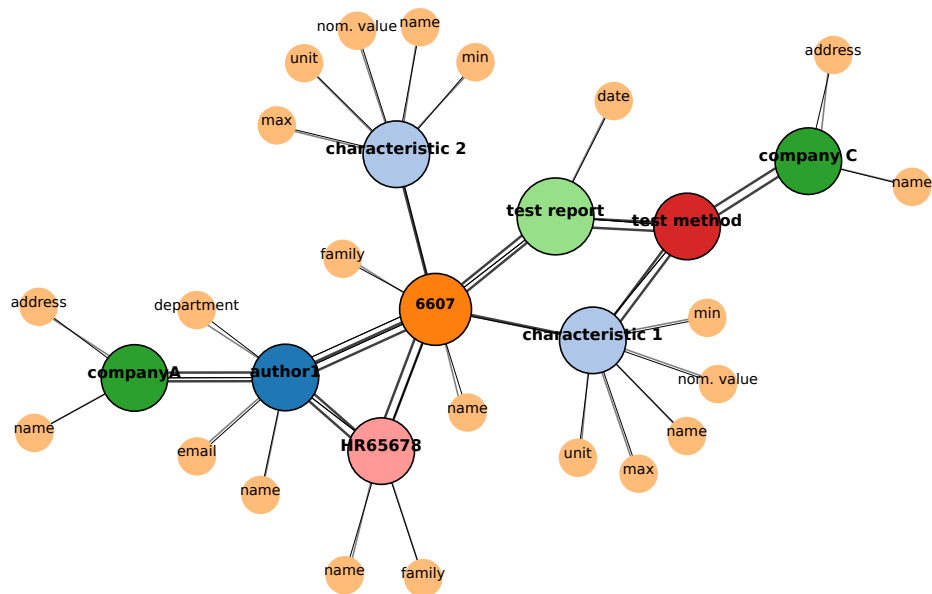
Nous pouvons aussi mieux comprendre l'imbrication des données structurées entre elles avec le graphe en réseau de la figure 2 qui met en évidence les données, leurs types (par les couleurs) et les liens entre elles.

Ontologie

Un travail important consiste à lister l'ensemble des données que l'entreprise souhaite structurer. Cet ensemble de termes et concepts représente l'ontologie de l'entreprise. Une ontologie peut être simplement définie comme étant un ensemble de terme utilisé pour représenter et décrire un domaine (ont 2014). Pour chaque donnée, il faut choisir sa méthode de structuration : page (avec champs) ou tableau (avec colonnes), et lister les propriétés nécessaires pour décrire la donnée.

Dans le cas concret de l'entreprise qui nous intéresse, une petite cinquantaine de données a été suffisante pour décrire la façon de travailler de l'entreprise. Le tableau 4 offre un exemple de rédaction d'une donnée structurée.

Figure 2: Graphe en réseau



Légende pour les noeuds :

- cercle gros = une donnée (structurée par une page ou par un tableau à ligne marquée)
- cercle petit = propriété de la donnée (structurée dans les champs de la page ou dans les cellules du tableau à ligne marquée)

Légende pour les liens :

- double lien : La relation entre les données est un lien hypertexte donc chaque donnée est une page.
- simple lien : La donnée est incluse dans la page et structurée par un tableau à ligne marquée

Table 4: Exemple de structuration

Data	Test		
Keyword	Test		
Structure	Labeled table		
Definition			
A test can be considered a technical operation or procedure that consists of determination of one or more characteristics of a given product according to a specified procedure (see test method). Often a test is part of an experiment or validation plan. The test result can be qualitative (yes/no), categorical, or quantitative (a measured value). It can be a personal observation or the output of a precision measuring instrument.			
Properties			
#	Name	Type	Definition
1	DVPnR(1)	page	link (hypertext) to DVP&R web page
2	Customer Requirement	text	Reference of customer requirement
3	Start	date	Start date of test
4	End	date	End date of test
5	Completion	number	
6	Laboratory	page	link to organization which makes test
7	Status	text	Pass, Fail or other comment
8	Test Report	page	link (hypertext) to Test Report page

(1) Design Validation Plan and Report

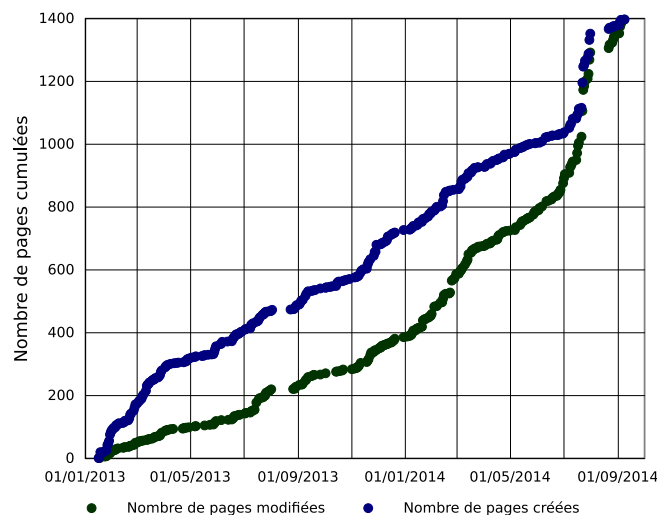
RÉSULTATS

Nous allons maintenant analyser comment ce wiki sémantique est utilisé. Quelle en est la fréquentation, le nombre de pages créées, le nombre de pages modifiées et quelques autres métriques.

Ce wiki a été mis en place dans un des sous-services de la direction R&D de l'entreprise. Ce service à dominante technique est composé de 4 personnes, le plus âgé ayant quarante ans. Le wiki a été mis en service en février 2013. En septembre 2014, 1400 pages étaient créées.

La courbe bleue du graphique de la figure 3 montre l'évolution cumulée du nombre de pages créées dans le temps. Nous pouvons observer deux accélérations : une au début de l'installation, c'est l'effort initial nécessaire pour débiter l'utilisation du wiki. La deuxième accélération en juillet 2014 est simplement due à l'embauche d'un salarié intérimaire pendant un mois pour mettre à jour quelques pages. Nous pouvons observer l'effet de son travail.

Figure 3: Nombre cumulé des pages créées et des pages modifiées dans le temps

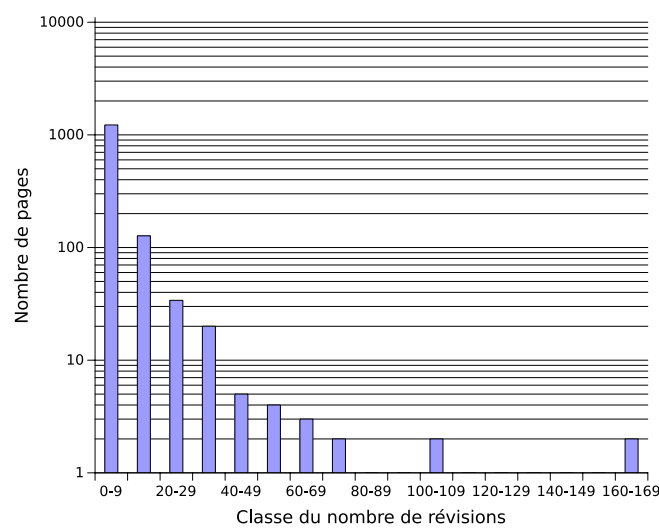


La courbe verte de la figure 3 représente la même courbe que la bleue mais chaque page n'est pas placée sur le graphique au jour de sa création mais à celui de sa modification. Plus la pente de la courbe verte est raide, plus le nombre de modifications durant la période est important. Il est intéressant de voir que la partie la plus raide de la courbe verte est la dernière partie, ceci signifie que le wiki vit : les utilisateurs l'utilisent et

modifient les pages. Plus l'écart entre la courbe verte (les modifications) et la courbe bleue (les créations) est important, plus le wiki est utilisé.

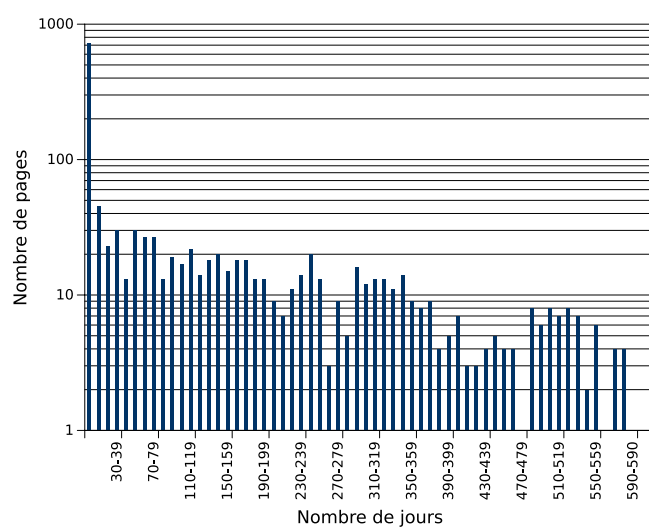
Nous allons maintenant nous intéresser au nombre de modifications ou révisions de chaque page. Il est en effet intéressant de savoir si les pages sont modifiées régulièrement ou, au contraire, si le wiki est statique. Un nombre important de pages (35%) n'a jamais été modifié et 51% des pages n'a pas connu plus de 10 modifications. Ce résultat n'est pas vraiment une surprise, il nous semblait évident que les utilisateurs travailleraient sur un nombre limité de page. De plus, certaines pages, comme la description d'une entreprise (client ou fournisseur) ne demandent pas de fréquentes mises à jour. Le graphique de la figure 4 montre la répartition du nombre de modifications.

Figure 4: Répartition du nombre de révisions



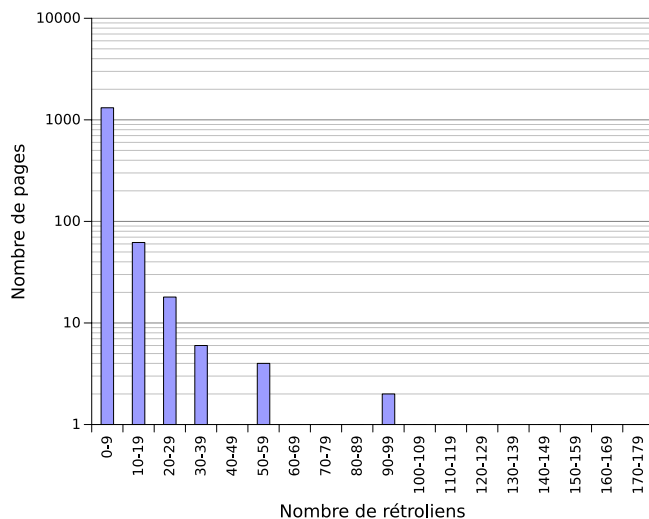
Maintenant, observons la durée qui existe entre la création d'une page et sa dernière modification. 51% des pages sont modifiées dans les 10 jours qui suivent leur création. Ensuite la répartition suit une légère tendance baissière (voir figure 5).

Figure 5: Répartition du nombre de jours entre la création d'une page et sa dernière révision



L'une des façons d'apporter du sens dans les données du wiki est aussi de relier les pages entre elles par des liens hypertextes. Plus il existe de liens entre les données, plus les données prennent du sens. Il est donc intéressant de voir si une page créée est utilisée dans une autre page. Pour cela nous étudierons la distribution du nombre de rétroliens : nombre de pages qui possèdent un lien vers la page étudiée (voir figure 6).

Figure 6: Répartition du nombre de rétroliens



63% des pages ne sont utilisées qu'une seule fois et 30% des pages sont utilisées entre 2 et 10 fois. Le tableau 5 donne un zoom sur le nombre de liens le plus faible. Plus de 1000 pages ont au maximum 4 liens. Ce sera dans l'avenir l'un des points à améliorer de ce wiki sémantique : comment obliger le rédacteur de page TWiki à mettre un nombre de liens significatifs ? L'idéal serait que le maximum se trouve dans la classe des 10 à 20 liens.

Table 5: Nombre de pages versus nombre de rétroliens

Nombre de rétroliens	Nombre de pages
0 à 4	1 158
5 à 9	141
10 à 14	46
15 à 19	21

Le lecteur aura s'en doute remarqué que quelques pages possèdent un nombre important de liens (supérieur à 50). Il s'agit en fait des pages qui définissent les types créés lors de l'ontologie : utilisateur, caractéris-

tiques techniques, entreprise, méthode d'essai, projet, rapport d'essai, plan de validation, Ces pages sont à voir comme des catégories plus que comme des informations utiles à stocker et à partager. Il est donc normal que ce type de page possède beaucoup de liens.

DISCUSSION

Qu'apporte une ontologie à l'entreprise ?

La rédaction de l'ontologie de l'entreprise a été un travail salvateur. Cette ontologie a permis de créer un vocabulaire métier commun à l'entreprise. Chaque terme de l'ontologie a fait l'objet d'une définition. Ce travail a permis de clarifier les façons de travailler de l'entreprise, de clarifier le "qui fait quoi". Nous nous sommes aussi aperçu qu'en définissant qui était responsable de la donnée, nous définissions le mode de fonctionnement de l'entreprise. Le travail suivant consistera à créer un flux de travaux (*workflow*) entre les données pour définir complètement les processus de travail de l'entreprise.

Mise en œuvre du wiki

Une fois le langage wikiML intégré, la mise en œuvre est facile. Dans le moteur TWiki, la fonction `%SEARCH{"text"...}%` permet de récupérer sous forme de tableau ou de liste à points n'importe quelle information. La fonction `%SEARCH{%}` peut aussi être imbriquée pour faire des recherches sur plusieurs critères. Toutes les inclusions d'information dans une page qui font appel à des liens sémantiques sont construites à partir de cette fonction.

Si au début de la construction du wiki l'information à structurer devait être essentiellement technique, nous avons réalisé rapidement que cela ne suffisait pas et il a fallu ajouter d'autres types d'informations. Des informations liées au mode projet, c'est à dire des dates, des saisies

d'heures passées sur les projets, des données logistiques sur la livraison des pièces reçues lors des phases de réalisation des pièces prototypes.

Quelle est l'acceptation du wiki par l'ensemble de l'entreprise ?

Tout le monde est d'accord pour dire que c'est la bonne façon de travailler. En lecture, les utilisateurs reconnaissent que l'utilisation est simple même si certains ont du mal à naviguer dans les pages. En général, ils ne savent pas ce qu'ils cherchent, il leur manque un point de départ. En écriture, les freins sont nombreux. Les réticences viennent du langage WikiML et ceci même si TWiki possède un éditeur WYSIWYG¹. La culture des outils commerciaux de traitement de texte ou de tableur est tellement ancrée dans les habitudes que le passage vers une logique *Markup Language*, ou langage de balisage, semble un obstacle insurmontable. Nous trouvons ici la vraie limitation à la pénétration des wikis dans l'entreprise (Chu 2008).

L'une des craintes la plus évoquée était la peur que les contributeurs inscrivent n'importe quoi dans les pages, voire sabotent le contenu des autres. Ici la réponse est simple : l'historique de toutes les modifications est conservé, ce qui revient pour l'utilisateur à signer toutes ses contributions, comme un courriel est signé et envoyé par une seule personne.

Les bonnes surprises de l'installation d'un wiki sémantique

Une fonction permet d'exporter une page au format PDF. Le logo de l'entreprise, le titre de la page, le nom de l'auteur de la page, le numéro de révision et la date sont automatiquement ajoutés. Il n'y a plus d'erreur possible sur la saisie de la date ou sur le numéro de révision. La diffusion des informations aux personnes ne disposant pas d'accès au wiki est ainsi rapide et sûre.

1. What You See Is What You Get : éditeur qui permet de voir directement à la volée les modifications de mise en forme

Voici une liste de fonctionnalités qui ont été ajoutées au wiki mais qui n'était pas prévues à l'origine :

- Organigramme automatique
- Tableau de bord avec indicateurs multiples
- Suivi des heures passées sur les différents projets

La souplesse du wiki et la structuration des données permet de faire de l'analyse complexe à l'exemple du *Big Data*.

CONCLUSION

Après avoir décrit le contexte de l'entreprise et ce que nous entendions par "données", nous avons vu comment il était possible de mettre en œuvre une solution de wiki sémantique avec le moteur de wiki TWiki. Ces principes peuvent être facilement transposables à d'autres moteurs de wiki. Il est maintenant évident pour cette entreprise que les technologies du web sémantiques sont une bonne réponse à la diffusion mondiale et à la gestion de ses données techniques. Cependant, les technologies web, même si elles pénètrent de plus en plus la sphère privée (notamment avec les réseaux sociaux), sont loin des habitudes de l'entreprise, des logiciels de traitement de texte ou des tableurs du commerce. Il faut donc faire un vrai travail de formation pour briser les freins des utilisateurs vers ces technologies. C'est à ce prix que l'entreprise pourra profiter pleinement de toute la force du web et de la sémantisation des données. L'une des conséquences inattendues de ce travail de sémantisation des données est l'apport de l'ontologie sur les descriptions des modes de fonctionnement de l'entreprise. Les procédures et les processus peuvent facilement être déduits de cette sémantisation en introduisant une notion de flux de travaux (workflow) entre les données structurées.

Références

chiffres-internet, jul 2014. URL <http://www.blogdumoderateur.com/chiffres-internet/>.

- Ontologies, sep 2014. URL <http://www.w3.org/standards/semanticweb/ontology>.
- Ben Adida and Michael Hausenblas. Rdfa use cases : Scenarios for embedding rdf in html. *W3C Working Draft, W3C Semantic Web Deployment Working Group*, 2007.
- Katrina Avers, Bill Johnson, Joy Banks, and Brenda Wenzel. Technical documentation challenges in aviation maintenance : A proceedings report. Technical report, DTIC Document, 2012.
- Michel Buffa, Guillaume Erétéo, and Fabien Gandon. Wiki et web sémantique. *IC 2007 : Ingénierie des Connaissances*, 2007.
- Laurence Caby-Guillet, Samy Guesmi, and Alexandre Mallard. Wiki professionnel et coopération en réseaux. *Réseaux*, (2) :195–227, 2009.
- S Chu. Twiki for knowledge building and management. *Online Information Review*, 32(6), 2008.
- Mark R Cutkosky, Jay M Tenenbaum, and Jay Glicksman. Madefast : collaborative engineering over the internet. *Communications of the ACM*, 39(9) :78–87, 1996.
- Elizabeth DaLio. Twiki-based facilitation in a newly formed academic community of practice. *WikiSym '05 Proceedings of the 2005 international symposium on Wikis*, 2005.
- Jim Giles. Fostering a data-driven culture. *Economist Intelligence Unit*, 2013.
- Bo Leuf and Ward Cunningham. *The Wiki Way : Quick Collaboration on the Web*. Addison-Wesley Professional, 2001. ISBN 078-5342714999.
- Olivia Montel-Dumont. Les entreprises dans la mondialisation. *Cahiers français*, (365), 2011.
- Charmaine C Pfaff and Helen Hasan. Overcoming organisational resistance to using wiki technology for knowledge management. *PACIS 2006 Proceedings*, page 110, 2006.
- Frédéric Simottel. La donnée, un actif stratégique de l'entreprise au service de tous les métiers. In Les Echos, editor, *LesEchosEvent*, 2014.
- Jean-Claude Tarondeau. La rapidité de développement des nouveaux produits. *Décisions marketing*, pages 71–79, 1994.
- Max Völkel, Markus Krötzsch, Denny Vrandečić, Heiko Haller, and Rudi Studer. Semantic wikipedia. In *Proceedings of the 15th international conference on World Wide Web*, pages 585–594. ACM, 2006.
- Christian Wagner and Narasimha Bolloju. Supporting knowledge management in organizations with conversational technologies : Discussion forums, weblogs, and wikis. *Journal of Database Management*, 16(2) :i–viii, 2004.

Hédia Zannad. L'individu et l'organisation projet. *Revue française de gestion*, (6) : 49-66, 2009.