



Valeur de l'optimisation des contenus pour les moteurs de recherche d'un portail de  
documentation technique

par

Roberto Roman

Essai présenté au CeFTI

en vue de l'obtention du grade de maître en technologies de l'information

(Maîtrise en génie logiciel incluant un cheminement de type cours en technologies de  
l'information)

FACULTÉ DES SCIENCES  
UNIVERSITÉ DE SHERBROOKE

Longueuil, Québec, Canada, mai 2016



## Sommaire

Pendant les vingt dernières années, Internet s'est élargi jusqu'à un point que l'on n'aurait jamais osé imaginer lors de sa création, passant d'un taux de pénétration sur la population mondiale de 0,8 % en 1995 à 14,1 % en 2004 puis à 40,4 % en 2014 [16].

Cette croissance a eu un impact sans précédent sur le monde des affaires et les corporations ont dû assurer leur présence sur Internet, dans beaucoup de cas sans trop savoir comment tirer avantage de celle-ci, ni obtenir un retour sur l'investissement. Quelle serait la valeur ajoutée et comment enligner cette présence avec les objectifs corporatifs?

Avec l'expansion d'Internet, les outils de recherche d'information durent évoluer, la croissance des sites, documents et données étant exponentielle, il fallut créer de nouveaux algorithmes de classement, de nouvelles méthodologies d'archivage d'information, le tout chaque fois plus performant, les résultats des recherches en ligne étant pertinents et rapides. Les utilisateurs, devenus plus exigeants, iront chercher ailleurs si un site n'offre pas des résultats pertinents.

Le lien entre la présence corporative sur Internet et les résultats de recherche n'est pas difficile à établir : un site corporatif devrait apparaître sur les résultats d'une recherche avec des mots-clefs spécifiques, ce qui n'est pas toujours le cas. Le site doit se démarquer dans les résultats de recherche, doit apparaître parmi les premiers. Pour ce faire, des nouvelles théories et des méthodes ont été développées dans le but d'obtenir un meilleur classement sur les résultats : les méthodes d'optimisation pour les moteurs de recherche (en Anglais : SEO, Search Engine Optimisation).

Des méthodes pour améliorer le classement sont donc apparues, le résultat étant facile à vérifier (classement avant et après l'optimisation). Cependant, la valeur de cet effort pour la corporation n'est pas toujours simple à mesurer, l'impact réel, sauf pour des sites transactionnels, le chiffre d'affaires étant censé augmenter après une amélioration du classement.

Dans le cadre de cet essai, la recherche porte sur la valeur ajoutée pour un scénario dans lequel ce n'est pas possible d'établir une corrélation directe entre l'optimisation du site et son impact sur les objectifs corporatifs, plus précisément, un département de soutien technique en technologies de l'information (TI) et son site de documentation technique en ligne libre d'accès sur Internet. Afin de valider les hypothèses de la recherche, le site de documentation du département de soutien technique TI d'une compagnie basée à Sherbrooke fut choisi.

Le site de documentation doit être le point de référence pour les utilisateurs avant qu'ils ne placent une requête de service, ce qui n'était pas le cas au début de la rédaction de cet essai. Les informations disponibles gratuitement en ligne répondaient déjà à beaucoup de requêtes reçues, ces requêtes avaient un impact non négligeable sur la charge de travail du département de soutien.

L'optimisation du site de documentation vise alors à améliorer le classement des documents techniques sur les moteurs de recherche. Un des buts est de diminuer les requêtes de base libérant ainsi du temps de l'équipe de soutien pour travailler sur des requêtes plus avancées. Le doute est si l'amélioration de l'optimisation du site a un effet réel et mesurable sur la charge de travail et si les résultats s'enlignent avec les objectifs corporatifs, ceci compterait comme un retour sur l'investissement.

Jusqu'à présent, il reste du travail à faire pour compléter l'optimisation du site. Cependant, des résultats sont déjà disponibles, différents de ce qui était prévu au début : le temps alloué à des requêtes de base a diminué mais le temps alloué à des requêtes plus avancées augmenta (requêtes non nécessairement répondues par la documentation). Ces résultats portent à croire que le fait d'avoir un bon accès à la documentation de base permet aux utilisateurs de mieux comprendre et de se servir des produits conçus et fabriqués par la compagnie. Ceci ayant comme résultat des exigences plus élevées de la part des utilisateurs : ce qui répond aux objectifs corporatifs, des utilisateurs plus à l'aise avec les produits et services seront des clients plus fidèles. De plus, pendant cette période, les revenus du département de soutien augmentèrent, il y a eu plus d'ententes de service signées par les clients ainsi que plus d'achats d'heures de soutien.

## Remerciements

Sans l'aide particulière de Claude Cardinal, cet essai serait encore inachevé, chaque fois que ce fut nécessaire, il était disponible pour fournir conseil et guide. Mes directeurs, Martin Dozois et Pierre-Martin Tardif, étaient toujours réceptifs et attentifs à mes courriels et appels. Leur expérience professionnelle me permit d'établir une bonne base théorique, le cours de Martin particulièrement me donna l'inspiration sur le sujet de recherche. Un merci spécial à Lynn Legault, qui m'aida à démarrer, donner forme et écrire une première ébauche de cette recherche.

Un grand merci à mes amis Paulina et Julio, ils passèrent des heures à lire mes ébauches, à prendre du temps avec moi pour me questionner et m'aider à regarder le tout sous un autre angle.

Mon père, Roberto, m'a soutenu à distance, il révisa je ne sais combien de brouillons, me fit part de ses conseils et commentaires, nous avons passé beaucoup de temps en ligne et au téléphone afin de peaufiner le tout, jusqu'à qu'il soit satisfait.

Je ne peux finir ces remerciements sans dédier les derniers mots à mon épouse. Andrea, qui m'a soutenu tout au long de ce processus en s'occupant de tout pendant que je passais des soirées et nuits en classe en étudiant ou en rédigeant. Finalement la patience de nos enfants, compréhensifs quand je leur disais que je ne pouvais pas partager plus avec eux car il fallait que je dédie du temps aux études.

## Table des matières

Sommaire .....	ii
Remerciements.....	iv
Table des matières .....	v
Liste des tableaux.....	vii
Liste des figures.....	viii
Glossaire .....	ix
Liste des sigles, des symboles et des acronymes.....	x
Introduction.....	1
Mise en contexte.....	3
1.1 Problème de recherche et contexte de réalisation .....	4
1.2 Concepts en jeu .....	4
1.3 La question de recherche et l'hypothèse .....	5
Revue de la littérature.....	6
2.1 Les moteurs de recherche .....	7
2.2 Liens commandités vs résultats organiques .....	8
2.3 L'optimisation pour les moteurs de recherche.....	10
2.4 Études scientifiques sur l'impact et retour du référencement.....	11
Problématique.....	13
3.1 Description .....	13
3.1.1 Objectifs et hypothèses .....	14
3.1.2 Le choix des mots-clefs .....	15
3.1.3 L'optimisation des contenus pour les moteurs de recherche.....	15
3.1.4 Limites.....	16
3.1.5 Sujets d'étude non couverts par cet essai.....	16
3.2 Méthodologie proposée .....	17
3.2.1 Type de recherche.....	17

3.2.2	Facteurs clef de succès .....	18
3.2.3	Approche de validation des résultats .....	18
3.3	Mise en œuvre .....	18
	Approche effectuée.....	20
4.1	Type de recherche.....	20
4.2	Approche.....	20
4.2.1	Temps de soutien alloué.....	21
4.2.2	Statistiques de visites .....	21
4.2.3	Classement .....	21
4.2.4	Niveau de requête .....	22
4.3	Public cible .....	22
4.4	Validation .....	23
	Présentation et analyse des résultats .....	24
5.1	Retour sur les hypothèses .....	24
5.2	Résultats obtenus.....	25
5.2.1	Classement du portail selon les mots-clefs de recherche.....	25
5.2.2	Temps de soutien gratuit alloué aux requêtes techniques.....	29
5.2.3	Requêtes de soutien technique gratuit traitées .....	31
5.3	Démonstration de la validité des résultats .....	35
	Conclusion.....	36
	Liste des références .....	40
	Bibliographie.....	43



## Liste des tableaux

Tableau 1 - Classement mensuel du portail par groupe de mots et moyenne globale.....	26
Tableau 2 - Classement trimestriel du portail par groupe de mots et moyenne globale.....	27
Tableau 3 - Heures de soutien gratuit par année .....	29
Tableau 4 - Heures de soutien gratuit par trimestre .....	30
Tableau 5 - Nombre de requêtes de soutien classées par type.....	32

## Liste des figures

Figure 1 - Partage du marché des moteurs de recherche en 2015 [15].....	8
Figure 2 - Classement mensuel du portail.....	27
Figure 3 - Classement trimestriel du portail.....	28
Figure 4 - Heures de soutien technique gratuit par année.....	29
Figure 5 - Heures de soutien gratuit fourni par trimestre .....	31
Figure 6 - Total de requêtes de soutien gratuit traitées par trimestre .....	32
Figure 7 - Nombre de requêtes niveau 1 par trimestre .....	33
Figure 8 - Nombre de requêtes niveau 2 par trimestre .....	33
Figure 9 - Nombre de requêtes niveau 3 par trimestre .....	34
Figure 10 - Nombre de requêtes d'accès au logiciel par trimestre.....	34

## Glossaire

Action Research	Recherche pratique ou recherche appliquée
Hyperlien	Une référence permettant de passer automatiquement d'un document consulté à un document lié
Internet	Le réseau informatique mondial accessible au public, un réseau de réseaux
Spider	Logiciel de suivi de liens Web d'un portail à un autre, ou d'une page à une autre
VoIP	Voice Over IP, voix sur IP, télécommunication sur Internet
Web	Abréviation de l'anglais World Wide Web, permet de consulter, avec un navigateur, des pages accessibles sur Internet
Wiki	Outil de création, modification et illustration collaboratives de contenus Web

## Liste des sigles, des symboles et des acronymes

SLA Service Level Agreement

TIC Technologies de l'Information et Communications

VoIP Voice Over IP

## Introduction

Cet essai cherche à mesurer les avantages de l'optimisation des contenus d'un portail Web ayant pour but d'améliorer son classement sur les moteurs de recherche d'Internet. Peu d'études cherchent à établir les bénéfices de l'optimisation des contenus, quelques-unes réussissent à établir le retour sur l'investissement pour des sites d'achat en ligne, tandis que des documents, études et analyses sur les méthodes d'optimisation et leur performance sont disponibles un peu partout.

Pas toutes les organisations visent à vendre ou à faire de la publicité en ligne, des cas où, l'atteinte des objectifs corporatifs n'est pas mesurée financièrement, un lien entre l'amélioration du classement et ces objectifs n'est pas nécessairement simple à établir.

Déterminer la valeur ajoutée d'une façon générale va au-delà de cet essai, le choix est de cibler un contexte particulier : le portail de documentation en ligne d'un département de soutien technique, libre d'accès sur Internet. Ce département s'occupe des requêtes client ainsi que du portail de documentation technique, portail à être utilisé comme première ressource d'information avant de placer une requête de soutien. La valeur ajoutée d'une optimisation du portail serait une diminution de la quantité de requêtes techniques de base soumises au département, permettant d'allouer plus de temps et ressources aux requêtes de plus haut niveau.

Pour les fins de cet essai, l'entreprise Media5 Corporation fut choisie. Media5 conçoit, développe, fabrique et vend des produits de voix sur IP (VoIP) à travers des distributeurs sur tous les continents. Les différents clients (distributeurs, intégrateurs, compagnies de télécommunication) ont des ententes de service (SLA) pour obtenir du soutien ainsi que du développement sur demande. Trois grandes lignes de produits font partie du portfolio offert par la compagnie :

- M5T : services de développement technologique VoIP;
- Mediatix : équipements VoIP;
- Media5-fone : application VoIP pour téléphones mobiles pour le grand public.

Le département de soutien technique s'occupe des lignes Mediatrix et Media5-fone, est le lien avec les clients, travaille sur la résolution des problèmes, répond à des requêtes techniques, documente les besoins et représente les clients à l'interne. Ce département est conçu pour un soutien de haut niveau. Cependant il y avait un grand flux de demandes de soutien de premier et deuxième niveau car la documentation disponible était désuète, difficile à comprendre et mal indexée : une amélioration du portail de documentation était donc de mise.

Un Wiki est mis en ligne pour remplacer l'ancien portail. La documentation migrée, améliorée et mieux classée, toute la documentation est libre d'accès sur Internet. Les documents sont simplifiés et remplacés par des articles courts et spécifiques, faciles à mettre à jour. La plupart de cette amélioration vise la ligne Mediatrix, une section de questions fréquentes est prévue pour l'application Media5-fone, cette dernière n'a pas encore vu le jour.

Au début de cette étude, le département de soutien avait quatre spécialistes à temps plein et jusqu'à 10 heures par semaine étaient dédiées en moyenne à du soutien gratuit (requêtes pour les lignes Mediatrix et Media5-fone, non couvertes par un SLA). Les changements visaient à couper ce temps de moitié. À titre informatif, une entente de service standard prévoit 10, 20 ou 40 heures de soutien par année.

## Mise en contexte

Depuis environ 20 ans, Internet est devenu accessible à toutes les institutions (avec ou sans but lucratif) ainsi qu'au public en général. Ce réseau global a évolué très rapidement pendant cette période de temps, donnant un accès rapide et direct à beaucoup d'information, de marchés, de magasins en ligne.

Cette évolution a créé un nouveau besoin pour les corporations : être visible sur Internet, avoir un portail, « si ce n'est pas sur le Web alors ça n'existe pas ». Les moteurs de recherche sont pris en compte dès qu'un portail Web commence à être ébauché (Google par exemple). Les internautes doivent être capables de trouver l'institution en question ainsi que ses produits ou services.

Beaucoup de documentation a été écrite sur la façon d'améliorer le référencement, optimiser un portail, adapter et optimiser les contenus en fonction des moteurs de recherche visés, sans compter les guides des meilleures pratiques Web et les guides pour améliorer le classement d'un portail sur les moteurs de recherche.

Reste à savoir quel est le véritable avantage de suivre et d'appliquer toutes ces pratiques, il faut être capable de mesurer les effets sur l'organisation, s'ils s'enlignent sur les objectifs et buts corporatifs. Il y a des cas pour lesquels c'est plus simple d'établir un lien, par exemple un magasin qui vend des produits en ligne, mais il y a d'autres qui ne seront pas simples, comme pour un centre de soutien à la clientèle.

Des méthodes, ou des règles, qui permettent de calculer l'impact des pratiques de référencement sur une organisation et ses objectifs institutionnels seraient très utiles, même de grande valeur. Avant de lancer un projet de refonte et d'amélioration d'un portail, il y aurait au moins un aperçu des résultats possibles, ce qui permettrait d'établir la pertinence ou non du projet.

## **1.1 Problème de recherche et contexte de réalisation**

Tel qu'indiqué, tout dépendamment des objectifs institutionnels de chaque organisation, il y a une grande diversité de scénarios possibles et de contextes. Établir des règles générales pour définir les liens, l'impact et la valeur des pratiques de référencement sur n'importe quelle organisation demanderait une recherche plus approfondie et large que celle visée lors d'un essai de maîtrise.

Un scénario particulier à analyser est celui d'un portail de documentation en ligne et un centre de soutien technique d'une compagnie de développement en TIC (Technologies de l'Information et Communications). Quel serait l'impact sur la charge de travail et le type de requêtes techniques reçues en fonction du classement de la documentation du portail sur les moteurs de recherche Internet, sans oublier la pertinence de ces résultats.

Le type de centre de soutien technique visé fait le lien avec les clients, répond à des requêtes techniques, documente les besoins et représente les clients envers le département de recherche et développement quand il y a besoin de nouvelles fonctionnalités ou corrections à inclure dans les produits. Ce centre était conçu pour un soutien de haut niveau, cependant des demandes de soutien de base arrivaient, car la documentation disponible était mal répertoriée et non facilement accessible sur Internet.

Le portail de documentation technique de ce département est ciblé par un projet de refonte visant à rendre la documentation plus vivante ainsi que plus facile à indexer et référencer par les moteurs de recherche sur Internet. Reste à se demander si c'est vraiment pertinent de s'embarquer sur un projet de refonte, si les demandes de soutien de base diminueraient ou non.

## **1.2 Concepts en jeu**

Des recherches de documentation sur les moteurs de recherche Internet sont documentées et réalisées, un suivi régulier est réalisé de façon à vérifier l'évolution des résultats à travers le temps.



En début de l'essai, il est défini que ce ne serait pas possible d'étudier différentes méthodes d'indexation et d'optimisation de contenus, une seule méthode est choisie et les résultats obtenus sont évalués.

Les indicateurs de nombre de requêtes d'avant et après sont analysés, avec un suivi régulier de façon à vérifier l'évolution de ce changement au fur et à mesure que les index des moteurs de recherche Internet sont mis à jour.

### **1.3 La question de recherche et l'hypothèse**

Quel est l'impact d'un référencement de documentation technique de niveau 1 et 2 sur la charge de travail d'un département de soutien par niveau de requête (1, 2 ou supérieur) :

- Quel est le classement des résultats;
- Quantité de requêtes techniques par trimestre;
- Niveau technique de chaque requête de soutien;
- Temps de soutien gratuit alloué par trimestre.

L'hypothèse étant qu'un bon référencement sur les moteurs de recherche a un impact sur la charge de travail d'un département de soutien technique : les requêtes techniques de base devraient diminuer et celles de plus haut niveau devraient se maintenir voire augmenter.

## Revue de la littérature

Beaucoup de guides pratiques sur l'optimisation pour les moteurs de recherche sont disponibles sur le Web, ainsi que des analyses et études sur les meilleures méthodes d'optimisation, études sur les méthodes et stratégies de référencement et les algorithmes utilisés par les moteurs de recherche pour minimiser le risque d'indexation de références biaisées, des études et analyses sur la valeur et le retour de l'investissement sur l'optimisation versus autres mécanismes de référencement payés (liens payés, publicité ciblée, paiement par clic).

Des études visant le référencement de documentation académique furent écartées, car le domaine d'application et les moteurs de recherche ciblés par ces études sortent du focus de cet essai.

La question qui se pose : est-ce qu'il y a eu des recherches sur l'impact et la valeur de l'optimisation pour les moteurs de recherche sur les objectifs organisationnels non seulement financières? Les études et matériel de référence trouvés mettent en valeur l'importance du référencement et le classement sur les moteurs de recherche Internet, mais manquent de fondements valides.

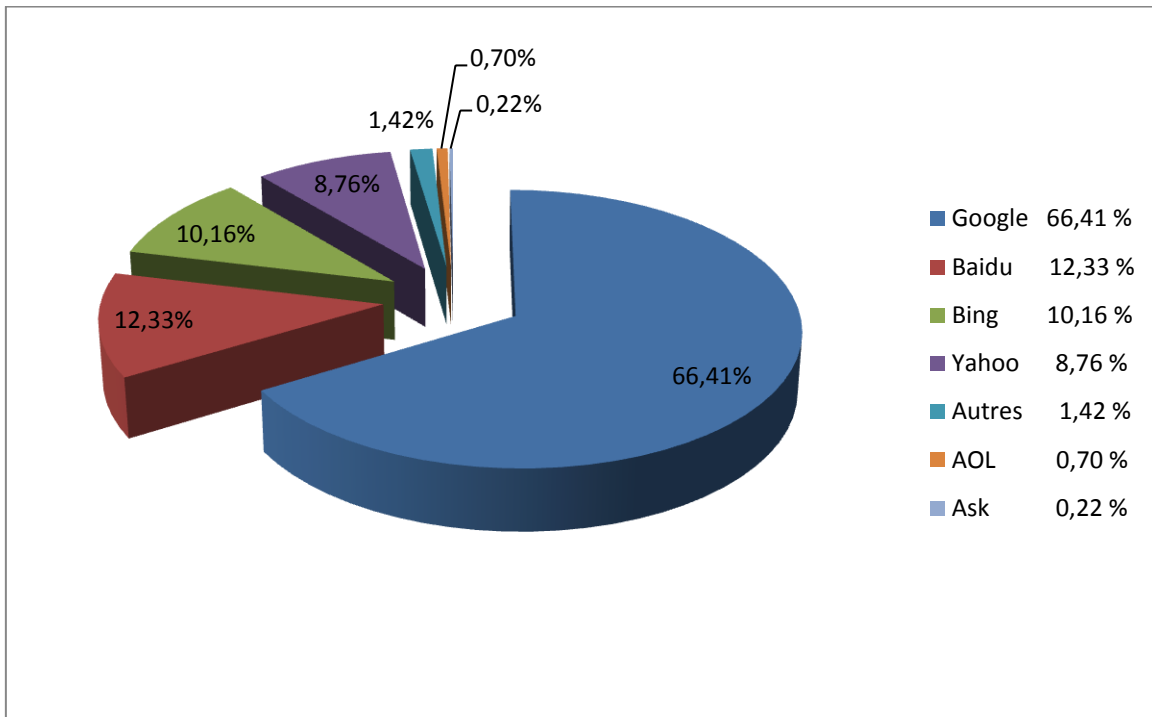
La revue est divisée en quatre sections distinctes. En premier lieu il faut définir ce qui est un moteur de recherche, quel est son utilité, comment ce moteur fonctionne et comment les utilisateurs s'en servent. Les deux types de résultats possibles sur les moteurs de recherche sont expliqués et comparés sur la deuxième partie, mettant en évidence les avantages de chacun. Les deux premières parties servent de base pour la troisième section, qui approfondit sur ce qui est l'optimisation pour les moteurs de recherche, qu'est-ce que ça implique ainsi qu'une vision globale des bonnes pratiques. La dernière section présente un bref sommaire de la seule étude scientifique trouvée jusqu'à présent qui a été réalisée sur un portail transactionnel non fictif, pendant laquelle des actions basées sur la théorie sont

appliquées sur le portail avec un suivi pendant plusieurs mois pour mesurer les résultats et vérifier si les objectifs corporatifs étaient satisfaits ou non.

## **2.1 Les moteurs de recherche**

Moteur de recherche : tout simplement une base de données de pages Web (un index), une méthode pour trouver des pages Web et les indexer et une façon de rechercher dans la base de données [1]. Les moteurs de recherche fonctionnent en deux étapes : premièrement des logiciels appelés « spiders » (araignées) trouvent et suivent des hyperliens, puis, en une deuxième étape les contenus sont indexés et mis à jour au besoin [10].

Si on parle de moteur de recherche, il faut penser aux résultats fournis par ce moteur en fonction des mots-clefs utilisés, les résultats sont classés selon la qualité de l'information ainsi que leur pertinence en fonction des algorithmes de classement utilisés par le moteur. Plusieurs facteurs rentrent en jeu : autorité, crédibilité, longueur du texte, contenu audiovisuel, bonne rédaction [3]. Le tiers ou plus du trafic obtenu par la plupart des sites Internet est déclenché par les moteurs de recherche [3].



**Figure 1 - Partage du marché des moteurs de recherche en 2015 [15]**

Faits saillants au sujet des moteurs de recherche et les résultats présentés sur ceux-ci :

- Plus de 60 % des utilisateurs vérifient les résultats de la première page de résultats [2];
- Moins de 10 % des utilisateurs vérifient les résultats au-delà de la troisième page [2];
- Pour être visible, il faut viser la première page de résultats, au maximum la deuxième.

Les moteurs de recherche les plus populaires aujourd'hui sont Google, Baidu, Bing et Yahoo (**Figure 1**). Google ayant deux tiers du marché mondial à lui seul devient le point de référence sur une bonne partie des études, documents et guides pratiques. Baidu est le deuxième moteur de préférence avec un 12,33 % du marché en 2015, un site qui vise quasi exclusivement les internautes en Chine pour des recherches en mandarin [15].

## 2.2 Liens commandités vs résultats organiques

Aujourd'hui il y a une telle quantité de sites et d'information disponibles sur Internet que c'est rendu très difficile, presque impossible, de se débrouiller sans les moteurs de recherche. Plus

de la moitié du trafic sur Internet commence via ces derniers [9] et les internautes s'en servent régulièrement.

Qui dit visibilité, pense moteur de recherche, deux types de résultats seront obtenus lors d'une recherche :

1. Liens commandités : lorsque des mots-clefs spécifiques sont recherchés, un ou plusieurs liens vers le site du commanditaire apparaîtront avant les résultats de recherche du moteur;
2. Liens organiques : les résultats de recherche sur la base de données du moteur pour des mots-clefs spécifiques souvent différents de ceux fournis par les résultats commandités.

L'avantage des liens commandités est le fait que sans modifier les contenus, le site apparaîtra en premier sur les résultats de recherche du moteur [4]. Cependant le coût de cet investissement dépend des mots-clefs à utiliser, il y a des mots plus convoités que d'autres, donc les frais peuvent être plus élevés. Il faut envisager aussi qu'il faut payer chaque fois que le lien commandité est suivi par un utilisateur, il y a des cas pour lesquels les frais seront plus élevés que les profits, cette option peut être utile en un premier temps pour commencer à avoir de la visibilité [4].

D'autre part, un travail sur les contenus d'un portail dans le but de le rendre visible demande un investissement initial qui peut avoir un retour, si le travail et la méthode d'optimisation sont bien faits et établis puis suivis d'un bon entretien du site Web. Avec des contenus à jour et indexés, les résultats de classement se maintiendront dans le temps et pourraient même améliorer.

Au sujet des résultats organiques :

- Attirent plus les internautes et sont perçus comme des résultats objectifs, non biaisés [7];
- 70 % des internautes choisissent les liens organiques au lieu des liens commandités [14];
- Les internautes sont plus enclins à acheter des produits ou services présentés sur les résultats organiques [9].

## 2.3 L'optimisation pour les moteurs de recherche

Optimisation pour les moteurs de recherche : le processus d'identification des facteurs dans un site Web qui auraient un impact sur son accessibilité par les moteurs de recherche, ainsi que l'ajustement des différents éléments du portail de façon à obtenir la visibilité plus élevée quand un moteur de recherche répond à une recherche reliée au sujet du site [5].

Pour réussir dans le processus d'optimisation d'un site et ainsi obtenir un meilleur classement sur les résultats de recherche des moteurs, il faut comprendre comment ceux-ci fonctionnent, quels sont les facteurs-clefs de succès, les meilleures pratiques et quelles sont les démarches à suivre ou pas.

Le portail de documentation technique étudié pour cet essai est en anglais, les guides, documents et matériel de référence étudiés analysent les deux moteurs de recherche plus importants dans cette langue, Google et Yahoo, la plupart ayant une emphase sur Google.

La première démarche avant de commencer une refonte d'un site ou mise à jour des contenus est de définir les objectifs à atteindre. Pour ce faire, il faut :

- Définir les buts corporatifs auxquels le site doit répondre [12];
- Comment l'atteinte des buts corporatifs sera mesurée [12];
- Quelles sont les balises de référence pour définir si les buts sont atteints ou non [4];
- Qui sont les utilisateurs visés par le site [4];
- Établir le budget pour le projet [12].

Points saillants pour obtenir un bon classement :

- Le titre et description pour chaque page sont clairs pour les utilisateurs ainsi que pour les moteurs de recherche [8];
- Choisir des mots-clefs que les internautes vont utiliser en premier lieu et qui permettront de positionner le site sur les résultats de recherche [11];
- Les mots-clefs choisis pour le site et les descriptions donnent un aperçu clair du contenu, ils sont extraits du texte principal [6];
- Il y a une corrélation entre le nom de domaine et les mots-clefs du contenu du site [9];
- La navigation doit être simple, du contenu plus général vers le plus spécifique [13];

- Utiliser du texte pour la navigation (hyperliens) au lieu d'images [13];
- Les contenus doivent être clairs, organisés, conçus pour des utilisateurs et non pas pour les moteurs de recherche [13].

## **2.4 Études scientifiques sur l'impact et retour du référencement**

Jusqu'à présent, une seule étude scientifique visant l'impact du référencement sur les objectifs corporatifs est trouvée. Un projet de recherche du professeur Ross A. Malaga [1] analysant l'effet de l'optimisation pour les moteurs de recherche sur un site de commerce électronique, la méthodologie décrite, les résultats obtenus ainsi que le suivi avant et après l'implantation s'enlignent avec l'objectif visé par cet essai.

Toutes les autres études scientifiques trouvées et analysées se basaient sur des recherches théoriques sur des sites Web fictifs (mis en ligne sur Internet lors de la durée de la recherche puis retirés), créés dans le seul but de valider ou réfuter des hypothèses ainsi qu'établir des algorithmes mathématiques qui permettraient de calculer l'effectivité d'une méthode de référencement par rapport à une autre.

La valeur particulière du projet de recherche du professeur Malaga est le fait que ce n'est pas seulement une recherche théorique, l'étude utilise comme laboratoire d'essai un nouveau site de vente en ligne. Les buts corporatifs à atteindre définis au point 2.3 sont établis et par la suite un plan d'action est mis en place. Des indicateurs sont définis et au fur et à mesure que l'étude avance, des actions basées sur la théorie sont appliquées en fonction des résultats obtenus à chaque étape, des mesures sont prises avant et après ces actions.

La méthodologie appliquée est ce que Malaga nomme en anglais « action research », ce qui pourrait être traduit librement par une recherche pratique ou recherche appliquée. Celle-ci étant une recherche réalisée en collaboration avec un particulier dans un contexte éthique bien défini. L'avantage est que les deux parties impliquées peuvent bénéficier de ce type de recherche, le particulier profite des connaissances des chercheurs et ceux-ci peuvent valider des théories et hypothèses sur un scénario réel. Il faut cependant bien établir les limites et

responsabilités ainsi que clarifier les attentes de chaque partie impliquée avant même de commencer.

Le but du travail est d'obtenir un bon classement sur les moteurs de recherche pour des recherches explicitement liées au domaine d'affaires du site, ce domaine étant très compétitif. Pour ce faire, il faut se démarquer de la concurrence en utilisant quasiment les mêmes mots-clés que les autres, ce qui fait vraiment la différence est la méthodologie et les bonnes pratiques de l'optimisation. Tout au long de l'étude, trois grandes actions sont réalisées sur le site, ses contenus et leur structure ainsi que des actions sur des sites externes pour « pousser » les moteurs de recherche à ajouter le portail sur leur index. Au bout de deux mois, le site passe d'une quasi inexistence sur les moteurs jusqu'au deuxième rang sur MSN (actuellement Bing) et troisième sur Yahoo, au bout de six mois c'était le premier. Cependant il n'apparaissait pas encore sur Google. Ceci s'explique par le fait que les algorithmes d'indexation de Google attendent que les sites aient au moins six mois d'existence.

Cette approche et méthodologie sont semblables à celles utilisées pour cet essai. Les buts organisationnels sont différents mais une analogie peut être faite.



## Problématique

### 3.1 Description

Tel qu'exposé au premier chapitre, cette étude vise un centre de soutien technique de haut niveau. Théoriquement, ce centre doit répondre aux besoins de clients ayant une entente de service. Cependant beaucoup de temps de soutien gratuit était alloué à des individus ou petites entreprises n'ayant pas et qui n'achèteraient pas un contrat de service.

Le soutien gratuit fait partie des stratégies d'affaires de l'entreprise. C'est une façon de faire connaître les produits ainsi que de gagner une part du marché. La solution en place n'était pas optimale, il y avait un impact sur le soutien alloué aux clients détenteurs de contrats de service, ainsi que sur l'équipe de soutien qui avait des ressources limitées.

Le groupe de soutien cherche alors d'autres approches pour diminuer la quantité de requêtes de soutien gratuit sans laisser de côté les demandeurs. Le plan est de donner un accès public et libre sur les bases de connaissance. Cette façon de faire devait avoir un impact direct sur la charge de travail du groupe.

Dans ce contexte, un Wiki est mis en place, la décision est prise sur le fait que c'est basé sur code source ouvert gratuit, bien documenté, avec beaucoup de ressources de connaissances et expertises disponibles sur Internet. De plus, c'est un outil connu par les utilisateurs qui seraient capables de s'y repérer. Le dernier argument pour ce choix, non moins important, est qu'au sein de l'équipe de soutien, il y a l'expertise et connaissances pour mettre en place et entretenir cet outil.

Une fois le choix fait, les membres de l'équipe de soutien ont pour objectif de rédiger et ajouter régulièrement des bulletins techniques sur le site, le tout en collaboration avec le département de rédaction technique. Les bulletins et documents ajoutés sont directement liés aux questions récurrentes reçues de la part des clients et des utilisateurs (en plus des

documents pour des nouveaux produits et scénarios). Une section de questions et réponses fréquentes est aussi prévue.

La question qui se pose est si un travail de référencement du Wiki pour les moteurs de recherche était de mise, est-ce que l'effort aurait un retour mesurable? Intuitivement un bon classement sur les moteurs de recherche devrait avoir un impact sur la quantité de requêtes récurrentes qui seraient répondues sur le Wiki, cette quantité devrait diminuer. Comment faire pour valider cette hypothèse?

Pour obtenir l'appui de la direction de l'entreprise, des données sont recueillies sur le système de gestion de soutien et requêtes techniques, des indicateurs sur la quantité de temps de soutien gratuit fourni depuis 2011 sont obtenus. Après la présentation des tableaux des données recueillies, la direction approuve les démarches proposées pour améliorer le référencement du portail de documentation technique.

### **3.1.1 Objectifs et hypothèses**

La question de recherche est donc :

« Quel est l'impact d'un bon référencement Internet du portail de documentation technique sur la charge de travail du groupe de soutien? »

L'objectif de la recherche consiste à répondre à cette question en se basant sur des faits issus d'une expérimentation scientifique.

Hypothèses :

- Un bon référencement sur les moteurs de recherche a un impact sur la charge de travail d'un département de soutien technique, les requêtes techniques devraient diminuer;
- La documentation est adéquate et à jour, bien indexée et optimisée pour les moteurs de recherche;
- La documentation est bien classée et organisée sur le site, avec une structure claire qui s'ajuste adéquatement aux moteurs de recherche.

Le nouveau portail de documentation est en place depuis le deuxième trimestre de 2013, ce qui donne de crédibilité au site du point de vue des moteurs de recherche. En fait, le site apparaissait déjà sur les classements, mais au-delà de la troisième page de résultats.

### **3.1.2 Le choix des mots-clefs**

Le portail doit non seulement être indexé par les moteurs de recherche mais il faut que les résultats aient un lien avec les contenus fournis (et recherchés par les utilisateurs). Pour ce faire, il faut savoir quels seraient les mots-clefs de recherche utilisés plus fréquemment, ces mots n'étant pas nécessairement ceux qui sont déjà inclus sur la documentation. Ces mots-clefs sont déjà disponibles mais il faut les recueillir, classer et organiser:

- Tous les courriels et requêtes de soutien reçus par le groupe de soutien sont analysés (plus de 500 par semaine);
- Un bilan des termes et thèmes récurrents est établi;
- Une liste des mots-clefs descriptifs des scénarios et besoins plus fréquents est bâtie;
- Par la suite, la liste est régulièrement révisée pour s'assurer qu'elle soit à jour.

### **3.1.3 L'optimisation des contenus pour les moteurs de recherche**

La méthode d'optimisation choisie est simple : intégrer les mots-clefs sur le portail, les articles ainsi que les documents publiés. Ces mots-clefs doivent permettre de retrouver ces documents rapidement lors d'une recherche sur Internet.

L'intégration des mots-clefs :

- Sur le titre des articles et documents publiés;
- Sur la description et sommaire de chaque article et document;
- Sur le nom des fichiers des documents à télécharger et
- Sur le nom des images affichées sur les articles et documents.

Plus il y a de cohérence entre les titres, sommaires et contenus, ainsi que sur le nom des fichiers mieux le classement sera. Si les images ont des noms clairs et bien décrits, celles-ci apparaîtront sur les recherches d'images pour le sujet, ce qui pourra déclencher une visite au site de documentation.

### **3.1.4 Limites**

Cet essai vise à établir un lien entre le classement sur les moteurs de recherche et la charge de travail du département de soutien technique. Pour ce faire, il faut des données mesurables :

- Classement sur les moteurs de recherche;
- Pertinence des résultats sur les moteurs de recherche;
- Niveau technique de la requête de soutien;
- Temps dédié aux requêtes techniques.

Un seul moteur de recherche est choisi pour l'analyse de classement : Google. Ce moteur avait un peu plus du 67 % du marché en 2014, les autres moteurs populaires sont Baidu, Bing et Yahoo. Baidu est un portail de recherche utilisé majoritairement en Chine, tandis que Bing et Yahoo avaient moins du 13 % du marché des moteurs de recherche en 2014 [15].

Lors de l'analyse préliminaire des données de temps de soutien, il est établi que seules les requêtes s'adressant aux produits Mediatix sont prises en compte. Le portail de documentation n'est pas prêt pour adresser les besoins de soutien de la ligne Media5-fone.

Lors des premières phases de l'essai, il est établi que ce n'est pas possible d'essayer différentes méthodes d'indexation et d'optimisation de contenus. Ce choix dû à des problèmes sur la base de données du Wiki déjà mis en place, le seul référencement possible étant d'intégrer les bons mots-clefs sur les titres des articles publiés.

### **3.1.5 Sujets d'étude non couverts par cet essai**

Un projet qui impactait la documentation est mis en place lors du déroulement de cet essai, celui-ci n'étant pas partie de ce dernier, il est mentionné juste à titre informatif :

- Amélioration de la documentation technique :
  - Processus simplifié et automatisé de génération et mise à jour de documents et
  - Mise à jour des bulletins techniques et notes de configuration.

Les objectifs visés ci-dessous par ce projet ainsi que celui de la refonte du portail en cours portent sur des sujets d'étude qui ne seront pas abordés dans l'essai. Bien que ces objectifs et résultats ont un impact direct sur le sujet de recherche.

Automatisation de création de contenus à partir des notes sur le code logiciel :

- Quelles seraient les méthodes disponibles et plus adéquates pour inclure les notes dans le code et quels seraient les meilleurs outils pour les extraire?
- Comment convertir les notes en documentation technique?

Gestion documentaire interne et externe :

- Amélioration de la gestion documentaire;
- Les outils et méthodes de gestion documentaire.

Mise en place d'un Wiki :

- Les outils de développement web disponibles : rechercher les plus adéquats pour un portail de documentation technique;
- Pourquoi choisir un Wiki au lieu d'un forum ou autre outil.

Convivialité du portail de documentation :

- Analyse des différentes méthodes d'amélioration de convivialité;
- Études de convivialité avec des groupes d'utilisateurs;
- Enquête auprès des clients : méthodes et approche.

## **3.2 Méthodologie proposée**

### **3.2.1 Type de recherche**

La recherche est de type quantitatif. Des statistiques sont prises, documentées, comparées et mises en relation :

- Classement des résultats sur les moteurs de recherche par mois et par trimestre;
- Nombre de requêtes de soutien reçues par trimestre :
  - Classement des requêtes par type;

- Temps de soutien gratuit alloué aux requêtes techniques par trimestre.

### **3.2.2 Facteurs clef de succès**

Pour bien établir les liens entre les différentes statistiques et valeurs recueillies, la qualité des données est critique. Les membres du département de soutien doivent bien documenter les requêtes sur lesquelles ils ont travaillé (type de client, temps alloué et type de requête) et cette qualité doit perdurer dans le temps.

La documentation est en processus de mise à jour, intuitivement ça devrait avoir un impact sur la charge de travail mais cet impact ne peut pas être mesuré car, en même temps, le portail du Wiki ainsi que le référencement sont mis à jour.

L'auditoire de cette recherche sont les utilisateurs des produits supportés par le centre de soutien, ils viennent demander de l'aide ou lire la documentation disponible: les outils de mesure permettraient de recueillir l'information nécessaire automatiquement.

### **3.2.3 Approche de validation des résultats**

La méthodologie de cueillette de données et d'analyse de résultats expliquée sur le prochain chapitre s'ajuste à des recherches semblables déjà réalisées antérieurement mais à d'autres contextes (voir section 2.4). Cette méthodologie donne un degré de confiance et certitude sur la validité des résultats à obtenir.

Ce n'est pas un échantillon de participants qui est étudié, l'essai tient compte de toutes les requêtes reçues de la part du public ciblé (documentées sur le système de suivi) minimisant ainsi le marge d'erreur des résultats.

## **3.3 Mise en œuvre**

Tout au long de la période d'échantillonnage des mises à jour et optimisations selon les bonnes pratiques furent appliqués méthodiquement sur le site :

- De nouveaux documents adressant des scénarios non couverts préalablement sont publiés;

- Les titres des pages, sommaires et contenus existants révisés plusieurs fois par mois de façon à mieux intégrer les mots-clefs recueillis;
- La structuration et classement des documents révisés, optimisés et organisés, afin de simplifier la navigation sur le site et trouver les contenus pertinents plus facilement;
- Les fichiers existants sont renommés en intégrant les mots-clefs descriptifs du contenu aussi;
- Les noms de fichier des images publiées sont en cours de mise à jour, de la même façon que les fichiers de documentation, le nom étant un indicateur de ce qui est représenté, ceci est de grande utilité lors de la recherche d'images de diagrammes ou scénarios sur les moteurs de recherche.

Tout ce processus est encore en place et fait maintenant partie des tâches du groupe de soutien.

Les indicateurs sont mesurés de façon régulière, les données sont recueillies pendant 19 mois moment où l'analyse des résultats est déclenchée.

## **Approche effectuée**

### **4.1 Type de recherche**

La méthodologie de recherche est quantitative corrélationnelle c'est-à-dire qu'il devrait avoir un lien entre le temps de soutien alloué aux clients non détenteurs de contrat de services, le nombre de requêtes techniques classé par niveau et le classement sur les moteurs de recherche.

### **4.2 Approche**

Dans le cadre de cette recherche, le nombre d'heures de soutien est extrait hebdomadairement, si jamais il y a besoin de plus de détail, étant donné la variation d'une semaine à une autre le choix est de rassembler les données par trimestre, ce qui permet de mieux observer les tendances et variations.

Le classement sur Google est documenté et pris mensuellement à partir de juillet 2014. Par la suite, les résultats sont rassemblés par trimestre de la même façon que pour les heures de soutien. Ceci dans le but de mieux croiser et analyser les données obtenues. La moyenne des classements obtenus est utilisée pour le tableau de résultats et plusieurs combinaisons de mots de recherche sont utilisées chaque fois.

L'analyse préliminaire des données ainsi que l'évolution du Wiki, démontre qu'il est difficile d'adresser les requêtes concernant la ligne Media5-fone, puisque celles-ci ont besoin d'être répondues directement sous le champ d'aide de l'application ainsi qu'un changement d'approche sur la façon de répondre aux questions reliées. La cible est alors la ligne de produits Mediatrix seulement.



### **4.2.1 Temps de soutien alloué**

Un système de suivi de requêtes de soutien existe et les outils d'extraction de données disponibles permettent d'extraire le temps de service fourni par :

- Client
- Ticket de service
- Jour

Un ticket de service générique est utilisé pour documenter le soutien gratuit fourni aux clients non détenteurs de contrat de service. Cette façon de faire permet d'obtenir le temps alloué à ce type de soutien pour une période spécifique.

### **4.2.2 Statistiques de visites**

Un outil d'analyse des visites ne peut être installé sur le site, dû à des problèmes de base de données système. L'objectif est d'obtenir des statistiques des visites telles que :

- La localisation géographique du visiteur : continent, pays
- La source : moteur de recherche, lien, redirection, visite directe
- Les mots-clefs utilisés sur le moteur de recherche (si la visite venait d'un moteur de recherche)
- La date et l'heure

### **4.2.3 Classement**

Le classement sur Google est suivi de façon manuelle. Pour ce faire, des termes récurrents sur les requêtes de soutien sont choisis pour effectuer des recherches et vérifier les résultats. Le but est d'apparaître parmi les 20 premiers résultats. Google affichant dix résultats par page, une 30<sup>ème</sup> place équivaut à la dernière place sur la troisième page de résultats.

Les combinaisons de mots de recherche sont choisies en fonction des questions sur les requêtes reçues par le département de soutien, celles-ci étant les plus récurrentes :

- Mediatix upgrade;
- Mediatix software download;

- Mediatrix firmware download;
- Mediatrix documentation;
- Mediatrix setup et
- Mediatrix configuration.

À noter que le portail de documentation n'apparaît pas sur les trois premières pages de résultats lors de recherches pour la combinaison « Mediatrix manual » ou seulement « Mediatrix », même plus loin (31<sup>ème</sup> ou plus). Le portail corporatif apparaît pourtant parmi les premiers (1<sup>er</sup>, 2<sup>ème</sup> et même 3<sup>ème</sup>), ce qui indique qu'il reste du travail à faire sur l'indexation du Wiki.

#### **4.2.4 Niveau de requête**

La quasi-totalité des requêtes de soutien technique sont gérées par courriel, les réponses fournies sont dûment documentées et enregistrées sur le système de suivi.

Les requêtes sont classées selon les trois niveaux standard de soutien TI, plus une quatrième type de requête :

- Niveau 1 – Soutien de base;
- Niveau 2 – Configuration VoIP et réseau de base;
- Niveau 3 – Configuration et débogage avancés;
- Demandes d'accès au logiciel.

Les requêtes niveau 1 et 2 sont adressées avec la documentation déjà publiée sur le portail.

Un tableau des requêtes reçues et leur niveau technique est bâti à partir des archives courriels du centre de soutien.

### **4.3 Public cible**

Tous les utilisateurs qui demandent du soutien pour les applications logicielles ou les équipements de la ligne de produits Mediatrix, ainsi que les entreprises qui déploient les solutions et techniciens qui déploient ou soutiennent les solutions.

Les informations sont déjà disponibles sur le système de suivi de requêtes, il n'y avait pas besoin d'autorisation des utilisateurs pour utiliser ces données, puisque celles-ci sont corporatives.

#### **4.4 Validation**

Le lien proposé entre les variables semble clair et intuitivement logique mais il reste à valider et démontrer par les résultats et le suivi (avant, pendant et après). Une analyse des données d'avant la mise en place du projet de référencement indique la présence du lien entre les variables. Il faut remarquer que le classement du portail de soutien au début de l'étude était non mesurable (pas de classement).

La quasi-totalité des requêtes est documentée, au moins tout ce qui passe par le système de suivi, ce qui minimise la marge d'erreur des résultats.

La cueillette des données débute au mois de juillet 2014 et continue jusqu'à janvier 2016 sur une période de 19 mois. Un premier aperçu des résultats démontre de grands progrès sur le classement. Toutefois, le nombre de requêtes de soutien ne diminue pas, par contre le chiffre augmente, ceci va à l'encontre des hypothèses de départ.

## Présentation et analyse des résultats

Une fois que la collecte de données est complétée, l'analyse doit être amorcée. Pour ce faire, il faut classer le tout en tableaux et graphiques représentatifs. À première vue les résultats sont contraires à ce qui était espéré c'est-à-dire que le nombre de requêtes de soutien a augmenté au lieu de diminuer, de même que pour les heures de soutien gratuit. Un retour sur la façon de classer et analyser les données est nécessaire.

### 5.1 Retour sur les hypothèses

La question de recherche au début est :

« Quel est l'impact d'un bon référencement Internet du portail de documentation technique sur la charge de travail du groupe de soutien? »

La recherche et collecte de données est basée sur l'hypothèse que le nombre de requêtes techniques de soutien gratuit doit diminuer, ce qui n'est pas le cas. L'erreur de cette hypothèse est le fait que toutes les requêtes sont visées, sans catégorisation, il faut donc les organiser d'une autre façon:

- Les requêtes sont déjà extraites et classées par date;
- Une nouvelle approche est réalisée sur les requêtes, celles-ci sont analysées plus en détail;
- L'analyse démontre que le type de requête n'est pas toujours le même, quatre groupes sont identifiés :
  - Niveau 1 – Soutien de base;
  - Niveau 2 – Configuration VoIP et réseau de base;
  - Niveau 3 – Configuration et débogage avancés;
  - Demande d'accès au logiciel.

Les requêtes sont alors classées à nouveau par date et type/niveau. Les hypothèses sont révisées en tenant en compte ces nouveaux éléments :

- Un bon référencement sur les moteurs de recherche a un impact sur la charge de travail d'un département de soutien technique : les requêtes techniques de base doivent diminuer et celles de plus haut niveau doivent se maintenir, ou même augmenter;
- La documentation est adéquate et à jour, bien indexée et optimisée pour les moteurs de recherche;
- La documentation est bien classée et organisée sur le site, avec une structure claire qui s'ajuste adéquatement aux moteurs de recherche.

## **5.2 Résultats obtenus**

Cette section présente les résultats de classement du portail, par la suite les statistiques des requêtes de soutien, classées par type, finalement le temps de soutien gratuit fourni dans le temps.

### **5.2.1 Classement du portail selon les mots-clefs de recherche**

Six combinaisons de mots-clefs sont identifiées. Le classement du portail est obtenu mensuellement pour chaque groupe de mots de façon à mieux représenter le classement global du portail sur un seul graphique. La moyenne est calculée, le but est d'avoir une vision simplifiée des tendances.

Les groupes de mots identifiés sont (numérotés de 1 à 6 sur le tableau 1) :

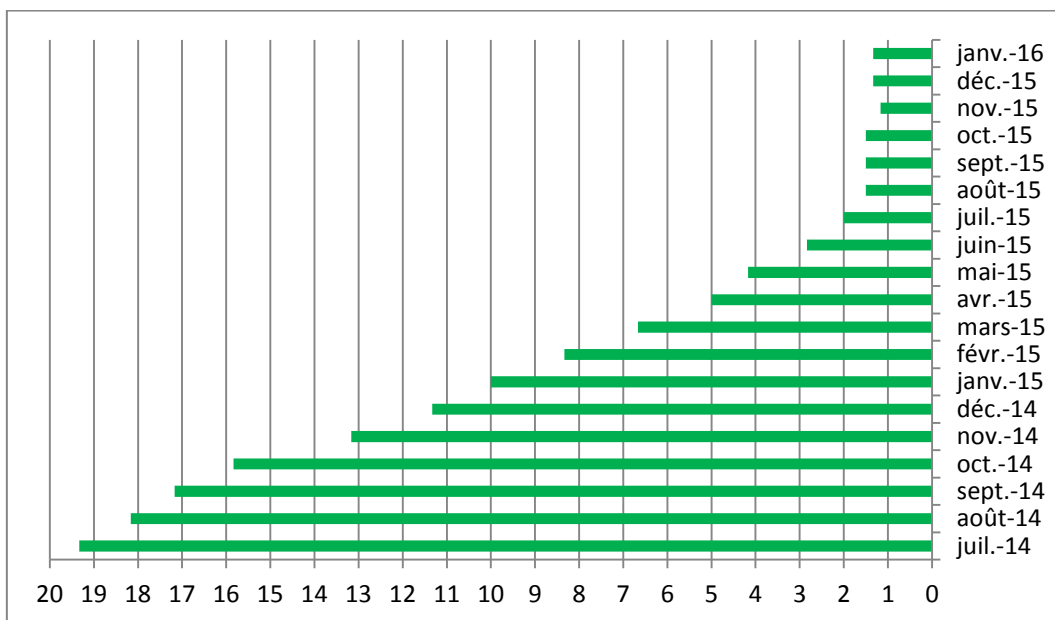
1. Mediatix upgrade;
2. Mediatix software download;
3. Mediatix firmware download;
4. Mediatix documentation;
5. Mediatix setup;
6. Mediatix configuration.

Le classement est la place du portail sur la liste de résultats, cette place est documentée seulement si les liens affichés sur les résultats de recherche sont pertinents, ce qui est mesuré est la qualité en plus de la visibilité.

Mois	Mots 1	Mots 2	Mots 3	Mots 4	Mots 5	Mots 6	Moyenne
juil-14	17	14	23	21	20	21	19,33
août-14	17	14	20	19	20	19	18,17
sept-14	16	14	21	17	19	16	17,17
oct-14	16	10	17	17	19	16	15,83
nov-14	13	10	12	15	16	13	13,17
déc-14	13	7	9	11	16	12	11,33
janv-15	10	7	9	10	12	12	10,00
févr-15	7	4	8	10	11	10	8,33
mars-15	7	4	5	7	9	8	6,67
avr-15	4	3	5	7	6	5	5,00
mai-15	4	2	2	6	6	5	4,17
juin-15	2	2	2	4	4	3	2,83
juil-15	2	1	2	4	2	1	2,00
août-15	1	1	1	3	2	1	1,50
sept-15	1	1	1	3	2	1	1,50
oct-15	1	1	1	3	2	1	1,50
nov-15	1	1	1	2	1	1	1,17
déc-15	2	1	1	2	1	1	1,33
janv-16	2	1	1	2	1	1	1,33

**Tableau 1 - Classement mensuel du portail par groupe de mots et moyenne globale**

Le classement mensuel pour chaque groupe de mots ainsi que la moyenne globale qui représente le classement global du site de documentation sont présentés sur le Tableau 1.



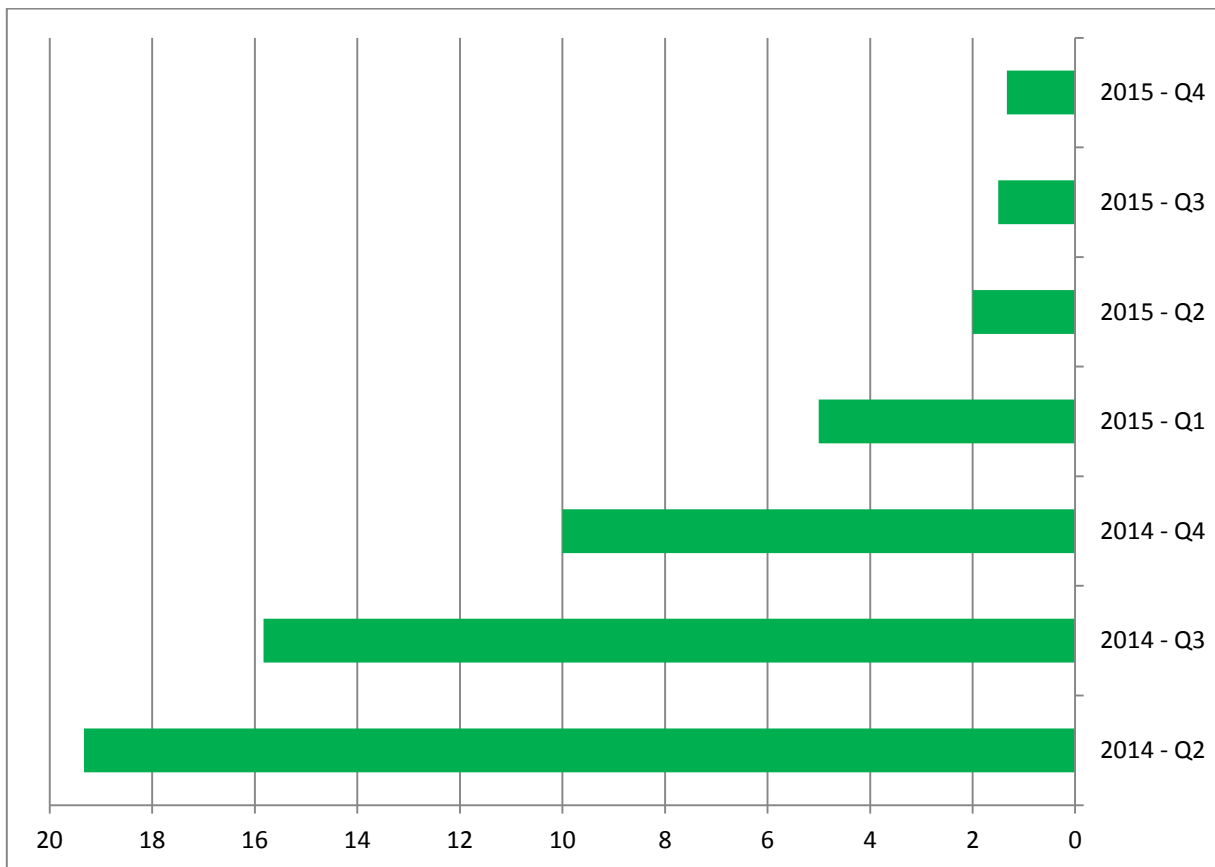
**Figure 2 - Classement mensuel du portail**

L'évolution de la moyenne globale mensuelle calculée au Tableau 1 est représentée visuellement sur la Figure 2.

Quart	Mots 1	Mots 2	Mots 3	Mots 4	Mots 5	Mots 6	Moyenne
2014 - Q2	17	14	23	21	20	21	19,33
2014 - Q3	16	10	17	17	19	16	15,83
2014 - Q4	10	7	9	10	12	12	10,00
2015 - Q1	4	3	5	7	6	5	5,00
2015 - Q2	2	1	2	4	2	1	2,00
2015 - Q3	1	1	1	3	2	1	1,50
2015 - Q4	2	1	1	2	1	1	1,33

**Tableau 2 - Classement trimestriel du portail par groupe de mots et moyenne globale**

Le classement pour chaque groupe de mots ainsi que la moyenne globale qui représente le classement global du site de documentation sont regroupés par trimestre sur le Tableau 2.



**Figure 3 - Classement trimestriel du portail**

L'évolution de la moyenne globale trimestrielle calculée au Tableau 2 est représentée visuellement sur la Figure 3.

Les tableaux et graphiques montrent clairement l'évolution de la visibilité du portail au fil du temps, les progrès ne sont pas les mêmes pour chaque combinaison de mots. Globalement, les ajustements réalisés pour l'optimisation ont eu un impact clair et l'effet n'est pas le même pour chaque composante.

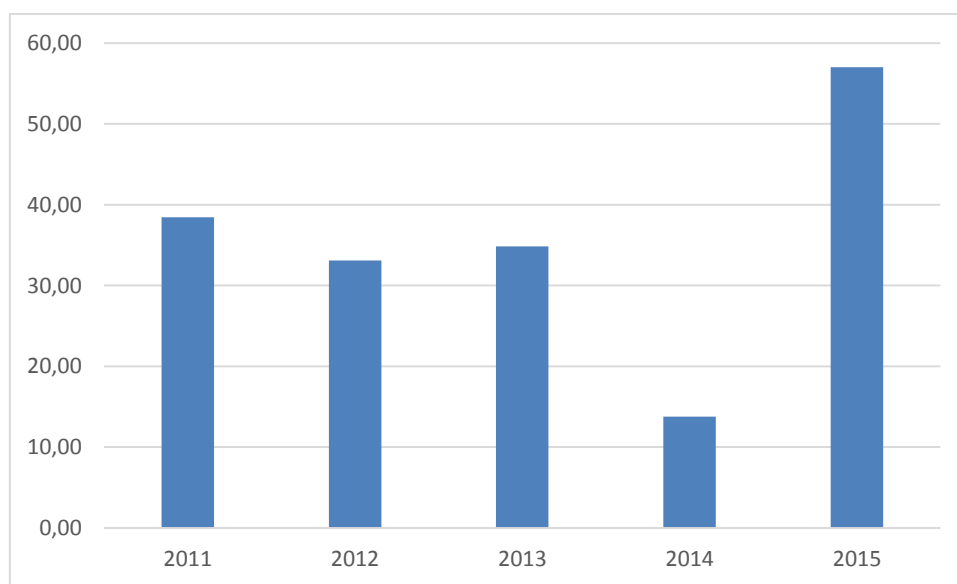
Le classement du portail démontre une nette amélioration sur une période de sept mois (juillet 2014 à janvier 2015), au début trois groupes de mots sont sur la troisième page des résultats (classement 21 ou plus). Rendus au mois de janvier 2015 sur les six groupes de mots deux seulement sont demeurés sur la deuxième page, les quatre autres sont affichés sur la première page.



Au début de janvier 2016, deux groupes sont affichés en deuxième place, tous les autres sont à la première place. Le Wiki s'affiche plusieurs fois sur les résultats de recherche pour les mots-clefs en pointant vers plus d'un document pertinent aux résultats. Le groupe de mots numéro 1 est descendu dans le classement pendant les deux derniers mois, de la première place à la deuxième alors qu'il était à la première place les quatre mois précédents.

### 5.2.2 Temps de soutien gratuit alloué aux requêtes techniques

Un tableau avec les heures de soutien gratuit est créé et entretenu bien avant le déclenchement du projet d'optimisation.



**Figure 4 - Heures de soutien technique gratuit par année**

Au mois de janvier 2016 le nombre total d'heures de soutien gratuit fourni depuis janvier 2011 jusqu'à la fin décembre 2015 est extrait puis représenté par année (Tableau 3 et Figure 4).

Année	Heures
2011	38,45
2012	33,08
2013	34,83
2014	13,75
2015	57,00

**Tableau 3 - Heures de soutien gratuit par année**

Un premier constat est la différence des valeurs sur 2014 et 2015 par rapport aux années précédentes :

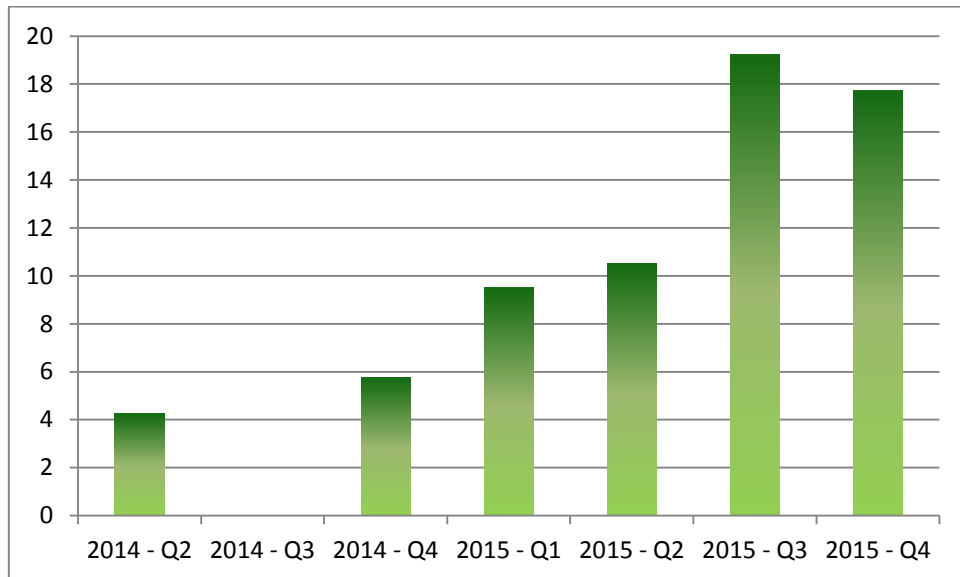
- Pendant la période 2011 à 2013 les valeurs restent semblables, stables entre 34 et 39 heures de soutien gratuit par année;
- Pour 2014 la valeur diminue à presque un tiers par rapport à 2013, passant de 34,83 heures à 13,75 heures;
- L'année 2015 compte presque 40 heures de soutien gratuit seulement pour les trois premiers trimestres. Le total dépasse de presque 50 % la valeur comptabilisée pour 2011 (Tableau 3).

Quart	Heures
2014 - Q2	4,25
2014 - Q3	0,00
2014 - Q4	5,75
2015 - Q1	9,50
2015 - Q2	10,50
2015 - Q3	19,25
2015 - Q4	17,75

**Tableau 4 - Heures de soutien gratuit par trimestre**

Les heures de soutien gratuit depuis le début du projet d'optimisation jusqu'à la fin de prise de données sont classées par trimestre (Tableau 4).

La première remarque est que pendant le troisième trimestre de 2014, il n'y a pas eu d'heures de soutien gratuit comptabilisées (2014 –Q3). Cependant l'analyse des requêtes de soutien gratuit répondues indique clairement qu'il y a eu du temps alloué à celles-ci. Une seule conclusion est possible : le temps n'est pas documenté adéquatement sur le système.



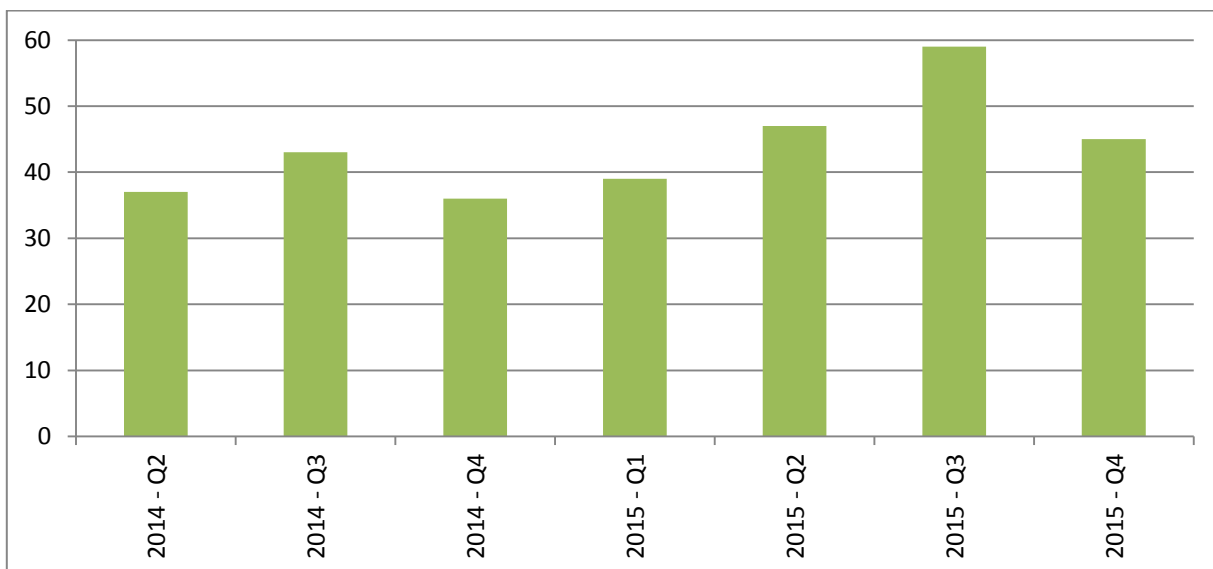
**Figure 5 - Heures de soutien gratuit fourni par trimestre**

La qualité des données extraites du système de suivi des tickets de service est en doute. Cependant les données pour les autres périodes semblent bien documentées et présentent un bon degré de cohérence d'une période à une autre. D'autres facteurs doivent être pris en compte alors pour expliquer le manque d'information. Au cours de l'année 2014, il y a eu un changement de système de suivi des requêtes, pendant une période le temps était documenté sur le vieux système et le nouveau simultanément, en plus de la période d'apprentissage pour se servir des nouveaux outils. Donc oui, la qualité des données fut impactée pendant une période, par la suite tout fut remis en place.

Le doute éclairci, il y a quand même un constat clair. En 2015, il y a eu beaucoup plus de temps de soutien gratuit fourni pendant les neuf premiers mois que pour chaque année précédente depuis 2011, et le chiffre final est encore plus élevé puisque le temps de soutien gratuit a augmenté au lieu de diminuer comme espéré.

### **5.2.3 Requêtes de soutien technique gratuit traitées**

Les dernières données à vérifier sont celles des requêtes techniques traitées. Les requêtes de soutien reçues sont automatiquement archivées par le système. Étant donné qu'il n'y a pas de gestion manuelle de ces requêtes, le risque d'erreur et de perte de qualité sur ces données est faible.



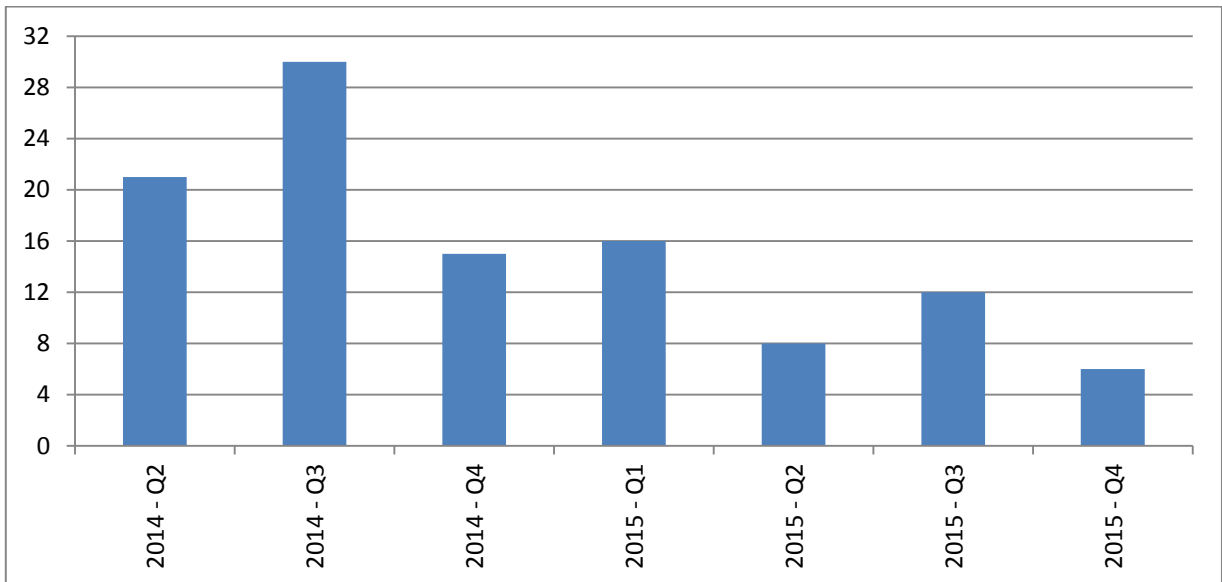
**Figure 6 - Total de requêtes de soutien gratuit traitées par trimestre**

Une augmentation du nombre de requêtes traitées par trimestre est remarquée (Figure 6). À nouveau, ça va à l'encontre de ce qui est espéré, cette tendance est cohérente avec l'augmentation de temps de soutien gratuit fourni (Figure 5).

Pour essayer de comprendre cette augmentation, les requêtes sont classées par type et cette façon de faire pourra apporter de la lumière sur les causes et fournir une explication.

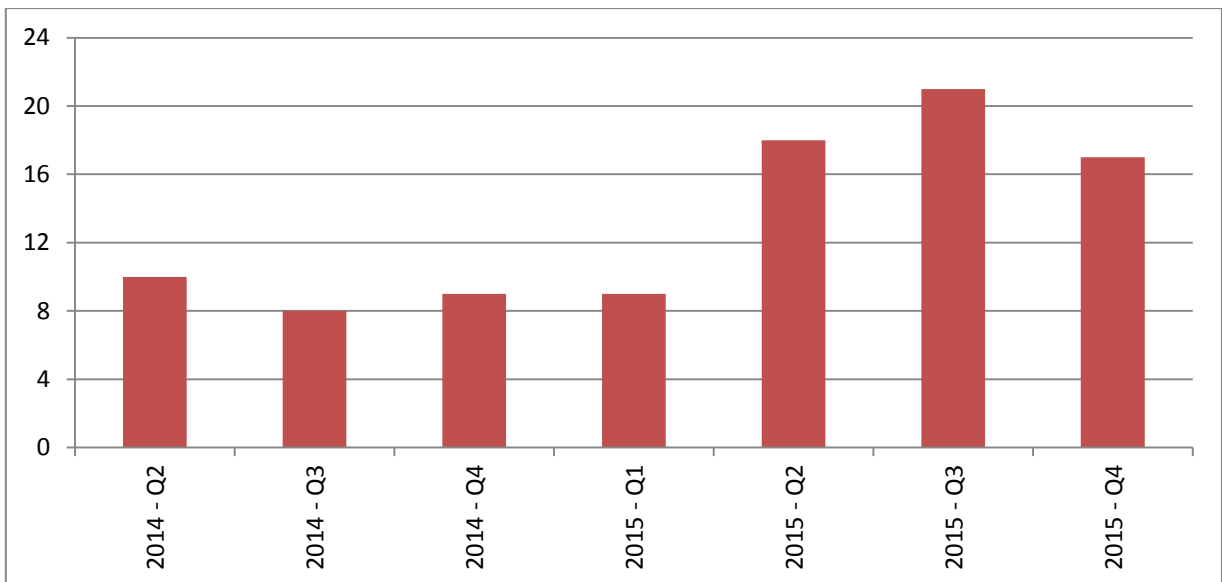
Trimestre	Accès logiciel	Niveau 1	Niveau 2	Niveau 3	Total
2014 - Q2	5	21	10	1	37
2014 - Q3	3	30	8	2	43
2014 - Q4	8	15	9	4	36
2015 - Q1	7	16	9	7	39
2015 - Q2	9	8	18	12	47
2015 - Q3	10	12	21	16	59
2015 - Q4	8	6	17	14	45

**Tableau 5 - Nombre de requêtes de soutien classées par type**



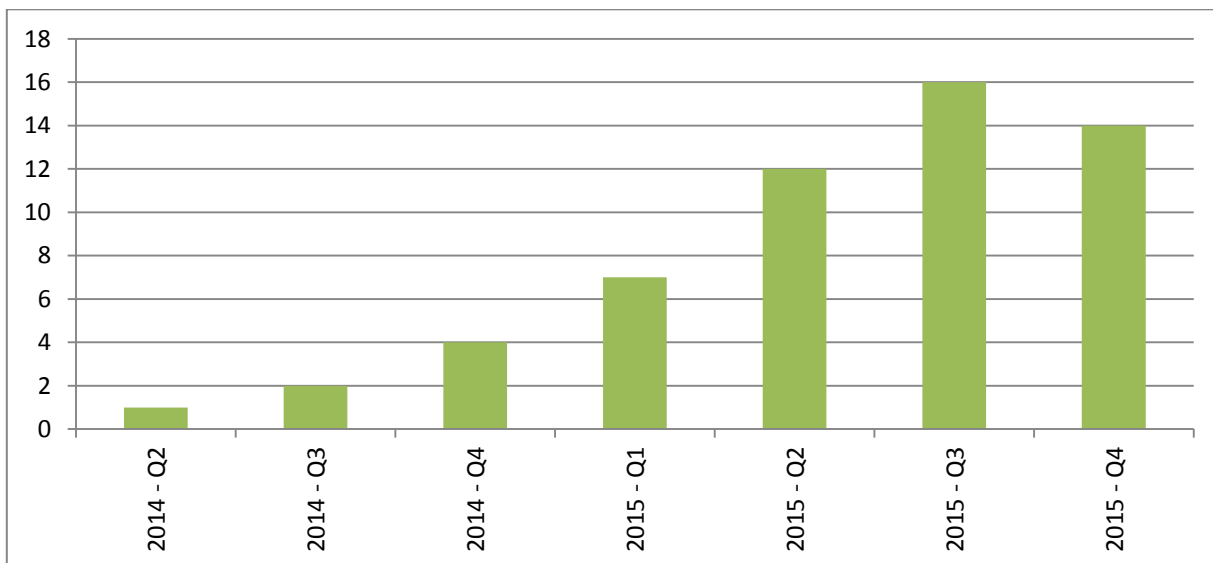
**Figure 7 - Nombre de requêtes niveau 1 par trimestre**

Le fait de classer par type permet de mieux visualiser les variations et comportements. En premier lieu le nombre de requêtes de niveau 1 a diminué à moins de la moitié en un an (Q4 2014 vs Q4 2015) (Tableau 5). L'hypothèse que les requêtes de soutien de base pourraient diminuer est validée.



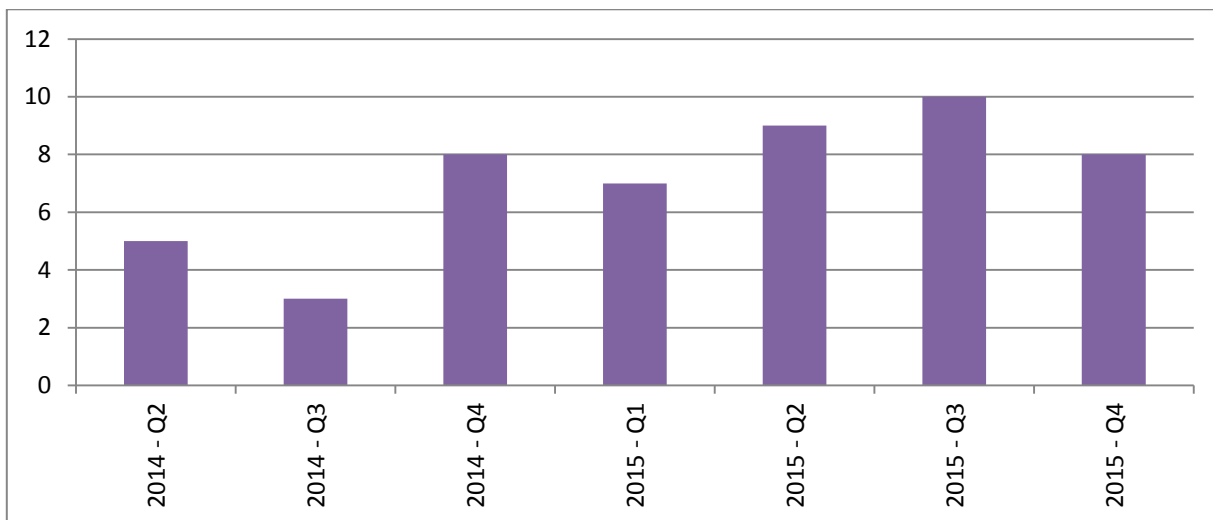
**Figure 8 - Nombre de requêtes niveau 2 par trimestre**

D'autre part les requêtes de niveau 2 diminuent un peu pendant trois trimestres puis le chiffre augmente significativement. À la fin de la collecte de données, le nombre a augmenté de 70 % par rapport au chiffre initial (Q2 2014).



**Figure 9 - Nombre de requêtes niveau 3 par trimestre**

L'augmentation de requêtes de soutien plus remarquable est pour celles de niveau 3. La tendance est quasi exponentielle, passant d'une seule jusqu'à 16 pour le troisième trimestre 2015, puis 14 pour le dernier trimestre (Tableau 5).



**Figure 10 - Nombre de requêtes d'accès au logiciel par trimestre**

Finalement, les demandes d'accès au logiciel on aussi doublé par rapport au début de l'étude, après une légère diminution lors du troisième trimestre 2014.

### **5.3 Démonstration de la validité des résultats**

La première question à se poser lors d'une recherche basée sur l'expérimentation pratique, dans un environnement non contrôlé, comme dans le cas présent, c'est si les données recueillies sont crédibles, non biaisées, valides et avec une marge d'erreur raisonnable. La façon d'assurer des données valides est d'utiliser des outils de mesure de qualité fiables, ainsi que des méthodes prouvées :

- La méthode d'obtention des valeurs du classement du portail sur Google est déjà prouvée et validée par le professeur Malaga;
- Le temps de soutien gratuit alloué est documenté sur un système de suivi informatique, des erreurs pourraient être introduites lors de la saisie des valeurs par les spécialistes de soutien;
- Les requêtes de soutien sont saisies automatiquement par les systèmes de suivi. Les seules qui pourraient échapper à ce système seraient celles envoyées directement aux spécialistes via courriel ou appel téléphonique. Celles-ci proviennent de la part des clients détenteurs de contrat de service.

Les tableaux obtenus démontrent des valeurs cohérentes, la seule erreur constatée est la saisie des heures de soutien pour le troisième trimestre de 2014. Celle-ci expliquée par le fait du changement de système de suivi et des données sont malheureusement perdues. À tenir en compte que les spécialistes veillent particulièrement à bien saisir leurs temps car le système de paye est connecté à celui de soutien avec des éléments de contrôle.

Malgré le problème de documentation des heures il y a des bons indicateurs qui prouvent la validité de toutes les autres valeurs saisies. L'erreur introduite est compensée par la qualité des autres données ce qui rend les résultats de cette étude pertinents.

## Conclusion

Cet essai vise à établir un lien entre la charge de travail d'un département de soutien technique et l'amélioration du classement de son portail de documentation technique sur les moteurs de recherche sur Internet, notamment Google pour le cas présent.

Pour réussir à ces fins, une démarche méthodique a été établie :

- Définir et encadrer le problème de recherche;
- Établir la question de recherche c'est à dire ce qu'il fallait répondre à la fin de l'étude;
- Rechercher, trouver et étudier la documentation pertinente :
  - Comprendre qu'est-ce qu'un moteur de recherche, son fonctionnement;
  - Comprendre les résultats de recherche obtenus sur les moteurs;
  - Comprendre les facteurs qui mènent à un bon classement sur les résultats;
  - Étudier les bonnes pratiques pour réussir des bons classements;
  - Étudier la méthodologie de recherche pour des cas d'études semblables au cas présent;

Une fois les enjeux identifiés et ayant une base de connaissances fiable comme référence, la problématique de l'étude est analysée en profondeur, la question de recherche, les objectifs et hypothèses définis, ainsi que les enjeux et limites :

- La question : « Quel est l'impact d'un bon référencement Internet du portail de documentation technique sur la charge de travail du groupe de soutien? »;
- L'objectif est de répondre à cette question en se basant sur des faits issus d'une expérimentation crédible;
- Les hypothèses :
  - Un bon référencement sur les moteurs de recherche a un impact sur la charge de travail d'un département de soutien technique les requêtes techniques devraient donc diminuer;
  - La documentation est adéquate et à jour, bien indexée et optimisée pour les moteurs de recherche;



- La documentation est bien classée et organisée sur le site, avec une structure claire qui s'ajuste adéquatement aux moteurs de recherche.

Une fois que la question de recherche et les hypothèses sont établies la documentation de référence utilisée apporte une base théorique crédible, ainsi qu'une méthodologie de recherche c'est à dire la quantitative corrélacionnelle. Par la suite l'approche, les données à saisir, les unités de mesure à utiliser, l'échantillon de recherche (public cible), puis un plan d'action sont définis. Restait à faire la collecte de données et l'analyse.

Une première analyse des données collectées a remis en question deux hypothèses de base, sur lesquelles l'étude se basait, mettant en question la validité des résultats obtenus :

- Un bon référencement sur les moteurs de recherche a un impact sur la charge de travail d'un département de soutien technique et les requêtes techniques devraient diminuer :
  - Le nombre de requêtes n'a pas diminué. Il y a eu au contraire une augmentation significative du nombre de requêtes, même chose pour le temps de soutien gratuit;
- Les données de temps de soutien documentées sur le système de suivi des requêtes est fiable et bien saisi :
  - Il n'y a pas d'heures de soutien gratuit documentées sur le système pour le troisième trimestre de 2014, par contre les données sur le nombre de requêtes de soutien gratuit traitées pendant cette période indique clairement qu'il y a eu du temps alloué, celui-ci n'est pas saisi adéquatement.

L'objectif d'une recherche est de soit valider ou de réfuter une hypothèse. Si le nombre de requêtes techniques a augmenté c'est qu'il doit y avoir une explication.

Une vérification en profondeur démontre que le problème de qualité de données est encadré dans un seul trimestre. L'explication est qu'il y a eu un changement de système de suivi lorsque l'étude fut réalisée et la transition de systèmes s'est étalée pendant plusieurs semaines lors du troisième trimestre de 2014. Il y a eu une période d'apprentissage et ajustement aux nouveaux outils qui a eu un impact sur la saisie des données pendant ce temps. Les données saisies pour les autres périodes étaient cohérentes avec les résultats, sauf pour la période en question le reste demeure valide.

Les données sont ensuite analysées plus en profondeur, avec des résultats intéressants :

- Le nombre total de requêtes de soutien augmente, ainsi que le temps de soutien gratuit :
  - Le nombre de requêtes de soutien de niveau 1 diminue effectivement;
  - Le nombre de requêtes de soutien de niveau 2 double;
  - Le nombre de requêtes d'accès au logiciel double aussi;
  - Le nombre de requêtes de soutien de niveau 3 augmente de façon quasi exponentielle.
- Le classement du portail de soutien s'est amélioré significativement c'est à dire de la troisième page des résultats de recherche à la première place.

Malheureusement, les outils de mesure de visite du portail ne peuvent être installés, seules des hypothèses peuvent être formulées en se basant sur les résultats :

- Une documentation plus visible et accessible permet aux utilisateurs de trouver une réponse aux questions fréquentes, sans avoir besoin de contacter le groupe de soutien;
- Le fait d'avoir accès à la documentation et répondre aux questions de base déclenche rapidement des questions plus avancées et les utilisateurs sont capables de déployer les produits et s'en servir.
- Les demandes d'accès au logiciel vient de la meilleure connaissance des produits, les versions plus récentes fournissent des améliorations et nouvelles fonctionnalités.

Une meilleure connaissance et maîtrise des produits permet d'avoir des clients plus fidèles aux solutions fournies, prêts à les utiliser à nouveau : ce n'est pas le nombre de ventes qui compte, mais la fidélité des clients.

Donc, oui, un bon classement du portail de documentation technique sur les moteurs de recherche a un impact sur la charge de travail du groupe de soutien : les requêtes de base diminuent mais celles de plus haut niveau augmentent.

Pendant cette période, on constate qu'il y a eu une augmentation du nombre de contrats de service vendus. Des clients qui n'avaient pas renouvelé depuis plusieurs années ainsi que des nouveaux clients, il y a aussi eu une augmentation d'achat de blocs d'heures de service pour des besoins ponctuels. Bref, les revenus de service du groupe de soutien ont augmenté.

L'impact sur la charge de travail allouée aux détenteurs de contrats de service, le niveau technique des requêtes de soutien soumises ainsi que le nombre de requêtes par type ne sont pas mesurés par cet essai. Tout porte à croire que ces clients sont en train de se servir du Wiki.

Un projet de remplacement du portail de documentation est actuellement en marche, le nouveau site devrait être mis en ligne au courant de 2016. Ce changement vise à répondre à plusieurs besoins que le site actuel n'est pas capable de répondre :

- Automatisation de la publication de la documentation technique;
- Automatisation de la mise à jour des contenus de la documentation technique;
- Automatisation de l'inclusion des mots-clefs et contenus pour les moteurs de recherche;
- Intégration d'outils de mesure de visites et navigation (statistiques);
- Format mobile (pour les appareils intelligents).

Des outils de mesure des visites peuvent fournir des informations telles que :

- Source de la visite : moteur de recherche, lien à partir d'un autre site, lien direct;
- Mots-clefs de recherche si la source est un moteur de recherche;
- Autres documents vérifiés ou téléchargés.

Pour les visites en provenance d'un moteur de recherche il est possible de vérifier si les mots-clefs mis en valeur sur les documents sont les plus adéquats, si les documents vérifiés ont un rapport avec les mots-clefs de la recherche. Ces données permettent de mieux ajuster les titres, sommaires et contenus des documents de façon à fournir des résultats pertinents à la recherche.

Le site actuel a sa base de données endommagée, ceci empêche d'activer une version mobile et d'installer des outils de mesure de visites. Avec les changements à venir un meilleur lien sera établi entre les visites, la documentation et la charge de travail du groupe de soutien.

## Liste des références

- [1] Malaga, Ross A., «The Value of Search Engine Optimization: An Action Research Project at a New E-Commerce Site», *Journal of Electronic Commerce in Organizations (JECO)*, Volume 5, Issue 3, 2007
  
- [2] Malaga, Ross A., «Worst practices in search engine optimization», *Communications of the ACM*, Volume 51, N°12, Décembre 2008
  
- [3] Fichter, D. et Wisniewski, J., «Being Findable: Search Engine Optimization for Library Websites», *Online Searcher*, Volume 38, Issue 5, Septembre/Octobre 2014 , Pages 74-76
  
- [4] Arsenault, T. et Feeny, D., «Your Website: Greater Visibility and Return of Investment through Search Engine Marketing», *Journal of Library and Information Technology*, Volume 28, N°4, Juillet 2008
  
- [5] Zhang, J. et Dimitroff, A., «The impact of webpage content characteristics on webpage visibility in search engine results (part-I)», *Information Processing and Management*, Volume 41, Issue 3, Mai 2005, Pages 665-690
  
- [6] Zhang, J. et Dimitroff, A., «The impact of metadata implementation on webpage visibility in search engine results (part-II)», *Information Processing and Management*, Volume 41, Issue 3, Mai 2005, Pages 691-715

- [7] Xing, B. et Lin, Z., «The impact of search engine optimization on online advertising market», *Conference Paper, ACM International Conference Proceeding Series*, 2006, Pages 519-529
- [8] Kumar, L. et Kumar, N., «SEO Techniques for a Website and its Effectiveness in Context of Google Search Engine», *International Journal of Computer Sciences and Engineering*, Volume 2, Issue 4, Avril 2014, Pages 113-118
- [9] Shih, B.-Y., Chen, C.-Y. et Chen, Z.-S., «An empirical study of an internet marketing strategy for search engine optimization», *Human Factors and Ergonomics In Manufacturing*, Volume 23, Issue 6, November 2013, Pages 528-540
- [10] Jain, Ayush, «The Role and Importance of Search Engine and Search Engine Optimization», *International Journal of Emerging Trends & Technology in Computer Science*, Volume 2, Issue 3, Mai/Juin 2013, Pages 99-102
- [11] Olivier Andrieu, «Référencement : optimisation d'un site web», *Ed. Techniques de l'Ingénieur*, Février 2005, 18 p.
- [12] Alliance Numérique, «Guide des meilleures pratiques Web», [En ligne]. <http://alliancenumerique.com/fr/publications/guide-des-meilleures-pratiques-web/>, V1, Octobre 2009, 24 p.
- [13] Google, «Search Engine Optimization - Starter Guide», [En ligne]. <http://googlewebmastercentral.blogspot.ch/2010/09/seo-starter-guide-updated.html>, 2010, 32 p.

[14] «State of Search Marketing Report 2012», SEMPO, <http://www.sempo.org>

[15] «Desktop Search Engine Market Share», NetMarketShare,  
<https://www.netmarketshare.com>

[16] «Internet Live Stats», <http://www.internetlivestats.com>

## Bibliographie

Ledford, Jerri L., «Search Engine Optimization Bible», *Wiley Publishing Inc.*, Second Edition, 2008, 507 p.

Grappone, J. et Couzin, G., «Search Engine Optimization: An Hour a Day», *Wiley Publishing Inc.*, Second Edition, 2008

Jones, Khristopher B., «Search Engine Optimization: Your visual blueprint for effective Internet marketing», *Wiley Publishing Inc.*, 2008, 288 p.

Blanco, R., Bortnikov, E., Junqueira, F., Lempel, R., Telloli, L. et Zaragoza, H., «Caching search engine results over incremental indices», *SIGIR '10 Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, 2010, Pages 82-89

Haas, K., Mika, P., Tarjan, P. et Blanco, R., «Enhanced results for web search», *SIGIR '11 Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, 2011, Pages 725-734

Killoran, John B., «How to Use Search Engine Optimization Techniques to Increase Website Visibility», *IEEE Transactions on Professional Communication*, Volume 56, Issue 1, February 2013, Pages 50-66

Castillo, C. et Davison, B.D., «Adversarial Web Search», *Journal: Foundations and*

*Trends in Information Retrieval*, Volume 4, Issue 5, Mai 2011, Pages 377-486

Özlük, Ö. et Cholette, S., «Allocating expenditures across keywords in search advertising», *Journal of Revenue and Pricing Management*, Volume 6, Number 4, Décembre 2007, Pages 347-356

Moreno, L. et Martinez, P., «Overlapping factors in search engine optimization and web accessibility», *Online Information Review*, Volume 37, Issue 4, 2013, Pages 564-580

Cholette, S., Özlük, Ö. et Parlar, M., «Optimal Keyword Bids in Search-Based Advertising with Stochastic Advertisement Positions», *Journal of Optimization Theory and Applications*, Volume 152, Issue 1, January 2012, Pages 225-244

Shi, J., Cao, Y. et Zhao, X., «Research on SEO strategies of university journal websites», *The 2nd International Conference on Information Science and Engineering*, Décembre 2010, Pages 3060-3063

Marszałkowski, J., Marszałkowski, J. M. et Drozdowski M., «Empirical study of load time factor in search engine ranking», *Journal of Web Engineering*, Volume 13, Issue 1&2, Mars 2014, Pages 114-128

Suchánek, Petr, «The fundamentals of prosperous e-shop in connection to search engine optimization», *Economics and Management*, 2010, Pages 92-103

Sau-ling Lai, Linda, «In Search of Excellence – Google vs Baidu», *World Academy of Science, Engineering and Technology*, Volume 60, 2011, Pages 1108-1111



