

Ministère des Enseignements Secondaire et Supérieure

Université Polytechnique de Bobo

Unité de Formation et de Recherche en
Science et Technologie

Licence Statistique et d'Informatique



Ministère de la Santé

Secrétariat Général

Centre MURAZ



Rapport de Stage Professionnel en Licence Statistique et d'Informatique
Sous le Thème :

GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES »

Préparé par :

LOYARA N. A. Azize

Maitre de Stage :

SOMDA Serge Aymar

Responsable du Data-management de l'UFAM

Directeur de mémoire:

OUATTARA Yacouba

Coordonnateur de la section Informatique à l'IEAM

Date de soutenance: 13/06/2015

**LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET
GESTION DES METADONNEES ».**

Dédicace

*Je dédie ce présent document à ma famille pour
l'accompagnement, le soutien et l'éducation que j'ai pu
recevoir d'eux.*

A mon père

LOYARA Télého

A ma mère

LOYARA née **LINGANI** Akissi

A ma sœur et à mes frères

Patricia, Yves, Yannick

A mes amis et camarades

A celui-là qui nous a quitté prématurément :

LOYARA Naziko Jacques Abdine.

Tu resteras toujours dans nos cœurs et dans nos mémoires.

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

Résumé

Le présent rapport de stage expose les résultats d'un travail de gestion et d'organisation des données des projets fait au Centre MURAZ.

Ce projet a été accompli en binôme dans la collaboration, le soutien, le partage des connaissances acquises durant ces dernières années universitaires dans un esprit de travail et de cohésion avec YARA Mimbouré. Il se présente en deux parties à savoir :

- organisation et gestion des métadonnées
- conception d'une application informatique pour gérer les métadonnées.

Le sujet du travail reposait sur la thématique du data management, qui consiste à gérer et organiser les données pour aboutir à une analyse statistique optimale. Pour ce qui est de ce rapport, il fait état de la première partie qui est l'organisation et gestion des métadonnées.

Dans le présent rapport, nous avons proposé une approche conceptuelle de l'état d'organisation et de gestion des métadonnées.

Nous avons ensuite proposé une stratégie afin d'améliorer la gestion des métadonnées au niveau de l'Unité de Formation et d'Appui Méthodologique (UFAM).

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

Avant-propos

Véritable allié, indispensable de toute société en quête de performance, le métier du statisticien vient répondre aux exigences des entreprises en termes de prise de décision, d'analyse des données, de recherche d'organisation et de gestion des données.

La filière de la Licence de Statistique et d'Informatique a été instaurée par l'Université Polytechnique de Bobo. Elle est intégrée dans l'unité de Formation et de Recherche en Science Technologie (UFR/ST). Elle a ouvert ses portes au cours de l'année universitaire de 2011-2012.

L'objectif de la filière est de former des cadres intermédiaires, dans le domaine du traitement statistique et informatique des données dans des secteurs d'activité (marketing, santé, assurance banque ...) la formation vise à développer les compétences suivantes chez l'étudiant:

1. comprendre les problématiques et les enjeux dans les différents domaines d'implication de la statistique ;
2. faire preuve d'autonomie en programmation, gestion des bases de données et utilisation de l'informatique communicante ;
3. avoir une démarche critique et réfléchie devant des données, savoir identifier celles qui sont pertinentes et les analyser avec les principaux outils de la statistique.

Dans le cadre de l'obtention du diplôme professionnel, les étudiants doivent obligatoirement effectuer un stage pratique d'une durée de trois mois. Les missions du stage sont les suivantes :

1. faire découvrir le monde du travail et de l'entreprise ;
2. permettre à l'étudiant de trouver ou consolider son projet d'orientation ;
3. utiliser ses acquis scolaires.

Ce présent rapport de stage que nous présentons, témoigne du travail accompli pendant ces trois derniers mois de stage en entreprise.

Nous avons effectué notre stage au Centre MURAZ, centre de recherche en santé situé à Bobo-Dioulasso. Nous étions dans le Département des Maladie Non Transmissible (DMNT) de la structure d'accueil, plus précisément dans l'Unité de Formation et d'Appui Méthodologique (UFAM). Le stage s'est tenu du 20 Mai au 20 Aout.

L'UFAM est une jeune unité du Centre MURAZ. Son objectif est d'accompagner les équipes de recherche dans la méthodologie de montage et de mise en œuvre des projets ainsi que dans le renforcement de leurs capacités techniques et opérationnelles et d'apporter son expertise technique dans l'administration et la gestion des bases de données de recherche ainsi que dans l'analyse statistique et la valorisation des résultats des projets.

Les missions de l'UFAM ne seraient une réussite sans une organisation efficace des données de recherches. C'est dans cette optique qu'il nous a été proposé un sujet d'étude sous le thème général: la gestion des données au Centre MURAZ avec pour option : **GESTION ET ORGANISATION DES METADONNEES.**

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

Remerciements

Je tiens à remercier avant tout, toute l'équipe pédagogique de la Licence Statistique Informatique et les intervenants professionnels responsables de la formation.

Je remercie également Monsieur **OUATTARA** Yacouba mon directeur de stage pour l'aide et les conseils qu'il m'a apporté lors des différentes missions de suivi.

Je remercie particulièrement mon maître de stage, Monsieur **SOMDA** Aymar Serge, méthodologiste biostatisticien, responsable de l'Unité de Formation et d'Appui Méthodologique (UFAM/DMNT) pour son accueil, sa disponibilité, la confiance qu'il m'a accordés dès mon arrivée, toutes choses qui ont rapidement facilité mon intégration dans l'entreprise. C'est le lieu pour moi de lui témoigner toute ma reconnaissance pour l'expérience enrichissante et pleine d'intérêt qu'il m'a fait vivre durant ces trois-mois au sein de l'institut du Centre MURAZ.

Mes remerciements s'adressent également aux personnes suivantes :

- Monsieur **BADOLO** Herman démographe à l'UFAM pour sa lecture, ses appréciations et son accompagnement apporté dans la rédaction de ce présent rapport en l'absence de Serge **SOMDA** ;
- Monsieur **Diallo** Ibrahim data manager de l'UFAM pour m'avoir accordé son temps, sa disponibilité et sa patience sur les concepts et technique d'organisation et de gestion des données;
- le responsable du service Informatique du Centre MURAZ, Monsieur **ZOUGRANA** Moumouni pour son temps et ses remarques pertinentes dans la mise en application de la méthode MERISE;
- l'ensemble du personnel de l'UFAM pour leur accueil sympathique et leur coopération professionnelle tout au long de ces trois mois.

Pour clôturer la liste, je tiens à remercier du fond de mon cœur ma famille, mes amis et camarades pour leur soutien tant acharné durant toutes ces années.

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

Table des matières

Dédicace	I
Résumé	II
Avant-propos	III
Remerciements	IV
Sigles et abréviations	V
Liste des figures	VI
Présentation de la Structure d'accueil	7
1.1 Historique du Centre MURAZ	8
1.2 Organigramme du Centre MURAZ	8
1.3 Missions du Centre MURAZ	8
1.4 L'Unité de Formation et d'Appui Méthodologique (UFAM)	9
2 Introduction et problématique	10
2.1 Contexte et justification	11
2.2 Problématique	11
3 Matériels et Méthodes	12
3.1 Fonctionnement de l'UFAM	14
3.1.1 L'appui méthodologique	14
3.1.2 La formation continue	14
3.1.3 Le traitement, gestion et stockage de données	14
3.1.4 L'analyse des données	14
3.2 Approche méthodologique	14
4 Résultats	15
4.1 Gestion des données à l'UFAM	16
4.2 Le recensement des données de recherche du serveur	16
4.3 Le modèle de gestion des métadonnées du centre MURAZ	16
4.4 La méthode d'organisation physique	17
4.4.1 Spécificités du modèle	18
4.4.2 Stockage physique des fichiers dans le serveur	19
4.4.3 Dictionnaire	19
5 Discussion	21
6 Conclusion	22
7 Bibliographie	23

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».**Sigles et abréviations**

ADISP : Archives de Données Issus de la Statistique Publique
D.S : Direction Scientifique
DATA : Données
DM : Data Manager
DMNT : Département des Maladies Non Transmissible
DSP : Département de Santé Public
EPA : Etablissement Public à caractère Administratif
E-P-S : Etablissement Public de Santé
OMS : Organisation Mondiale de la Santé
PC : Personal Computer
PDA : Personal Digital Assistant (Assistant Numérique Personnel).
PNUD : Programme des Nations Unies pour le Développement
PROTO : Protocol
SGAMS : Service Général Autonome de la Maladie du Sommeil
UFAM : Unité de Formation et d'Appui Méthodologique
UFR/ST : Unité de Formation et de Recherche en Science et Technologie
UPB : Université Polytechnique de Bobo-Dioulasso

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

Liste des figures

FIGURE 1: ORGANIGRAMME DE LA DIRECTION SCIENTIFIQUE	1-9
FIGURE 2: RESEAU D'UFAM	4-16
FIGURE 3: FORMULAIRE ET SOUS FORMULAIRE.....	4-17
FIGURE 4: SCHEMA RELATIONNEL.....	4-17
FIGURE 5:MODELE DE REPRESENTATION DES METADONNEES DES PROJETS	4-18
FIGURE 6: MODELE PHYSIQUE.....	4-20

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

1 Présentation de la Structure d'accueil

1.1 Historique du Centre MURAZ

Le Centre MURAZ a vu le jour en 1939, sous la désignation de « **Service Général Autonome de la Maladie du Sommeil (SGAMS)** ». Sa première mission était de faire reculer la trypanosomiase humaine ou maladie du sommeil qui sévissait dans toute l'Afrique Occidentale Française (AOF) et tout particulièrement à Bobo-Dioulasso. C'est en 1956 SGAMS prendra le nom de "**Centre MURAZ**" en l'honneur de son premier directeur, le médecin militaire Français Gaston MURAZ.

Ensuite, et pendant longtemps, la recherche au Centre MURAZ s'est intéressée à la lutte contre d'autres grandes endémies qui constituaient les principales préoccupations sanitaires de l'époque.

C'est ainsi qu'au fur et à mesure que la vaccination de routine faisait reculer certaines de ces pathologies, le **Centre MURAZ** s'adaptait au contexte et aujourd'hui, il s'intéresse essentiellement à des maladies telles que le paludisme, les IST/VIH/SIDA et maladies associées (telle que la tuberculose), la fièvre jaune et la méningite.

1.2 Organigramme du Centre MURAZ

Depuis 2001, le **Centre MURAZ** est un Etablissement Public de Santé (EPS) placé sous la tutelle technique du Ministère de la Santé et celle, financière du Ministère des Finances.

Il a été d'abord érigé en Etablissement Public à caractère Administratif (EPA) puis est devenu en 2006 un EPS doté d'un statut particulier à partir de 2008. Ces transformations institutionnelles ont été imprimées par les pouvoirs publics pour garantir au **Centre MURAZ** les conditions d'une pleine contribution par la recherche au développement sanitaire du Burkina Faso.

Le centre est ainsi placé administrativement sous l'autorité du Secrétariat Général du Ministère en charge de la Santé.

Pour accomplir ses missions, le **centre MURAZ** peut selon les besoins :

1. créer des directions, des unités de recherche et/ou de production ;
2. assurer des prestations de services à but lucratif ;
3. exploiter des brevets et licences ;
4. conclure des conventions/accords de coopération avec d'autres institutions de recherche au niveau national, régional et/ou international ;
5. participer à des regroupements d'intérêt public avec des personnes physiques ou morales, publiques ou privées, nationales ou internationales.

Les organes d'administration et de direction sont : le Conseil d'Administration (CA), la Direction Générale, les cadres de concertation et les organes consultatifs.

Le personnel du centre constitue une équipe pluridisciplinaire d'environ quatre cents (400) personnes.

Le **Centre MURAZ** possède un nouvel organigramme dans la recherche scientifique. La Direction Scientifique coiffe ainsi trois départements de recherche.

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

1. Le département des maladies transmissibles (DMT).
2. Le département des maladies non transmissibles (DMNT).
3. Le département de santé publique (DSP).

Chaque département est constitué d'unités de recherches. Ces 3 grands départements seront chargés de la coordination de la recherche, de la prospective, de l'animation et de l'évaluation au niveau des unités de recherche et seront de ce fait, l'interface entre les unités de Recherche et la Direction Scientifique.

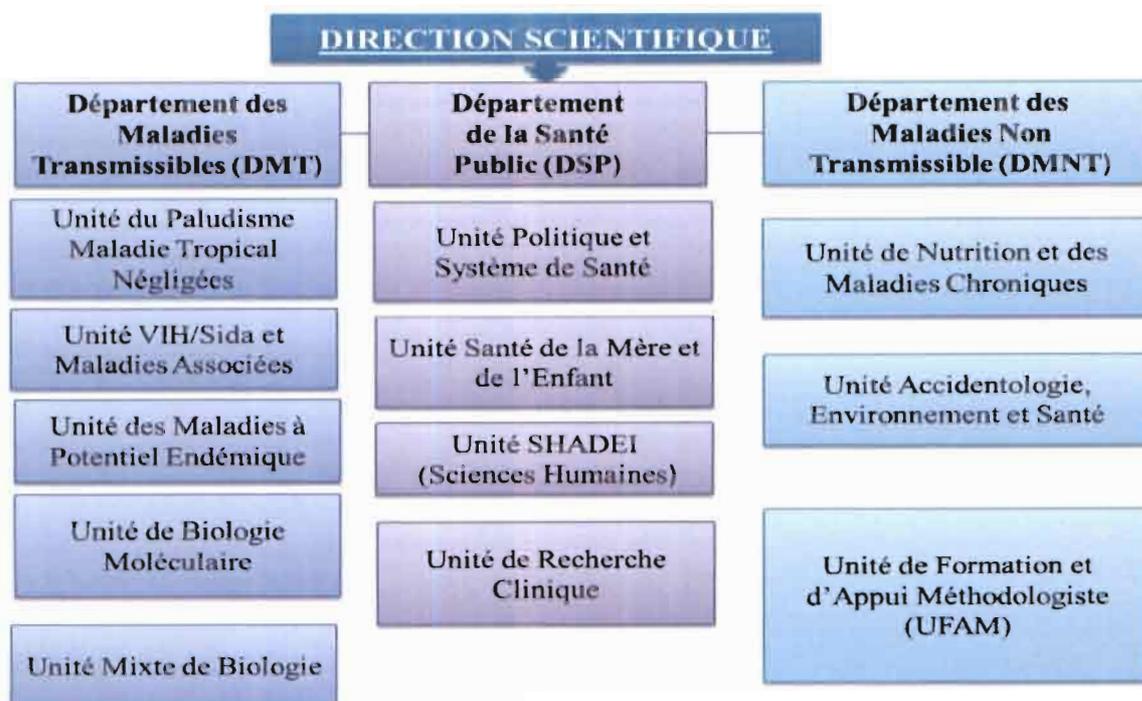


Figure 1: Organigramme de la Direction Scientifique

1.3 Missions du Centre MURAZ

Les missions essentielles du **Centre MURAZ** sont axées selon trois axes stratégiques. Il s'agit de:

La recherche :

Mission première et aussi la plus connue du **centre MURAZ**. Elle est organisée autour de huit axes de recherche.

La formation :

Participation effective à la formation des personnels de santé à travers des stages pour paramédicaux, doctorants (mémoires, thèses), post-doctorants.

L'expertise :

Renforcement des capacités des compétences nationales (en priorité) de sorte à disposer sur le plan technique d'experts de haut niveau. De plus, le centre est sollicité par des organismes tels

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

que l'OMS, L'UNICEF, la Coopération française, le PNUD dans des domaines aussi variés que possible comme l'animation de séminaires, de formation, la planification et l'évaluation des programmes, le contrôle de qualité de laboratoire etc.

1.4 L'Unité de Formation et d'Appui Méthodologique (UFAM)

Notre stage s'est déroulé dans l'Unité de Formation et d'Appui Méthodologique (UFAM), dans le Département des Maladies Non Transmissibles. Il s'agit d'une unité, ambitieuse, participative, aux objectifs de développement du **Centre MURAZ**.

L'UFAM a été créée pour répondre aux missions de recherche, d'expertise et de formation dans une perspective d'excellence et pour répondre au mieux aux priorités nationales et internationales au bénéfice de la santé des populations. Ses objectifs sont les suivants :

- 1 assurer l'accompagnement méthodologique dans la conception, la mise en œuvre et la valorisation de la recherche ;
- 2 assurer la formation des acteurs de la recherche et des professionnels en méthodologie de la recherche ;
- 3 apporter son expertise scientifique et technique en méthodologie de la recherche ainsi qu'en organisation, gestion et analyse des données quantitatives.

Pour se faire l'UFAM dispose d'un système informatique permettant la saisie, la gestion et le traitement des données ainsi que d'une salle d'archive pour le stockage des données physiques.

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

2 Introduction et problématique

2.1 Contexte et justification

Au cours de ces dernières décennies on a constaté une explosion, à l'échelle mondiale, des données numériques produites par les chercheurs dans de nombreuses disciplines scientifiques. Pour faire face à ce « déluge de données », les institutions de recherche ainsi que leurs tutelles doivent relever des défis de plusieurs ordres : organisationnel, normatif, juridique, technologique et financier. A cet effet, la gestion des données devient aujourd'hui l'une des fonctions primordiales des entreprises.

Le métier du Data Management ou la gestion de données qui est une discipline de gestion regroupant les différents processus mis en place pour définir, stocker, maintenir et garantir une prise en compte correcte et fiable des données au sein d'un référentiel vient pour répondre aux défis organisationnels et de gestion des données.

En effet, le Data Manager se définit comme le responsable technique d'entretien, et de la normalisation et juridique de fichiers de données. Ce métier peut se retrouver dans tous les secteurs d'activité qui génèrent de grande quantité de données. La mission du Data Manager est d'organiser de manière optimale les données pour faciliter la recherche d'informations à des fins d'analyse et de prise de décision. Pour que cela soit possible le data manager se doit de mettre en place un système de gestion qui puisse répondre aux exigences du métier. C'est-à-dire :

1. mettre en place un référentiel de gestion des données, une sorte de lexique au niveau de l'entreprise ;
2. garantir l'unicité de la donnée et de son descriptif ;
3. centraliser le cycle de vie des donnée : création, mise à jour et suppression;
4. structurer les données de la recherche dans une base ;
5. vérifier qu'elles sont cohérentes, les préparer pour une analyse statistique afin de pouvoir les interpréter.

La bonne gestion et le meilleur partage des données de recherche est important pour le Centre MURAZ dans la mesure :

1. d'accélérer le progrès scientifique :
la bonne gestion des données de recherche permettra aux chercheurs d'obtenir et de comprendre les données des autres et de les réutiliser pour leurs propres fins scientifiques, ce qui accélérera les nouvelles découvertes ;
2. d'accroître la visibilité et les retombées de la recherche :
les données rendues visibles par l'entremise d'un dépôt de données peuvent de beaucoup accroître les retombées de cette recherche;
3. de respecter les politiques des organismes subventionnaires :
de plus en plus d'organismes subventionnaires exigent que les chercheurs et les établissements hôtes conservent, gèrent et rendent accessibles leurs données à la fin d'un projet de recherche ;
4. d'éviter la répétition des travaux de recherche :
lorsqu'un ensemble de données est disponible publiquement, il est beaucoup plus probable que ces données ne seront pas créées de nouveau, ce qui évite des travaux coûteux et inutiles de collection et de production des données ;

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

5. permettre la reproduction et la validation des résultats de la recherche :
lorsque les données sont archivées et connues, les résultats peuvent être répétés et les données peuvent servir à une nouvelle analyse, ce qui permet de confirmer les premières constatations de la recherche. Elles peuvent aussi servir à faire état d'erreurs ou d'incohérences dans la première analyse des données ;
6. d'intensifier la collaboration :
les données disponibles publiquement permettent aux chercheurs de collaborer les uns avec les autres en mettant en commun les ensembles de données, le contexte et les outils de recherche.

Durant notre période de stage avec l'assistance des gestionnaires de données de l'UFAM nous avons eu à constater ensemble des difficultés dans l'infrastructure de conservation, et d'organisation des données des projets de recherche.

En effet la conservation des données est l'une des étapes les plus importantes d'une enquête statistique. Mais fort malheureusement, elle est la phase la plus souvent ignorée.

La conservation des données d'étude des projets à l'UFAM manque d'efficacité dans la mesure où elles ne permettent pas:

1. d'élaborer un index pour les fichiers;
2. de mettre au point une structure de stockage dans la machine principale d'archivage;
3. d'envisager des procédures efficaces de sauvegarde;
4. de viser des procédures physiques et techniques de sécurité; et
5. de permettre le contrôle continu de toutes les procédures visées.

2.2 Problématique

La métadonnée se définit comme une « donnée servant à définir ou décrire une autre donnée, quel que soit son support (papier ou électronique) ».

La mise en application d'un système de gestion des métadonnées est d'une très grande importance au sein de l'UFAM. Les métadonnées joueront les différents rôles à savoir :

1. de permettre aux gestionnaires de données de naviguer rapidement à travers des référentiels de données afin de trouver, de gérer et de suivre l'information ;
2. de permettre la création des associations entre différents objets et utilisateurs à travers une ou plusieurs applications, tout en ajoutant une consistance dans la façon dont l'information est utilisée, stockée et partagée ;
3. de fournir une grande précision sur les origines et les historiques de chaque donnée;
4. d'organiser et suivre l'ensemble du cycle de vie numérique des informations importantes, y compris les procédures, les processus et les utilisateurs à qui les tâches sont affectées ;
5. d'offrir une traçabilité précise, protéger et organiser les données;
6. d'habiliter à classer et à organiser l'information, de manière plus efficace, ainsi que de donner un aperçu rapide sur les actions qui ont lieu au sein de l'unité ;
7. de donner une meilleure qualité de l'information, des automatismes, la conformité aux normes, le partage des données, la collaboration et plus encore.
8. permettre aux utilisateurs (data manager) de rechercher des informations sur un projet

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

de la manière la plus logique pour eux (par exemple, le nom du projet, la date, le type du projet, le type de l'étude, etc.).

L'ex Centre de Calcul aujourd'hui devenue UFAM avait précédemment tenté la mise en œuvre d'un système de représentation des métadonnées des projets. Il s'agissait d'une méthode basique créée sur les logiciels tels qu'Excel, Word, et MS-Access qui tentait de faire une brève description des projets. Mais forts malheureusement ces différents procédés de représentation des métadonnées se sont révélés insuffisants. Ils étaient surtout basés sur des systèmes statiques, donc permettant juste d'enregistrer des métadonnées pour un nombre limité de projets déjà effectués, ne prenant pas en compte les projets à venir. D'où leurs échecs.

Cependant des propositions ont été faites pour l'adoption d'un logiciel d'organisation et de gestion des métadonnées des projets. Il s'agissait entre autre de la mise en application d'un logiciel libre ou payant qui utilise un programme du web sémantique et intégrer d'une base de données répondant à la gestion des contenus d'entreprise.

Comme exemple nous avons :

1. IMIS : plate-forme d'interaction entre producteur et utilisateur de données via l'Internet. Il est libre et est utilisé par l'INSD.
2. Nesstar : qui est une application du web sémantique pour les données et les métadonnées utilisé par ADISP.
3. M-Files : Application du web sémantique est utilisé par des milliers d'organisation dans la gestion du contenu de leurs documents.
4. WinDev : Programme informatique de gestion des données et du personnel d'entreprise généralement utilisé par les banques et les assurances.

L'UFAM se propose aujourd'hui de mettre en œuvre un système de gestion de métadonnées fait maison. Ce système devra permettre d'organiser les métadonnées des projets afin de faciliter le travail des data managers. Nous avons effectué ce travail en deux étapes, à savoir définir le système et le mettre en œuvre.

La partie du travail qui est ici présentée concerne l'approche conceptuelle pour la création d'un système de gestion des métadonnées.

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

3 Matériels et Méthodes

3.1 Fonctionnement de l'UFAM

L'Unité de Formation et d'Appui Méthodologique s'est donné de nombreux défis à relever : adaptation à des contraintes légales nationales, africaine ou internationales, lancement de nouveaux services innovants, intégration de nouvelles technologies, ou la mise à jour de technologies déjà implantées. L'unité s'organise suivant les objectifs suivants: Appui méthodologique à la recherche, formation continue, traitement, gestion et stockage des données, l'analyse statistique des données dans l'optique de répondre efficacement aux missions, et attribution qui lui sont assignées.

3.1.1 L'appui méthodologique

L'UFAM accompagne les projets de recherche en leur offrant son expertise méthodologique. Cet appui peut prendre plusieurs formes :

1. conception d'un projet de recherche (rédaction de protocole, définition de schémas d'études, calcul du nombre de sujet nécessaires...);
2. élaboration des outils du projet (outils de collecte, outils informatique de gestion) ;
3. appui pour la mise en œuvre des projets (organisation des opérations de terrain) ;
4. appui pour la valorisation de la recherche ;
5. etc.

3.1.2 La formation continue

L'UFAM propose de plus à la communauté des chercheurs (jeunes, seniors, étudiants, travailleurs administrateurs...) des séances de formation continues à la méthodologie de la recherche aussi bien dans divers domaines de la recherche. Un programme présentant au moins un atelier de formation par mois a ainsi été élaboré pour l'année 2014.

3.1.3 Le traitement, gestion et stockage de données

Le traitement, la gestion et le stockage de données est assuré par les data managers de l'unité. L'UFAM dispose pour ce faire d'une salle de saisie et d'une équipe d'opératrices de saisie pour l'encodage des données. Elle a de plus développé une expertise dans la capture automatique des données au moyen d'outils numériques (PDA, Smartphones, tablettes tactiles) sous les environnements Windows mobile et Androïde. Les principaux logiciels utilisés sont EPI Data, CS Pro, MS Access, Pendragon, Opendatakit.

L'UFAM administre les données de la quasi-totalité des projets du Centre MURAZ ainsi que de nombreux projets extérieurs, parfois hors du Burkina Faso.

3.1.4 L'analyse des données

L'analyse statistique des données de projet est effectuée selon un plan d'analyse conçu d'accord entre les investigateurs et les statisticiens. L'UFAM dispose de l'expérience et de la compétence pour proposer des modèles bio statistiques ou biomathématiques avancés. Les logiciels statistiques les plus couramment utilisés sont Stata et R.

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

3.2 Approche méthodologique

Afin de proposer un système efficace et adapté au contexte de l'UFAM, nous avons suivi la démarche méthodologique suivante.

Nous avons tout d'abord établi un état des lieux de l'existant au niveau de l'unité. Cela a consisté à décrire le réseau informatique mis en place. Nous avons ensuite recensé le matériel informatique disponible pour la gestion des données et notamment les serveurs que nous avons décrit. Nous avons enfin effectué un recensement exhaustif de l'ensemble des informations stockées dans le serveur concernant les projets de recherche.

La démarche suivante a consisté à étudier les expériences vécues par d'autres institutions pour la gestion de leurs métadonnées. Nous avons, pour ce faire, recherché sur Internet les informations sur les systèmes de gestion des métadonnées de plusieurs structures burkinabé et des institutions internationales.

Les caractéristiques propres du Centre MURAZ ainsi que les leçons apprises des autres structures nous ont permis de formuler des propositions sur le système de gestion des métadonnées à l'UFAM.

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

4 Résultats

4.1 Gestion des données à l'UFAM

Les différents départements du Centre MURAZ ne sont pas interconnectés. L'UFAM possède son propre réseau de communication. Ce réseau se caractérise par une conjonction de câbles reliés à trois Switch avant d'être relié aux différents serveurs de l'unité. Le réseau utilise un routeur pour connecté les différents machines à internet (câble, wifi). Il y'a plus de 20 machines qui sont connectées sur le serveur.

L'unité a une salle spécialement équipé de quatre 4 serveurs de marque Dell-PowerEdge 700, d'une capacité de 400 Go chacun. Ils sont universels et répondent parfaitement aux besoins de la structure. Ces serveurs supportent le RAID SATA qui contribue non seulement à accroître la disponibilité du système mais améliorent également les performances des disques. Ces serveurs sont incorporés d'un système d'exploitation Windows Server 2003. Il y'a deux serveurs qui ne sont pas en service. Quant aux serveurs qui fonctionnent il y'a un serveur principal et l'autre joue le rôle de relais de sauvegarde, pour plus de sécurité des données.

Les disques des serveurs sont partitionnés en deux : C et D. La sauvegarde des données dans le serveur n'est pas automatique elle se fait de façon manuelle et généralement les données sont stockées dans le disque D qui a une capacité de 200.92 Go. Les administrateurs définis accèdent aux données enregistrés à travers un mot de passe et un nom d'administrateur.

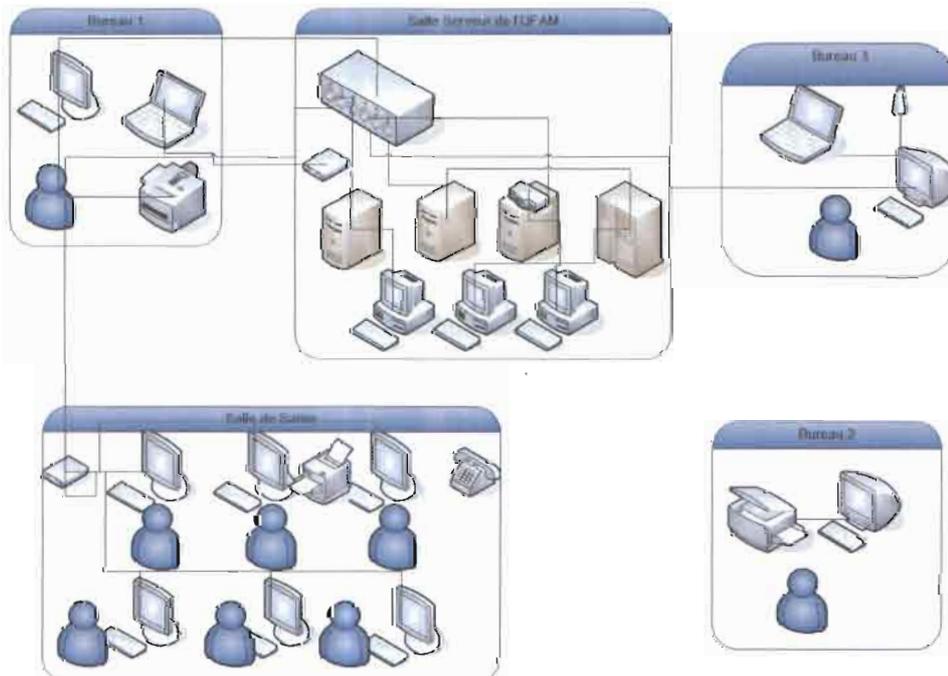


Figure 2: Réseau d'UFAM

4.2 Le recensement des données de recherche du serveur

Nous avons effectué le recensement des différentes données de recherche présentes dans le serveur de données de l'UFAM afin d'établir un état des lieux sur l'existant. Ce recensement a été effectué au moyen d'une base de données construite sous MS Access. La figure 3

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

présente un exemple de formulaire pour le renseignement de cette base tandis que la figure 4 présente un exemple d'association entre deux tables de la base.

Figure 3: Formulaire et sous formulaire

Ce recensement a permis de répertorier l'ensemble des fichiers de données, des protocoles de recherche, des rapports et articles, etc. stockés pour chaque projet. Les emplacements desdits fichiers ont été notés afin de permettre de les retrouver aisément.

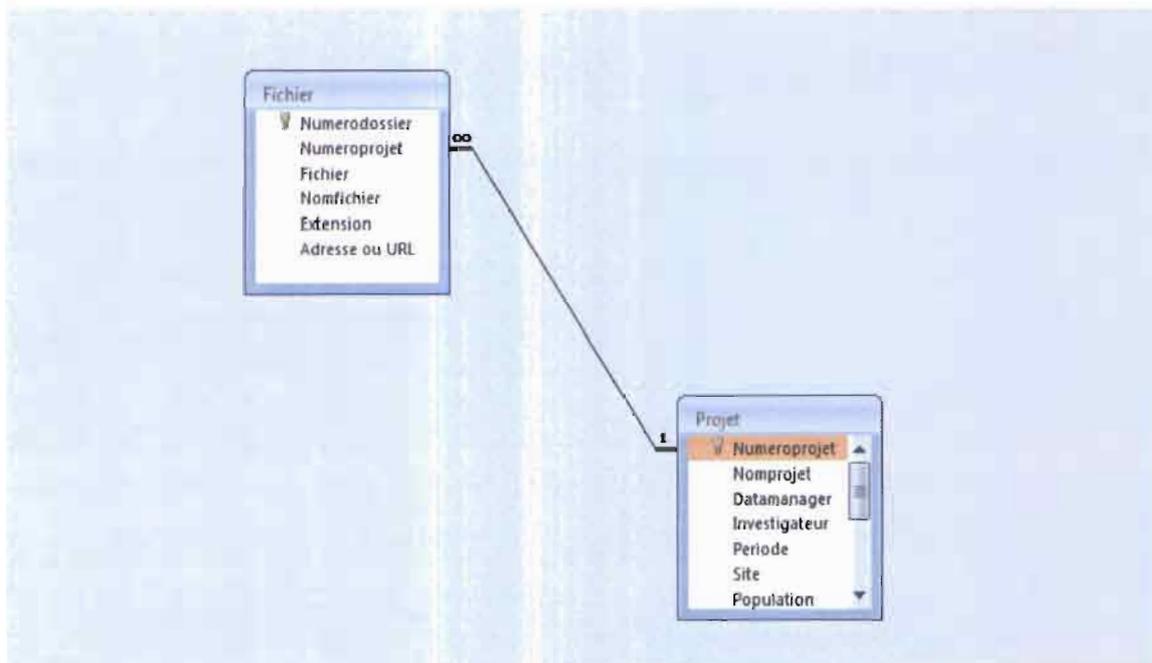


Figure 4: Schéma Relationnel

4.3 Le modèle de gestion des métadonnées du centre MURAZ

Le modèle de gestion des métadonnées du centre MURAZ s'inspire du modèle de

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

représentation des métadonnées du Dublin CORE (voir figure 6). Composé de plusieurs items qui ont pour but de d'écrire un projet dans son intégralité à travers les points focaux qui permettrons pour chaque gestionnaire de l'identifier comme il le faut.

Le modèle de Dublin CORE se définit comme un schéma qui permet de représenter la métadonnée suivant 15 items (le titre, le nom, l'ouvrage, le propriétaire....) bref tout ce qui peut permettre d'identifier un objet données.

Métadonnées d'un Projet

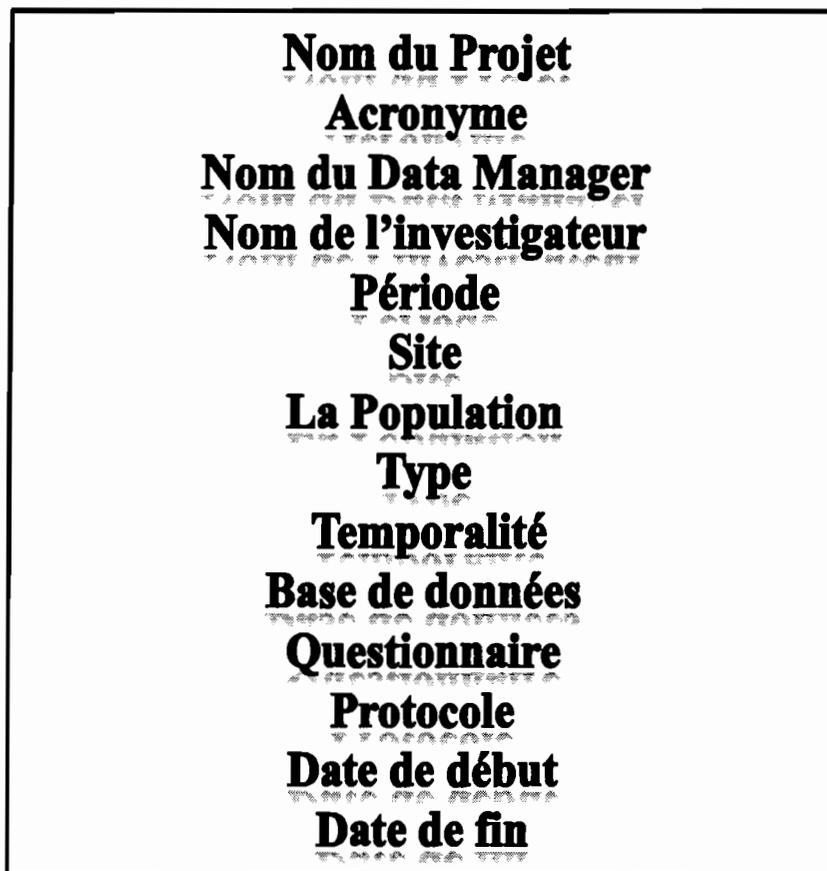


Figure 5:Modèle de représentation des métadonnées des projets

4.4 La méthode d'organisation physique

Le modèle physique présente le travail d'organisation qui a été mené sur le serveur. L'objectif visé par l'application de la méthode était de centraliser les données dispersées au niveau du serveur à une adresse url servant de référence pour tous les projets. Pour que cette vision soit possible, il a été question de faire un compte rendu de l'état exact des données existants sur chaque projet étudier dans le temps par l'établissement à travers le recensement exhaustive,

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

par la suite passé à la reconstruction et à la restauration des projets dont les fichiers étaient en mauvais état, où qui présentaient quelques informations manquantes (le type de l'étude, la temporalité, le gestionnaire des données...).

Notons que l'organisation physique est un procédé qui découle des exigences, et des besoins de l'entreprise. Elle montre une arborescence du contenu des projets et a été la première phase conceptuelle du projet de stage d'étude. Dans le souci de satisfaire les recommandations du personnel. Il a été préférable d'adopter une organisation structurelle, hiérarchisée. En effet, ce modèle hiérarchique a permis de rendre compte sur l'état des métadonnées de tous les projets.

4.4.1 Spécificités du modèle

1. assurer une bonne identification et archivage des projets ;
2. présenter une simplicité pour être mise à jour ;
3. réduire le temps de recherche ;
4. faciliter l'exploitation et la sauvegarde des données ;
5. garantir la pérennité des données.

4.4.2 Stockage physique des fichiers dans le serveur

Le stockage physique des fichiers dans le serveur sera défini selon un protocole standard afin d'assurer leur lisibilité et de permettre une navigation plus intuitive. L'architecture suivante sera retenue pour la partition de données du serveur :

1. Le niveau 1 : Comportera un dossier principal nommé PROJETS. Il fera référence à tous les autres projets d'étude.
2. Le niveau 2 : Représente la vue d'ensemble des projets.
3. Le niveau 3 : Chaque projet comportera deux sous dossiers à savoir DM et DOC.
4. Le niveau 4 : les sous dossiers DM et DOC seront constitués également de quelques sous dossiers nommées respectivement DATA, PROG, FINAL CR, et PROTO.
5. Le niveau 5 : Sera l'accomplissement final du travail. L'utilisateur pourra accéder à tous les fichiers de types (Access, PDF, EXCEL, Word, Stata, TEX, etc.) qui l'intéresse.

4.4.3 Dictionnaire

Ce petit dictionnaire de sigle est mis à la disposition des utilisateurs, dans le souci de faciliter la compréhension du schéma.

1. DM : Data Managent

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

2. DOC : Documents
3. DATA : Bases de Données
4. PROG : Programmes
5. FINAL : pour tout ce qui a été retenu comme documents final sur un projet
6. PROTO : protocoles, dossiers d'appels d'offre
7. CRF : Questionnaire, cahiers d'observations...

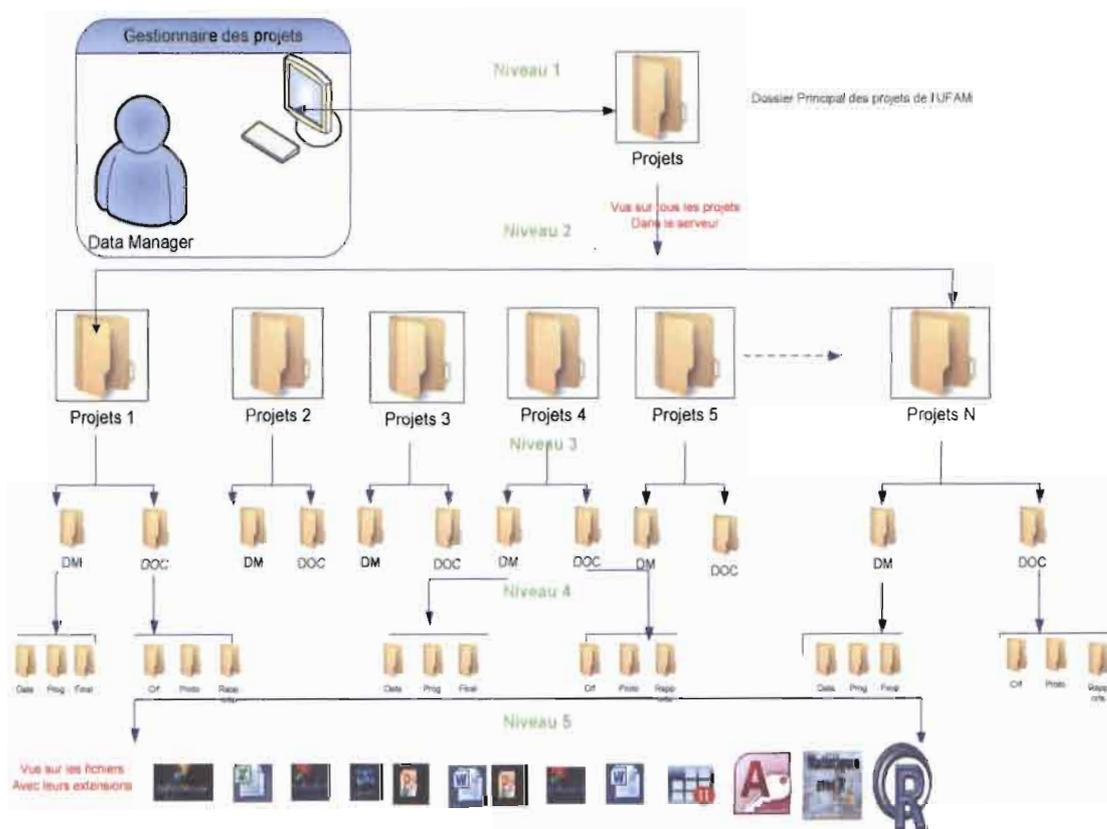


Figure 6: Modèle Physique

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

5 Discussion

Le besoin d'organisation et de planification du travail du data manager est une nécessité pour l'unité. Dans le souci de répondre aux exigences de l'institut en besoin de soutien et d'appui méthodologique en accompagnement des chercheurs dans les différentes disciplines de recherche en santé.

Très souvent les enquêtes importantes menées par l'unité du centre MURAZ ne débouchent que sur un nombre limité de rapports basés sur quelques tables, alors que les données brutes utilisées pour l'élaboration de ces rapports ne sont guère exploitées. De telles données, cependant, devraient rester disponibles pour une utilisation plus large par des analystes secondaires et cela exige que soit établie une stratégie clairement définie pour une conservation et une diffusion efficaces.

L'UFAM devrait pouvoir le gérer, l'archiver, l'historier de façon définitive est une réussite gagnante dans l'atteinte de ses objectifs. Dans la mesure où il doit pouvoir être retrouvé, ressorti et consulté pour apporter une preuve. En effet un document sans mise à jour n'est-il pas un document mort ?

Par ailleurs recueillir, classer, conserver et protéger des volumes considérables de données et de documents pour les restituer et les exploiter, tels sont les enjeux des data manager et les obligations auxquelles les gestionnaires de données de l'UFAM sont confrontées.

Conserver sur des supports évolutifs ou sur des supports qui « passeront » avec l'évolution technologique ne peut pas être satisfaisant pour des documents qui doivent être conservés sur du long terme. Il faut bien évidemment s'assurer également de la relecture d'un document enregistré. Pouvoir mettre en application des systèmes de sauvegarde externe tels que les CD-Rom, les disques durs externes qui pourront faire état de l'existence en cas de perte de données car la fiabilité dans la gestion des données est primordiale pour l'unité.

La preuve est que un document mal indexé, sauvegardé sur un support altéré ou obsolète, dont il n'existe plus de moyen de lecture par exemple, est alors partiellement ou totalement inexploitable et bien sûr sans garantie d'intégrité.

Chaque défi d'organisation et de gestion des données des projets doit être relevé par chaque gestionnaire sous forme de projet qui peut se définir comme une suite d'actions délimitée dans le temps, en vue de produire un résultat bénéfique dans la recherche en santé.

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».**6 Conclusion**

Le stage que nous avons effectué a été une expérience enrichissante pour moi, tant sur le plan d'acquisition de nouvelles connaissances professionnelles, que sociales. En réalité au début de ce stage je n'avais aucune notion sur ce qui voulait dire «gestion des données ». Grâce, à la collaboration participative de l'équipe, à leur dynamisme, leur convivialité, l'ambition débordante de nos encadreurs j'ai pu m'adapter et j'ai acquis toutes les connaissances théoriques et pratiques relative au traitement des métadonnées.

Tout le privilège a été pour moi d'intégrer une unité regorgeant de savoir-faire et de savoir être au cours de nos premiers pas en entreprise. En effet nous avons eu à renforcer nos connaissances en data-management, en méthodologie de gestion de base de données, la conception de site web implantée en local basée sur du PHP, MySQL serveur aussi nous avons découvert ce qu'est un projet de recherche en santé.

Il y eut bien évidemment des difficultés auxquelles j'ai été confronté à savoir : la mauvaise qualité des métadonnées, les problèmes à pouvoir structurer les métadonnées existantes. Mais avec la collaboration des intervenants aux différents projets de recherche ces difficultés ont pu dans la majorité être surmontées.

Je garde en moi, une immense satisfaction, de reconnaissance pour les personnes qui ont œuvré à la mise en application de ce travail et également aux différents atouts qui sont un plus pour moi en matière de connaissances.

LA GESTION DES DONNEES AU CENTRE MURAZ : « ORGANISATION ET GESTION DES METADONNEES ».

7 Bibliographie

1. Betremieux, A. (2009). *Les métadonnées: définition et présentation générale*. Paris, 75007, France.
2. DURABLE, O. N. (juin 2004). *POUR COMPRENDRE ET UTILISER LES INDICATEURS DE SUIVI DE LA PAUVRETE ET DU DEVELOPPEMENT HUMAIN*.
3. Geneve, B. i. (decembre 2003). *Traitement des données et stockage des fichiers électroniques concernant une enquête sur le travail des enfants*.
4. Leresche, F. (21 juin 2013). *Dublin Core*. Paris.
5. Perkins, A. (2014). *Les Clés du Succès en Gestion des Métadonnées*. Canada.
6. hearer, K. (2009). *les données de recherche: un potentiel insoupçonnée*. Canada.