HAL
open science

Analyse des activités en ligne et circulation de l'information : faciliter la prise de décision des entreprises afin d'orienter leurs stratégies numériques

Guillaume Reynier

To cite this version:

Guillaume Reynier. Analyse des activités en ligne et circulation de l'information : faciliter la prise de décision des entreprises afin d'orienter leurs stratégies numériques. Sciences de l'information et de la communication. 2016. dumas-01778761

HAL Id: dumas-01778761

https://dumas.ccsd.cnrs.fr/dumas-01778761

Submitted on 26 Apr 2018

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

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

Mémoire de fin d'études – Université Lille 3, UFR DECCID

Stage réalisé sein de l'agence altima° du 18/04/2016 au 31/09/2016

Analyse des activités en ligne et circulation de l'information : Faciliter la prise de décision des entreprises afin d'orienter leurs stratégies numériques.

Mémoire constitué pour l'obtention du diplôme :

Master 2 Ingénierie Documentaire, Edition et Médiation Multimédia (IDEMM)

Par :

Guillaume Reynier

Soutenu le :

Mercredi 14 septembre 2016

Sous la direction de :

Mme Marie Després-Lonnet, tuteur universitaire

Mme Julia Ferhat, tuteur professionnel

Année universitaire 2015-2016

REMERCIEMENTS

L'élaboration de ce mémoire a été grandement facilitée grâce aux conseils et au soutien de plusieurs personnes à qui je voudrais témoigner toute ma reconnaissance.

Je tiens tout d'abord à remercier altima° et les différents collaborateurs de cette agence. Un grand merci à l'équipe WAO regroupant les expertises d'analyse numérique, de conversion et d'acquisition de trafic avec laquelle j'ai passé l'ensemble de mon temps et avec qui j'ai pu apprendre et échanger dans une ambiance conviviale.

Des remerciements plus particuliers vont vers Mme Julia Ferhat, tuteur de stage professionnel qui a su me faire confiance et qui m'a très rapidement donné la possibilité de travailler avec des clients importants tout au long du stage et M. Antoine Tissier pour ses conseils avisés. Ces deux personnes n'ont pas hésité à prendre sur leur temps pour me transmettre le savoir nécessaire permettant de mener à bien l'ensemble des missions qui m'ont été confiées.

Merci à Mme Després-Lonnet, directrice de mémoire pour les échanges constructifs dans le cadre de la rédaction du mémoire, permettant d'orienter la réflexion et dans le cadre des différents cours dispensés.

Je tiens enfin à remercier mon amie m'ayant soutenu tout au long de mes études et ma famille m'ayant permis de les réaliser avec succès.

Table des matières

Présentation de L'agence altima° 6
Introduction 7
I. Récolte et fiabilité des données quantitatives 10
I.1 Les données quantitatives : histoire, définition et précisions 11
1.a Histoire de l'analyse numérique 11
1.b Définitions et caractéristiques des types de données quantitatives 13
1.c Biais des données quantitatives et facteurs de fiabilité 17
I.2 Élaboration d'un plan de tracking 19
2.a Définition 19
2.b Méthodologie 20
2.c Regard critique sur les difficultés de cette activité 22
I.3 Automatisation de l'extraction des données 24
3.a Objectifs et avantages 24
3.b Stratégie adoptée 25
3.c Perspectives 27
Conclusion de partie I 29
II. Analyses Qualitatives de l'écosystème client 30
II.1 Traitement et utilisation des données brutes 31
1.a Méthodologie de type bottom-up 31
1.b Méthodologie de type top-down 34
II.2 Analyse Ergonomique 37
2.a Analyses par un dérivé de test utilisateur: Hotjar 37
2.b Apports théoriques 41
II.3 Intelligence économique 43
3.a Benchmarking concurrentiel 43
3.b Veille multiforme 45
II.4 Analyse Technique 47
4.a Justifications 47
4.b Démarche de type audit 48
4.c Démarche exploratoire 51
Conclusion de partie II 53
III. Capitalisation et transmission du savoir 54
III.1 Transmission des connaissances : une vulgarisation rigoureuse 55

1.a Supports de communication utilisés 55
1.b Vulgarisation des analyses 56
1.c Structurer les connaissances transmises 59
III.2 Capitalisation sur les connaissances 61
2.a Les types de savoirs constituant les connaissances 61
2.b Processus de capitalisation 63
III.3 Gestion du temps liée au travail d'analyse 65
3.a Définition des différentes activités 65
3.b Optimisation du temps des différentes activités 67
3.c Procédure d'optimisation du temps 69
Conclusion de partie III 70
Conclusion & ouverture 72
Bibliographie 74
Webographie 75

PRESENTATION DE L'AGENCE ALTIMA°

altima° est une agence digitale pionnière dans le domaine e-commerce depuis 1997. Cette agence possède toutes les compétences pour accompagner tous ses clients lors des différentes étapes d'un projet de développement lié à une problématique e-commerce. Cet accompagnement peut s'effectuer soit à l'échelle nationale, soit à l'échelle internationale. En effet, cette agence possède aujourd'hui 8 bureaux répartis dans 5 pays et 3 continents.

Depuis 1997, date de fondation de l'agence, sa croissance ne cesse de croître. Aujourd'hui, l'agence compte 270 collaborateurs et pas moins de 250 clients de tous types comme Petit Bateau, CORA, Verbaudet. Le positionnement sur le marché est singulier, car cette agence est essentiellement reconnue pour son approche R.O.I.ste. C'est-à-dire que les clients qui investissent dans un service, doivent généralement profiter d'un bénéfice bien supérieur à l'investissement de base.

Parmi les grandes familles d'expertises, nous pouvons retrouver par exemple :

- La Gestion de projet : pour permettre aux différents clients de réaliser leurs sites e-commerce dans les meilleures conditions.
- Optimisation et analyse web : études sur les sites web déjà existants permettant de définir des axes à optimiser pour obtenir un bénéfice plus important.
- Acquisition : Développement de stratégies permettant d'obtenir un trafic qualifié sur les sites web grâce à l'utilisation de régies publicitaires payantes.

Réaliser un stage au sein de cette agence est une chance qui va permettre d'obtenir rapidement les connaissances nécessaires à la rédaction de ce mémoire sur l'analyse des activités en ligne.

INTRODUCTION

La rédaction de ce mémoire de fin de cursus de master Ingénierie Documentaire Edition et Médiation Multimédia (IDEMM) s'est effectué dans le cadre du stage réalisé au sein de l'agence altima° située à Roubaix. L'objectif principal de ce document est d'appréhender et de comprendre précisément le parcours de l'information permettant de comprendre les processus d'analyses des activités en lignes afin d'optimiser la stratégie de présence en ligne des entreprises. L'étude va donc permettre de suivre ce parcours en partant de la donnée la plus simple jusqu'à la constitution du savoir nécessaire à la mise en place des différentes actions.

Face à la croissance globale du web et des activités commerciales qui y sont liées, beaucoup de métiers, de pratiques et de stratégies se créent et évoluent. Le *digital analyst*[1] est un métier clé lié à cette croissance car cette personne est sensée comprendre de manière claire les enjeux et les besoins des entreprises afin de participer au pilotage des activités en ligne au travers de ses analyses.

Ce métier est relativement récent et c'est au travers de cette l'activité que nous allons étudier le parcours de l'information dans l'activité d'analyse des activités en ligne comme les sites web ou encore les applications mobiles. Ce professionnel est théoriquement chargé d'analyser les données quantitatives. Il est aussi capable d'étudier précisément la navigation des visiteurs pour évaluer la performance du site de manière plus générale. Ceci a pour but d'évaluer l'activité d'une entreprise sur le web. Ce professionnel se doit de sélectionner l'information pertinente afin de la traiter de manière pragmatique et d'en tirer les enseignements nécessaires qui pourront à la fois orienter les entreprises présentes sur le web à optimiser leurs stratégies tout en façonnant ses propres connaissances afin d'entrer dans un processus d'amélioration permanent le rendant ainsi plus efficace.

La rédaction de ce mémoire est donc la transcription de la réflexion liée à une problématique intéressante d'une part pour les sciences de l'information et de la communication et d'autre part pour l'activité d'analyse numérique. Cette réflexion a été orientée grâce à la problématique suivante : Comment les informations produites à partir

[1] L'expression anglaise *digital analyst* est toujours utilisée pour désigner le professionnel chargé de réaliser les analyses. Analyste numérique serait son équivalent en Français mais cet intitulé de poste n'est jamais rencontré.

l'analyse des activités en ligne permet de faciliter la prise de décision des entreprises afin d'orienter leurs stratégies ? À terme, nous devrions comprendre précisément le parcours de l'information, de la donnée brute, presque binaire sans intérêt lorsqu'elle n'est pas placée en contexte jusqu'à l'information stratégique qui permet de faire évoluer la performance de l'entreprise sur le web. Trois parties distinctes sont développées permettant de comprendre au mieux la réflexion sur ce sujet :

- Il est tout d'abord utile d'obtenir une meilleure connaissance de la matière première permettant de réaliser les études : la donnée brute. La provenance de ces données, les différentes manières de les obtenir, mais aussi l'optimisation de l'extraction de ces données à faible valeur vont être autant de sujets qui seront abordés. Obtenir ces données est un processus complexe et avant de réaliser des traitements sur ces dernières, il est important de s'assurer que cette base de travail est fiable et exempte de tous biais.

- Nous nous focaliserons ensuite sur une partie plus hétérogène portée sur l'essentiel du travail qualitatif de l'activité d'analyse. Le consultant effectue un métier complexe puisant sa connaissance dans des métiers très variés mais complémentaires. La connaissance du contexte, du traitement des données brutes ou de l'ergonomie sont autant de types de connaissances qui permettent de réaliser une exploration complète des dispositifs. Ces savoirs apportent des réponses précises aux objectifs d'analyse communiqués par les commanditaires.

- Enfin, nous porterons notre regard sur la transmission et la capitalisation des différents savoirs acquis durant ses analyses. Le consultant est un interlocuteur privilégié des entreprises, car cet analyste est capable de comprendre et de communiquer sur la performance des sites de manière objective. Une partie de la qualité de son travail passe majoritairement par la communication des résultats issus des différentes investigations. La qualité de son travail est aussi liée à un processus d'amélioration continu. Ce professionnel doit être capable capitaliser sur ces connaissances afin d'enrichir ses analyses tout en étant capable de les réaliser rapidement.

Au terme de la lecture de ce mémoire, nous devrions tout d'abord comprendre l'approche de ce professionnel et son rapport à l'information tout en permettant de voir comment cette information permet aux entreprises de générer à la fois un meilleur retour sur investissement et une satisfaction des visiteurs accrue.

Avant d'aborder le développement des différentes parties, plusieurs mises en garde sont à faire. Ce métier étant assez récent, il existe assez peu d'études scientifiques relatives à l'activité du consultant car la confidentialité des informations ou encore le savoir capitalisé constituent des éléments à forte valeur stratégique dans ce secteur très concurrentiel. L'ensemble des apports théoriques apportés pour soutenir les propos ne seront donc pas toujours liés directement à des articles de recherche mais à des livres, billets de blogues ou des connaissances acquises au coté de professionnels reconnus dans ce domaine. Le travail d'analyse numérique lié à ce métier emprunte toutefois des savoirs issus de domaines de compétences très différents comme les mathématiques, l'ergonomie ou encore au marketing. Il est alors plus aisé de trouver des sources pertinentes liées à ces domaines de compétence.

Enfin, la seconde remarque souligne le fait que la majorité du travail a été réalisée avec les outils fournis par la filiale de la société Alphabet : Google (Google Analytics, Google Tag management) et par la société Microsoft (Excel). L'ensemble des exemples fournis et les propos tenus sont donc majoritairement associés à ces outils. Il existe un certain nombre de différences entre les outils présents sur le marché et les outils avec lesquelles le travail s'est effectué au niveau des fonctionnalités, des calculs réalisés, et des possibilités mais, le travail sur la circulation de l'information devrait rester objectif car les modes opératoires et les méthodologies resteront les mêmes.

I. RECOLTE ET FIABILITE DES DONNEES QUANTITATIVES

L'activité du digital analyste base en partie sa réflexion sur l'investigation des données dites «brutes» ou encore quantitatives. Ces données qui n'ont subi aucun traitement sont extraites à partir des informations de navigation enregistrées sur les différents sites. La première tâche de cet acteur du digital est alors de s'assurer que la majorité des données nécessaires à l'analyse sont bien présentes et que les données récoltées sont fiables. Ce travail effectué, le consultant pourra avoir une première approche de la façon dont les usagers naviguent ou interagissent avec le site.

Ces données quantitatives sont essentielles afin d'installer une méthode s'approchant fortement d'une démarche scientifique afin de guider le processus de production des connaissances. Les analyses et les conclusions tirées ne seront pas alors uniquement réalisées à partir d'intuitions mais bien par l'observation rigoureuse des grandeurs et des mesures puis par les traitements effectués sur ces dernières. Tout au long du travail, ce consultant doit donc s'assurer de l'intégrité des données récoltées afin qu'elles s'approchent au plus près de la réalité.

I.1 Les données quantitatives : histoire, définition et précisions

1.a Histoire de l'analyse numérique

Il est trop périlleux d'avancer une date à laquelle le poste *digital analyst* en tant que tel a été créé. Cependant, il est possible d'obtenir une vision plus précise de l'évolution des techniques d'analyse numériques à partir de l'historique des outils permettant d'extraire les données relatives au trafic d'un site.

1993 : L'analyse des logs

Les prémices d'une analyse d'audience sur le web ont tout d'abord étés mises en place par l'analyse des fichiers de logs générés principalement grâce au **protocole http**[2]. Ce protocole est en mesure d'enregistrer toutes les interactions faites avec le serveur. Ces fichiers de logs permettent alors de répertorier l'ensemble des interactions réalisées entre le visiteur et le site web lors des différentes navigations.

```
70.45.2.1 – Guillaume [02/Jul/2016:03:20:10+0100] "GET
/memoire/version1 http/1.0" 200 25465
```

L'encart représente une ligne du contenu type que nous pouvons retrouver dans un fichier de log. Cette ligne représente une action effectuée entre un visiteur et un site. Nous pouvons constater la présence d'informations telles que l'adresse IP du visiteur (70.45.2.1), le nom du visiteur dans le cas où, le visiteur serait enregistré, la date et l'heure à laquelle la connexion a été établie, la nature de l'action (GET lors d'une consultation et SET lors d'une modification ou d'une mise en ligne de contenu), la page cible ou encore la taille de la page retournée. Ces fichiers de logs constituent les premiers outils ayant permis de réaliser des recherches sur l'audience des sites. Toutefois, des limites ont rapidement été rencontrées, car la qualité des données remontées par ces fichiers s'est dégradée principalement à cause des raisons suivantes :

[2] *HyperText Transfer Protocole* (ou Procotole de transfert hypertexte) : Ensemble de règles et de fonctionnalités permettant la mise en communication entre n'importe quel ordinateur disposant d'une connexion et un serveur sur lequel est stocké le site web.

- L'apparition des **robots de *crowling***[3] permettant **d'indexer**[4] les ressources sur le web qui se confondaient avec la navigation d'un visiteur et qui faussait les données.
- L'émergence de nouvelles technologies comme l'Ajax qui permet d'obtenir des informations différentes sans rechargement de la page. Il est alors impossible de définir précisément la ressource consultée.
- La présence de plus en plus importante des éléments multimédias comme les vidéos ou encore les présentations en lignes. L'analyse de log correspondant uniquement à une URL, il est impossible d'analyser les interactions avec ces éléments
- Certains internautes possèdent des adresses IP dynamiques et non une unique adresse IP fixe, il est donc impossible de définir, par exemple, précisément le nombre d'utilisateurs ayant consulté une ressource.
- A cause de l'ensemble de ces problèmes, l'analyse de log a été peu à peu dépréciée afin de laisser place à une technologie née en 1995 permettant une remontée de données plus efficace : Le **JavaScript**[5].

1997 : Première utilisation de JavaScript pour l'analyse d'audience

La principale réponse apportée aux soucis de précision liés à l'utilisation des fichiers de logs pour la mesure d'audience est l'utilisation d'un tag JavaScript. Un tag est une portion de code, intégré directement au sein du code d'un site déjà existant et a permis de pallier à la majorité des problèmes rencontrés par les fichiers de logs tout permettant la remontée de nouvelles informations. Ainsi, la zone géographique, le navigateur utilisé, la résolution du terminal sont autant d'informations qui peuvent être remontées sur un serveur et être accessible pour réaliser des analyses plus complètes.

L'un des avantages de l'utilisation de la technologie JavaScript pour la remontée de données est sa flexibilité. En effet, après configuration, il est possible d'obtenir de très nombreuses données différentes. Les principales limites sont la précision de certaines données remontées (comme la localisation d'un visiteur) et la législation en vigueur qui interdit l'extraction de certaines données comme son historique de navigation.

---

[3] Robots de *crawling* : Désigne les programmes permettant d'indexer le contenu sur le web.
[4] Indexer : provenant du terme indexation. Ensemble de techniques permettant de répertorier des collections de documents numériques afin d'en faciliter la recherche.
[5] Javascript : Langage de programmation de script essentiellement utilisé pour mettre en place de l'interactivité sur le web. Depuis peu, Javascript est aussi utilisé pour faciliter les communications avec le serveur.

Cette technologie est toujours utilisée de nos jours pour récolter les données de navigation. L'ensemble des évolutions notables que nous avons pu observer depuis cette innovation majeure se focalise davantage sur l'optimisation des différents outils d'analyse de trafic.

2005 : Entrée de Google Analytics sur le marché

Avant 2005, l'analyse d'audience sur les sites nécessitait un investissement important, car la majorité des outils permettant de réaliser les analyses étaient payants. C'est à cette période que la société Google décide d'acquérir la société Urchin Software Corporation produisant essentiellement des logiciels d'analyse d'audience. Par la même occasion, Google décide de mettre à disposition son outil d'analyse d'audience gratuitement pour les sites ayant un trafic ne dépassant pas les dix millions de pages vues par an et par compte. On pourrait s'interroger sur la rentabilité de cette stratégie, mais la réponse est simple : Google ne mise pas sur cet outil afin de générer son capital mais sur l'utilisation d'outils périphériques. Pour citer le principal, adWords une régie publicitaire en ligne payante permet de générer des revenus beaucoup plus importants.

2016 : Le marché des outils analytics aujourd'hui

À l'heure actuelle, il existe beaucoup d'outils différents permettant d'extraire et d'analyser le trafic d'un site. La majorité des outils proposés sont malgré tout payants et Google reste l'acteur majeur du marché, car il est le seul à proposer une version complète d'analyse d'audience gratuitement. Google Analytics est plus attractif, car il permet aux différents acteurs du web d'apprendre à maîtriser cet outil sans investissement inutile. Les technologies d'extraction de données sur le web sont assez matures et fiables. Le regard des entreprises se focalise de plus en plus sur la recherche de personnes capables d'analyser et d'interpréter les données tirées de ces outils.

1.b Définitions et caractéristiques des types de données quantitatives

L'activité d'analyse d'audience possède un vocabulaire qui lui est propre. Une fois les données remontées directement dans les outils d'analyses ou les différents

programmes utilisant ces données, il est important de connaître l'ensemble du vocabulaire utilisé et la manière dont sont calculées les données. Cela permettra de manipuler efficacement les données et de savoir où se diriger pour réaliser les investigations. Cette partie va permettre de définir la majorité du vocabulaire de base utilisé. Cette partie est essentiellement composée de vocabulaire mais, est intégrée directement au sein du mémoire, car la compréhension de ces informations fait partie intégrale de la maîtrise de l'activité d'analyse numérique.

Il est important de rappeler que ce vocabulaire est en partie lié à l'outil Google Analytics. Les définitions et les calculs réalisés pour calculer certaines métriques sont différents en fonction des outils du marché même si, après lecture, l'adaptation sur d'autres outils se fera plus facilement.

Tag : Partie de code JavaScript inséré au sein d'un site permettant de récupérer puis de transmettre les données de navigations standard sur des serveurs de stockage. A terme, ces informations permettent d'alimenter les outils d'analyse d'audience comme Google Analytics.

Google Analytics : Outil d'analyse d'audience le plus utilisé dans le monde. Cet outil permet d'interpréter et d'afficher les différentes informations stockées sur les serveurs ayant été préalablement extraites sur les sites.

Métrique : la métrique est l'unité quantitative de base utilisée pour les analyses. La métrique définit une caractéristique permettant de mesurer l'audience. La métrique peut définir simplement un temps passé sur une page, le nombre de personnes quittant une page ou le montant d'une transaction. Ces métriques peuvent prendre différentes formes mais elles se trouvent toujours sous forme de nombre entier, flottant ou encore en pourcentage.

Dimension : La dimension est un objet définissant plusieurs caractéristiques de la navigation comprenant ses propres métriques. Une dimension peut aussi bien définir le type de dispositif utilisé pour la navigation ou encore le navigateur utilisé.

Key Performance Indicator (KPI) : En Français, indicateur de performance clé, cette expression désigne la valeur clé répondant à la mesure des objectifs d'une entreprise. Les KPI sont définis par l'analyste faite en fonction des objectifs stratégiques

d'une entreprise. Ils peuvent représenter le nombre de transactions, le chiffre d'affaires ou encore le nombre de fois où est utilisé un formulaire de contact. Les KPI permettent de piloter les activités en ligne car ce sont des indicateurs simples mais parfaitement représentatifs de la performance de ces activités.

Segment : Ce terme issu du marketing permet de définir une portion de la population analysée, soit par des critères sociodémographiques, soit par des critères arbitrairement choisis en fonction d'une action effectuée par les visiteurs. Segmenter le trafic d'un site en fonction d'un critère sociodémographique peut être assez difficile car les données remontées par les outils ne permettent pas de s'assurer de la fiabilité des données comme, lorsque la segmentation se fait à partir de l'âge et du genre. Il est plus courant de segmenter la population en fonction des actions effectuées sur le site comme la visite d'une page précise ou encore l'interaction avec une fonctionnalité présente sur le site car cette segmentation fournit des données plus fiables.

Evénement : Un événement n'est pas présent de base dans les outils analytics. Un événement correspond à une action réalisée par un utilisateur dans un cadre bien précis. Par exemple, un événement déclenché lorsque le visiteur clique sur un bouton donné. Les événements sont mis en place par le plan de tracking qui va être développé dans la partie **I.2 Élaboration d'un plan de tracking.**

Google Tag Manager (GTM) : La gestion d'un grand nombre de tags peut être assez fastidieuse car tous ces tags doivent être insérés directement dans le code du site. Un autre souci posé est celui du déclenchement des tags. Certains tags ne peuvent être déclenchés que si d'autres tags ont étés préalablement déclenchés. Etant donné la difficulté apportée par la lisibilité du code JavaScript, il est préférable dans certains cas de faire appel à un gestionnaire de tags. Google Tag Manager est le gestionnaire le plus connu et permet de simplifier cette activité. Cette application en ligne permet de gérer la mise en place des tags et leurs déclenchements directement au sein d'une interface plus simple.

Session : La session est définie par la visite d'un site effectuée par un usager sur un unique navigateur. Si le visiteur ne navigue plus sur le site pendant 30 minutes (temps paramétrable dans les outils d'analyse d'audience), une nouvelle session débutera si le visiteur réalise une nouvelle navigation sur le site. C'est précisément cette métrique qui permet d'appréhender au mieux le trafic d'un site puisqu'il n'est pas possible de définir le

trafic à partir du nombre de visiteurs. La mise en place d'une mesure de trafic à partir du nombre réel de visiteurs est techniquement trop difficile.

Taux de rebond : Cette métrique dont le résultat est affiché en pourcentage est calculée de la manière suivante : nombre de sessions arrivant sur une page X / nombre de sessions quittant la page X sans interaction avec cette dernière. Si deux sessions entrent sur la page d'accueil du site et qu'une seule de ces deux sessions quitte cette même page, le taux de rebond sera alors de 50%.

Taux de sortie : à la différence du taux de rebond, le taux de sortie désigne le pourcentage de sortie d'une page donnée, peu importe la page d'atterrissage. Si deux sessions naviguent sur le site et quittent le site sur une page X, le taux de sortie de cette page sera de 100% si le site n'a enregistré aucune autre session passant par cette page.

Taux de conversion : Cette métrique définie en pourcentage représente le rapport entre le nombre de sessions et le nombre d'achats réalisés par ces sessions. Si sur 4 sessions, un seul achat est réalisé, le taux de conversion sera de 25%.

Pages vues : Le nombre de pages vues comptabilise l'ensemble du trafic étant passé par une page donnée. Si une session arrive sur la page d'accueil, navigue sur le site puis retourne à la page d'accueil, le nombre de pages vues pour la page d'accueil sera de 2.

Pages vues uniques : à la différence de la métrique de pages vues, cet indicateur mesure uniquement la visite d'une page durant la durée d'une session. Si une session arrive sur la page d'accueil, navigue sur le site puis retourne à la page d'accueil, le nombre de pages vues uniques pour la page d'accueil sera de 1 car l'adresse IP de la personne est enregistrée et l'outil détecte que le visiteur a déjà vu cette page.

Il n'est pas possible de définir ici tous le vocabulaire nécessaire à l'activité du digital analyste. Ces indicateurs permettent tout de même de comprendre la façon d'interpréter les données dans ces outils. Les indicateurs définis ici permettent de réaliser des premières analyses relatives à la performance des sites. Malgré tout, il n'est pas toujours possible de se fier aveuglément à ces données pour des raisons qui vont être développées dans le point suivant.

1.c Biais des données quantitatives et facteurs de fiabilité

Il existe beaucoup de raisons pour lesquelles, les données récoltées ne sont pas toujours fiables. Durant ce stage, nous avons pu observer un certain nombre de soucis rendant les analyses de données peu pertinentes ou dans le pire des cas inutiles. Tout d'abord, les problèmes de données remontées sont essentiellement dus aux outils permettant d'extraire les données de navigation. Si l'outil n'est pas intégré correctement dans le code ou s'il n'est pas correctement configuré, nous pouvons voir apparaître des irrégularités dans les chiffres qui rendent les analyses sur ces sites presque inutiles. Dans d'autres cas, ce sont les limites de l'outil qui ne permettent pas de travailler sur des données fiables. Une partie des biais dans les données qui sont facilement identifiables vont être développés dans cette partie.

Le premier problème rencontré est lié aux limites de l'outil et a été décelé lors de l'utilisation de Google Analytics dans sa version gratuite. Lorsqu'une analyse doit s'effectuer sur un segment précis de la population, la mention échantillon apparaît en haut à droite comme sur l'image ci-dessous :

Cette mention rend compte du fait que la génération des données n'a pas été faite sur l'ensemble des données concernant la portion de la population sélectionnée mais seulement sur une partie (ici, 22.99%) de l'ensemble des sessions sélectionnée arbitrairement. Une fonction d'extrapolation est ensuite appliquée aux informations pour fournir des données approximatives représentant l'ensemble de la période. En fonction de la longueur de la période sélectionnée, de la précision de la segmentation et des traitements appliqués, l'échantillonnage sera plus ou moins important ce qui ne permet pas de réaliser des analyses avec certitude.

La manière la plus simple d'éviter ce problème est tout simplement de réaliser les extractions de données sur des intervalles de temps plus faibles. Un intervalle de temps plus faible permet de limiter l'échantillonnage, mais ne permet pas de réaliser des analyses avec une approche plus globale. Effectuer une étude sur une période trop courte, risque de représenter des comportements inhabituels liés une offre spécifique ou des soldes. La solution pouvant être adoptée pour pallier à ce problème est d'extraire les données de

périodes courtes et de les combiner pour obtenir les données sur la période souhaitée. Il est alors chronophage de travailler sur des périodes de plusieurs mois. Une autre solution est tout simplement d'investir dans la solution payante de Google Analytics pour ne plus avoir de problème d'échantillonnage. La solution la plus efficace trouvée a été de réaliser un fichier d'extraction automatique de données qui permet d'obtenir des résultats très précis ce qui apporte un avantage important aux entreprises car cela permet de s'affranchir des coûts relatifs à l'achat d'une solution payante. Ce fichier est développé plus précisément dans la partie **I.3 Automatisation de l'extraction des données.**

Le second souci rencontré concerne cette fois la mauvaise configuration et la mauvaise mise en place du tracking. Un problème qui peut être souvent rencontré est le système de double tracking. Si deux codes JavaScript distincts destinés à récolter des informations sont présents sur une page et envoient les données au même compte de récolte, une grande partie des informations seront inexploitables. Les données concernant la consultation de page seront dupliquées et des métriques comme le taux de rebond seront anormalement basses ou élevées.

Concernant ce problème, il n'y a pas de solution directe à apporter. En effet, il n'est pas possible de nettoyer ou de modifier les données qui ont été récoltées par ce système. L'outil n'est donc pas exploitable dans sa globalité. La seule façon de pallier à ce problème est de supprimer un des deux codes de tracking le plus tôt possible pour rétablir une remontée d'information correcte. La détection de cette erreur peut se faire de plusieurs manières. Soit en visualisant les taux de rebond du site qui seront anormalement bas soit en utilisant un outil permettant de détecter l'ensemble des outils de tracking mis en place sur le site. L'outil utilisé durant le stage permettant de remonter les différents systèmes de tracking mis en place sur un site est une extension du navigateur Chrome appelée ObservePoint. Avant de débuter l'analyse des données d'un site, l'idéal est de réaliser une analyse du tracking rapide répertoriant les erreurs qui sont souvent rencontrées sur le compte des clients.

Un autre aspect important à prendre en compte pour réaliser les investigations constitue les biais inhérents à la façon dont les données sont récoltées. Ces données sont souvent fournies par les outils, mais il n'y a pas d'avertissements explicites permettant de dire si ces renseignements peuvent être utilisés pour décrire précisément un comportement. Pour comprendre ces difficultés, il est nécessaire de connaître la définition des différentes métriques mais aussi quand et comment ces données sont récoltées.

Dans le cas d'une session, par exemple, cette métrique définit la navigation d'un visiteur sur **un** seul ordinateur et **un** seul navigateur sans interruption durant un intervalle de temps précis effectué avant minuit (car le compteur de session se réinitialise à minuit). Autant dire qu'il est inutile d'associer le nombre de sessions au nombre de visiteurs d'un site. 3 sessions peuvent refléter : la navigation de 3 visiteurs différents **OU** La navigation d'un seul visiteur ayant navigué sur son mobile, sa tablette et son ordinateur **OU** la navigation d'un seul visiteur ayant navigué sur le site tout au long de la journée. La façon dont les métriques sont comptabilisées peut aussi devenir des biais si leurs méthodes de calcul ne sont pas parfaitement comprises.

I.2 Élaboration d'un plan de tracking

2.a Définition

Dans la partie précédente le tag JavaScript de base a déjà été mentionné. Ce tag JavaScript mit en place au sein du code d'un site web permet de transmettre les informations relatives à la navigation vers l'outil d'analyse d'audience. Toutefois, certaines de ces informations ne peuvent pas être obtenues par défaut et la mise en place d'une autre stratégie est nécessaire. Cette stratégie va permettre de définir et de mettre en place du code permettant de récolter les nouvelles données quantitatives qui seront nécessaires à une analyse plus fine.

Dans le cadre d'un site e-commerce par exemple, il n'est pas possible de connaître les produits qui ont été ajoutés au panier. Il n'est pas non plus possible de savoir si un visiteur a visionné une vidéo présente sur le site et si cette vidéo a eu un impact sur sa conversion. Toutes ces informations sont importantes, car elles permettent de connaître l'efficacité des fonctionnalités et du contenu mis en place sur le site. En effet, si les vidéos présentes sur un site ne sont pas visionnées, ou si la mise en place des vidéos ne permet pas d'augmenter la conversion, l'impact global est négatif puisque la présence d'une vidéo sur un site baisse les performances. En retirant ces vidéos, l'entreprise pourra aussi réaliser des économies car elle n'aura plus à réaliser ces vidéos.

Le JavaScript va permettre de définir de nouvelles informations concernant à la fois l'activité des visiteurs et les caractéristiques de ces derniers. Ces données seront ensuite remontées dans les outils d'analyse de la même manière que les données standards.

Par ailleurs, il est inutile de commencer à implémenter le code qui sera nécessaire à la remontée de données sans apporter une vision stratégique au travers d'une méthodologie. Il est important de définir en amont l'ensemble des données à remonter et la manière dont ces données seront affichées dans les outils d'analyse afin de faciliter l'accès à ces données pour l'entreprise d'une part et pour le professionnel qui sera en charge d'analyser ces données. Il est aussi important de limiter le nombre de données qui sont remontées car la mise en place de JavaScript sur les pages a toujours un impact sur le chargement des pages ou la fluidité de navigation.

L'ensemble du processus de réflexion sur les informations à remonter dans les outils d'analyse d'audience et l'activité de traduction de cette réflexion dans un langage informatique qui permettra de remonter ces données est appelé le plan de tracking (plus rarement appelé plan de marquage ou encore de taggage en français).

Le plan de marquage est finalement un ensemble de documents structurés et réalisés suivant une méthode rigoureuse décrivant les objectifs, les actions et les données qu'il est nécessaire de «tagger» afin de récolter les différentes informations.

2.b Méthodologie

La méthodologie présentée dans cette partie est intrinsèque à la stratégie de réalisation d'un plan de tracking mis en place au sein de l'agence. La réalisation d'un plan de tracking de son élaboration à sa mise en application est entièrement maîtrisée, mais les différentes étapes mentionnées dans ce dossier ne seront pas forcément celles que nous pourrions retrouver au sein d'autres entreprises.

Le processus de réalisation du plan de marquage a été l'une des activités les plus complexes rencontrées lors du stage étant donné les approches très différentes liées aux savoirs faire des entreprises mais aussi liées aux savoirs-faires appris des différents professionnels réalisant cette tâche. Plus globalement, le plan de tracking représente une façon de catégoriser les données et les pratiques à évaluer sur un site. De manière plus théorique, « *la catégorisation est une activité mentale qui consiste à placer un ensemble*

d'objets dans des catégories en fonction des similarités détectées » .[6] Cette activité repose donc en partie sur une perception subjective de son environnement et cette définition peut expliquer les différences constatées dans les différentes méthodologies de tracking.

Dans le cas choisi, le plan de tracking est séparé en 4 étapes distinctes ressemblant fortement à la gestion d'un projet web. Ces 4 phases ne sont pas nécessairement réalisées par la même agence mais nous allons les développer comme si la même agence était chargée du plan de tracking complet :

- L'étape préalable à la rédaction du plan est la définition du périmètre d'analyse et des objectifs de l'entreprise. Théoriquement, il est possible de traquer n'importe quelle action réalisée par un usager sur les sites ou les applications mais toutes, n'ont pas besoin d'être traquées car les informations remontées ne permettent pas de répondre aux besoins des entreprises. Durant cette étape, et en fonction de la maturité du client, la définition des besoins est réalisée à partir d'un document rédigé par l'entreprise ayant besoin d'un plan de marquage. Elle peut être aussi réalisée durant un entretien afin de définir les besoins de manière collaborative et de répondre aux interrogations/incertitudes des clients.

- La rédaction des spécifications fonctionnelles fait office de seconde étape. C'est le premier document qui sera rédigé pour réaliser le plan de tracking. La rédaction s'effectue toujours après un certain temps d'analyse car, il est important d'apporter le maximum d'objectivité. Cette étape correspond à la définition des différents événements à tracker (exemple : clic sur une vidéo, ajout d'un produit ua panier) et à la catégorisation du contenu du site. Au terme de cette activité, le premier livrable est réalisé. Il est censé être compris par le client qui pourra se faire une idée des données qui vont être traquées afin de voir si ces données correspondent bien à ses objectifs.

- Après validation du client, l'étape suivante concerne la traduction des différents besoins qui ont été consignés dans le plan de marquage fonctionnel. Cette activité permet simplement de transformer l'ensemble des fonctionnalités en JavaScript. Ce langage technique devra être compris facilement par les équipes techniques pour la partie implémentation du plan de tracking.

---

[6] Article conforme sur Fr.wikipedia.org. (2016). *Catégorisation*. [online] Accessible à l'adresse : https://fr.wikipedia.org/wiki/Cat%C3%A9gorisation.

- La dernière étape nommée implémentation ne rentre pas en compte dans l'activité du plan de tracking. Elle est toutefois nécessaire pour que l'ensemble de ce plan puisse être mis en œuvre afin que les données puissent remonter dans les outils d'analyse. Cette implémentation se fait soit en utilisant un système de *tag manager*[7]soit en programmant les différentes fonctionnalités directement à l'intérieur du site.

2.c Regard critique sur les difficultés de cette activité

Lors du stage, plusieurs plans de tracking ont pu être réalisés pour des clients posant des problématiques et des difficultés diverses:

- Les premiers plans de tracking ont été réalisés sur le client Vichy, marque internationale de cosmétique. Les besoins de mesure d'audience sont essentiellement liés au suivi précis de la navigation des internautes sur le site de la marque et sur le blogue. L'objectif était de réaliser la seconde étape du plan de tracking à savoir la réalisation du plan de tracking fonctionnel désignant les nouvelles fonctionnalités nécessaires à la remontée d'information. La principale difficulté de cet exercice résidait dans le fait que les plans de tracking fonctionnels était à réaliser sur les nouvelles pages du site et non sur le site en général, car un plan de tracking était déjà en place. Le plan de tracking devait alors s'adapter à une structure très rigoureuse. La difficulté liée à la réalisation de cette étape est alors liée à la nécessité de comprendre l'intégralité du plan de tracking existant avant de débuter les différents ajouts.

- Le second plan de tracking a été réalisé sur Pro-Duo. Pro-Duo est une enseigne traditionnelle de vente de produits professionnels de coiffure possédant aussi une activité en ligne. Pour ce client, toutes les étapes du plan de tracking ont dû être réalisées mais celle de l'implémentation technique n'a pas encore débuté à l'issue de la rédaction de ce mémoire. L'intention principale de cette enseigne en matière d'analyse est le suivit de la performance des ventes. La seule formation obtenue pour réaliser l'intégralité du plan s'est effectuée sur Vichy et plusieurs freins à l'élaboration étaient présents.

[7] Tag manager : Fait référence à Google Tag Manager, l'outil permettant d'organiser facilement le code du plan de tracking à mettre en place sur le site.

Tout d'abord, lors de l'élaboration du plan de tracking fonctionnel, le site était seulement en cours d'élaboration et le seul document aidant pour réaliser ce travail était la documentation technique récapitulant l'ensemble des fonctionnalités pouvant être mis en place sur le *Content Management System* [8]utilisé. Il a alors été difficile de réaliser ce plan sur des informations précise. Le travail de réalisation du plan de tracking a uniquement pu être débuté lorsque les maquettes du site ont été fournies. Malgré tout, même avec ces maquettes, il était difficile de réaliser ce plan de tracking puisqu'il n'était pas possible d'interagir avec ces dernières et avoir ainsi, obtenir une visibilité des différentes interactions possibles.

- Cofidis.be a été le dernier plan de tracking réalisé durant le stage. L'objectif ici était de réaliser la rédaction du plan de marquage fonctionnel afin de suivre la navigation des visiteurs sur leurs comptes clients. Ce plan a été beaucoup plus facile à réaliser que les plans précédents car la logique de réalisation du plan de marquage était assimilée. Comme sur le plan de tracking réalisé sur Vichy, La réalisation du plan de tracking n'est en fait qu'une amélioration d'un plan de tracking existant et doit respecter les différentes règles ayant été établies. De même, seules les maquettes du site ont été fournies et les problèmes rencontrés sur le plan de tracking du site Pro-Duo ont aussi été retrouvés sur cette activité.

La rédaction d'un plan de tracking est une activité stratégique pour les entreprises qui permet d'avoir une vue bien plus précise de la performance du site. Malgré tout, ce travail intellectuel est complexe et demande beaucoup de temps et d'apprentissage. En plus de cette complexité, la réalisation de ce type de document possède souvent des entraves comme nous avons pu le voir dans les trois différents cas présentés.

L'analyste doit malgré tout composer avec ces problèmes et s'arranger pour répondre aux différents besoins. Heureusement, l'efficacité d'un plan de tracking n'est pas liée à une mesure binaire. Il n'y a donc pas vraiment de plan de tracking raté et de plan de tracking réussi. L'idéal est de pouvoir suivre au mieux les différents objectifs. De plus, le plan de tracking n'est pas un document figé. Il est destiné à évoluer en fonction des modifications du site. Si un problème est rencontré ou si un élément n'est pas correctement configuré, le tracking pourra toujours être modifié.

---

[8] Content Management System (C.M.S.) : en Français, le Système de Management de Contenu est un outil permettant de créer et d'administrer des sites web de manière simple et intuitive.

I.3 Automatisation de l'extraction des données

3.a Objectifs et avantages

Les livrables qui peuvent faciliter la prise de décision des entreprises sont assez variés. Une entreprise peut par exemple avoir besoin d'un diagnostic complet permettant de comprendre pourquoi les résultats du site ne répondent pas assez aux différents objectifs à atteindre. Les objectifs dans le cas d'un site e-commerce sont en généralement assez similaires car le diagnostic s'oriente alors vers la compréhension de la navigation des visiteurs et la relation avec le chiffre d'affaires ou le nombre de transactions. Dans le cas d'un site qui n'a pas d'objectif de vente, le but est d'évaluer précisément le trafic du site et de donner des pistes permettant de diriger les visiteurs vers l'utilisation de certaines fonctionnalités.

Au sein de l'agence, ce livrable est appelé diagnostic critique. Le diagnostic critique a pour objectif de comprendre et de sonder les données de navigation du site afin de voir où se situent les problèmes de performance. Ensuite, ce livrable apporte des solutions ou des pistes d'amélioration afin de rendre le site plus performant. Dans ce cas, il est nécessaire de réaliser une analyse personnalisée permettant de voir s'il y a assez de visites sur le site puis dans le site, quelles sont les différentes pages visitées et quels sont les points de fuite rencontrés lors de la navigation.

Il y a, dans le cas des sites e-commerce, une certaine redondance dans les analyses. En effet, entre deux diagnostics nous pouvons souvent retrouver des méthodes d'analyse de données similaires. À partir du moment où nous pouvons remarquer une redondance, nous pouvons nous poser la question de l'automatisation de l'extraction de certaines données. Cette philosophie de réflexion est inspirée par le fameux principe **Don't Repeat Yourself**[9]. Durant le stage, plusieurs heures ont été nécessaires afin de réaliser les mêmes traitements mais sur des données différentes. Rapidement il a paru important de pouvoir gagner le maximum de temps sur ce type d'activité à faible valeur intellectuelle.

Dans d'autres cas, l'entreprise peut avoir besoin de mesurer sa performance dans le temps afin d'adapter sa stratégie en ligne. Il est alors nécessaire de réaliser un reporting

[9] Don't Repeat Yourself (DRY) : traduit en Français par l'expression *ne vous répétez pas,* Cette philosophie tirée de la programmation informatique consiste à mettre en place des stratégies permettant de supprimer toutes les activités redondantes.

régulier qui permettra de réaliser des analyses sur une période afin de suivre l'évolution des différents indicateurs clés.

La réalisation de programmes d'extraction automatique de données permet de minimiser les efforts nécessaires afin de pouvoir récolter les données. Ces programmes doivent permettent aux bénéficiaires de se focaliser sur les analyses et sur la préparation des données à analyser. L'élaboration de ce type de programme permet aussi de combattre « *l'accroissement constant du volume d'informations à traiter* » [10]

La réalisation de ce type de programme permet aussi de minimiser les erreurs possibles. En effectuant des extractions à la main et en les reportant dans des tableurs cela augmente les risques de réaliser des erreurs. Plus le nombre d'extraction augmente, plus la probabilité de faire des erreurs est grande à cause du facteur humain.

Par contre, mettre en place une extraction automatique des données est un chantier qui nécessite beaucoup de temps. Il est obligatoire de s'assurer que ce système pourra répondre à une grande partie des besoins tout en évitant d'effectuer des traitements qui ne seront utiles que dans une infime partie des cas. La réussite de ce type de programme est déterminée principalement par le retour sur investissement en termes de temps.

3.b Stratégie adoptée

Pour réaliser ce travail, Il est important de choisir un socle technique permettant d'être très proche du résultat attendu dans un temps assez limité. Étant donné le contexte de réalisation du programme, il n'est pas possible de réaliser un programme trop complexe afin de garder le temps nécessaire à l'apprentissage du métier de consultant et à la réalisation des livrables de qualité pour les clients pour ne pas corrompre la réputation de l'agence. Après analyse des différents livrables réalisés par l'agence, nous avons pu observer beaucoup de redondance dans les analyses et dans les extractions manuelles.

[10] Kalika, H. (2016). Surcharge informationnelle, urgence et TIC. l'effet temporel des technologies de l'information. *Management & Avenir,* [online] n° 13(3), pp.149-168. Available at: http://www.cairn.info/revue-management-et-avenir-2007-3-page-149.htm [Accessed 4 Sep. 2016].

Un cahier des charges fonctionnel a par la suite été réalisé pour exprimer le besoin et définir les données à récolter. Notre programme devant être en mesure de répondre aux besoins suivants :

- Génération de tunnels de conversion réalisés à partir de toutes les sessions peu importe la segmentation. Ce tunnel de conversion permet de visualiser rapidement le parcours des sessions et les différents points de fuite pour obtenir des informations concernant la performance de vente du site.
- Génération des informations relatives au panier permettant d'obtenir des chiffres très difficiles à obtenir sans utiliser cet outil comme le panier médian. Ce panier médian peut être comparé au panier moyen afin de déterminer si la série est homogène ou si au contraire le panier moyen est fortement influencé par des données non cohérentes avec la série.
- Génération des informations concernant la **valeur d'une session** ou la **valeur d'une page**[11].
- Génération de graphiques directement exportables dans les fichiers destinés à la présentation des analyses.
- Comportement d'achat des visiteurs en fonction de l'heure et du jour et du jour de la semaine.
- Consultation de toutes ces informations en fonction du type de terminal utilisé
- Extraction de toutes ces informations en un seul clic et temps de préparation du fichier en 30 minutes maximum.

Une fois le cahier des charges réalisé, l'objectif est d'effectuer une veille permettant d'obtenir un outil stable qui permet de développer ce type de fichier avec toutes les fonctionnalités citées. L'outil doit permettre d'extraire les informations contenues sur les bases de données de Google Analytics vers le tableur Excel pour réaliser facilement des traitements supplémentaires sur ces données.

Analytics Edge est l'add-on Excel qui a permis de répondre aux besoins du cahier des charges. Cette extension du tableur permet tout simplement de former des requêtes sur les bases de données de Google en sélectionnant nos métriques, variables et nos segments pour extraire toutes les informations voulues. En utilisant toutes les fonctionnalités de l'outil, nous avons pu extraire les données sans échantillonnage. Pour se

[11] La valeur de session ou la valeur de page est une métrique avancée des analyses. Cette métrique permet de donner une valeur pécuniaire à une session ou a une page afin de déterminer les éléments les plus stratégiques d'un point de vu rentabilité. La valeur d'une session est donnée par le calcul suivant : (Chiffre d'affaires+Valeur des objectifs) / nombre de sessions. Ce calcul est le même pour la valeur d'une page.

faire, nous avons simplement appliqué une règle permettant d'obtenir une précision d'extraction des données au jour près. Nous avons ensuite généré des formules Excel complexes de ce type :

=SOMMEPROD(BA:BA;AV:AV;(AS:AS="mobile")+0)/SOMME.SI(AS:AS;"mobile";AV:AV)

Ces formules permettent de rassembler toutes les données d'une plage au sein d'une seule valeur afin de synthétiser les données et faciliter l'interprétation. Cette formule permet par exemple de rassembler toutes les données concernant la navigation sur mobile.

Dans Excel, une page nous a permis de rassembler toutes les données nécessaires puis nous avons redirigé les données de cette page vers d'autres pages afin de structurer les analyses en fonction des besoins et de simplifier l'affichage des données. Par exemple, une page a été générée pour rediriger les données concernant le comportement des visiteurs ou encore une page a été générée pour rediriger les données d'achat.

Une fois le programme terminé, une phase de recette rigoureuse est mise en place pour vérifier si toutes les données et les traitements sont fiables. Durant cette phase, le programme est testé sur différents comptes puis une comparaison est effectuée avec les données de Google Analytics pour voir si les chiffres concordent. Ensuite le fichier est testé avec des intervalles de temps très longs pour s'assurer que ce fichier pourra répondre à des besoins d'analyse hétéroclites.

3.c Perspectives

Le fichier réalisé a répondu aux attentes tout en permettant de réaliser de nouvelles analyses. Il permet tout d'abord de gagner du temps en supprimant le travail d'extraction. Il apporte aussi l'assurance de travailler avec des données de qualité car il est inutile de réaliser des extractions à la main et de vérifier la qualité de ces données. Ce fichier a permis de faciliter le travail des diagnostics et a permis de se concentrer uniquement sur l'analyse des données. Il n'a pas été possible d'utiliser ce fichier sur d'autres activités comme le *reporting* mais ce travail ne devrait théoriquement pas poser problème.

La réalisation de ce fichier ouvre la voie à de nouvelles pistes de réflexion liées aux possibilités offertes par l'élaboration de programmes informatiques utilisant les ***Application Programming Interface***[12]. Développer un programme en le couplant à des API à la place d'un fichier Excel permet de combiner l'extraction de données avec les possibilités offertes par les langages de programmation, les bases de données et les différentes librairies. Ceci permet facilement de mettre en place des fonctionnalités de visualisation de données afin de fournir des analyses toujours plus complètes. La programmation permet d'ajouter bien plus de liberté que celles qui sont octroyées par l'utilisation d'un tableur. Ceci permet de pousser la réflexion sur l'automatisation des extractions.

Ce fichier a justifié la réflexion menant à l'élaboration d'un programme beaucoup plus performant permettant une bien meilleure flexibilité dans les analyses, une fiabilité technique accrue et une adaptabilité aux différentes problématiques posées par les clients. De plus ce fichier permettra d'être utilisé bien plus facilement par un public néophyte ce qui n'est pas le cas pour le fichier Excel. Même si ce programme permet d'extraire automatiquement les données, la configuration de ce fichier Excel nécessite une bonne connaissance du programme utilisé pour cette activité.

Rendre un fichier de ce type utilisable par un grand nombre de personnes sans nécessiter de connaissances techniques devrait théoriquement rendre le programme attrayant et commercialisable, car cela permettrait de rendre les entreprises plus autonomes face aux différents outils. L'objectif n'est pas de remplacer le consultant par un programme mais de permettre aux entreprises d'accéder plus rapidement à des données permettant d'adapter les stratégies le plus rapidement possible. Ce programme permettra aussi plus de confort pour le digital analyste. Il n'aura pas à extraire les données manuellement et ce programme lui fournira des données qu'il est impossible d'obtenir par lui-même.

Étant donné le milieu concurrentiel dans lequel la réflexion sur ces problématiques s'opère, Il n'est pas possible de développer plus précisément la stratégie d'élaboration de ce programme. Nous avons tout de même pu obtenir une première approche de sa réalisation au travers du fichier basique d'extraction de données réalisé sur Excel.

---

[12] Application Programming interface (A.P.I.) : En Français, Interface de programmation applicative est un ensemble cohérent de fonctionnalités pouvant être utilisées dans divers programmes. Ces fonctionnalités permettent de programmer plus rapidement en utilisant un code déjà réalisé par d'autres professionnels.

Conclusion de partie I

Cette première partie nous a permis de comprendre précisément la manière dont se construit le pilier principal de l'activité d'analyse d'un dispositif numérique. Les données brutes sont essentielles pour réaliser des analyses pertinentes et objectives. Même si nous avons vu qu'il était possible de faciliter ce travail d'analyse par l'extraction des données, nous avons aussi remarqué qu'il est par moments, difficile de travailler avec les données. Premièrement car elles ne sont pas toujours fiables et ensuite car il est nécessaire de réaliser des plans de tracking pour pouvoir pousser les analyses sur d'autres aspects de navigation. Il n'est pas possible de se reposer uniquement sur ces données pour réaliser les analyses les plus complètes.

Enfin, les données permettent en général de mettre en avant les différents problèmes rencontrés sur le site mais elles ne permettent pas d'apporter des réponses claires aux soucis relevés. Pour pouvoir apporter des réponses autres que la simple détection d'un problème, il est souvent nécessaire de faire intervenir des domaines de compétences différents. Etant donné les savoirs éclectiques nécessaires à la réalisation d'un site, les analyses doivent être capables de prendre en compte ces savoirs pour être complètes et précises. Ces analyses pourront alors accompagner efficacement la stratégie des entreprises.

II. ANALYSES QUALITATIVES DE L'ECOSYSTEME CLIENT

Même si la collecte et l'analyse des données quantitative doivent constituer en grande partie le savoir à transmettre par le digital analyste, il est nécessaire de faire intervenir un panel varié de compétences pour répondre efficacement aux besoins rencontrés lors des analyses. Les réponses que nous pouvons préconiser ne sont pas toujours apportées par l'unique analyse des données et il est important d'avoir une connaissance bien plus large de l'écosystème client et du digital en général pour pouvoir agir même lorsque les données ne permettent pas d'apporter une réponse concrète. Le digital analyste doit pouvoir s'adapter aux besoins des clients pour réaliser des benchmarkings concurrentiels, ou encore des analyses liées plus particulièrement à l'ergonomie des interfaces.

Dans cette partie, nous allons développer un grand nombre de connaissances et compétences qui sont développées par le digital analyste. L'ensemble de ces connaissances permettront de fournir des réponses aux clients lorsque les questions dépassent le périmètre d'investigation. Ces informations complémentaires que nous pourrions définir comme qualitatives sont essentielles pour faire le lien entre l'analyse des données et l'existant.

II.1 Traitement et utilisation des données brutes

1.a Méthodologie de type bottom-up

L'étape la plus importante de l'analyse des sites consiste dans l'interprétation et le traitement de ces données récoltées. Les traitements peuvent être séparés en deux méthodologies distinctes. La première est celle que nous allons développer dans cette partie. Le premier traitement est appelé l'analyse de type bottom-up ou en français, analyse ascendante. L'objectif de cette analyse est de se concentrer sur les indicateurs fournis par Google Analytics ou un autre outil afin de détecter des anomalies dans les données fiables ou de lier précisément les données à des comportements.

Cette analyse bottom-up est la plus complexe à appréhender, car elle nécessite une certaine expérience de l'investigation des données. Dans une analyse de type bottom-up, la navigation s'effectue au sein de l'outil de mesure d'audience et non au sein du site web même si une navigation préalable peut être effectuée. De par son expérience, l'analyste pourra obtenir des renseignements sur la performance du site et la façon dont les visiteurs interagissent avec ce dernier en analysant et en traitant uniquement les différentes données. Il pourra ensuite tirer les informations nécessaires à la détection des problèmes sur le site. Plusieurs difficultés sont tout de même à mettre en avant. En admettant que les données sont fiables et qu'il n'y a pas eu d'erreur lors de la remontée des informations, chaque site est différent et il est important de se replacer dans le contexte d'utilisation du site afin de ne pas réaliser d'analyses subjectives.

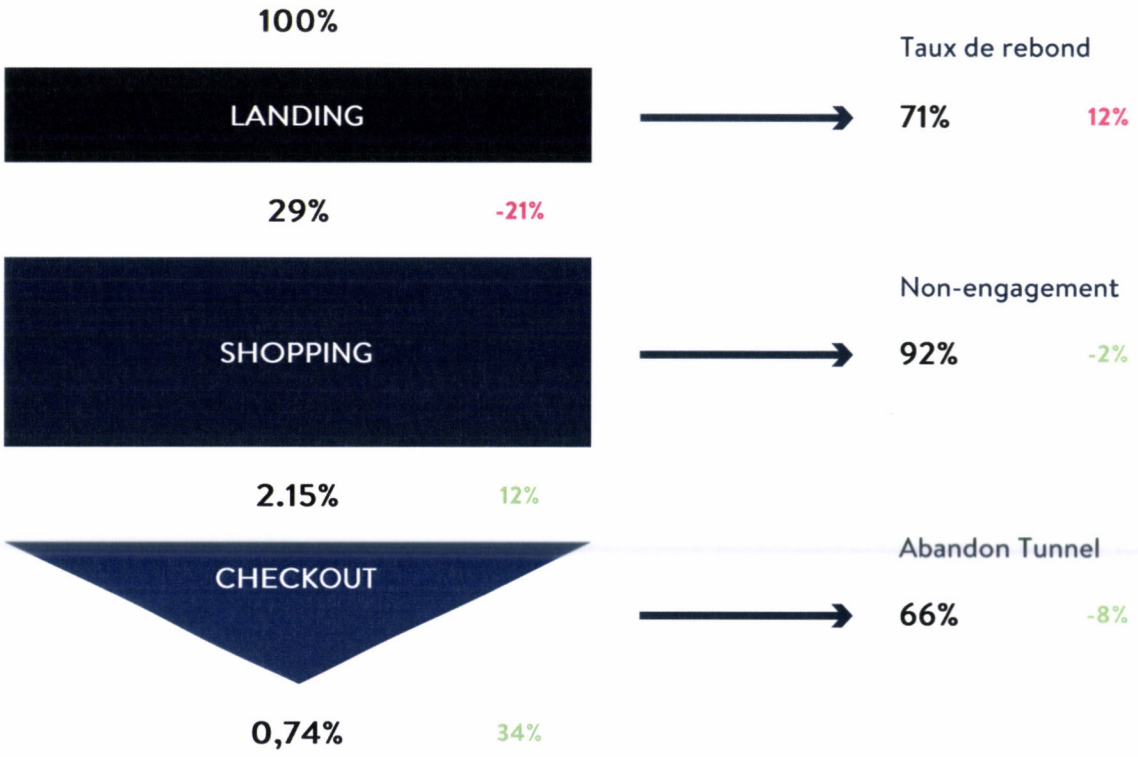
En prenant un exemple concret, un site e-commerce d'une marque de vêtements connue comme Brice et un blog d'articles destinés à faire le « buzz » ne fourniront pas du tout les mêmes comportements car l'utilisation est en effet très différente. Sur le site de vente, les personnes ont pour habitude de naviguer sur le site, de comparer les vêtements et de visionner plusieurs pages. Le taux de rebond doit donc être modéré. A l'inverse, sur le blog, les visiteurs accèdent généralement au site par les réseaux sociaux, consultent l'article puis quittent le site pour continuer la navigation sur les réseaux sociaux. Le taux de rebond sera alors bien plus important que sur le site de vente.

L'analyste doit avoir une bonne connaissance des différents usages afin de réaliser des analyses pragmatiques. C'est pourquoi, il peut être plus intéressant de débuter en agence plutôt que chez l'annonceur. Travailler en agence peut apporter beaucoup

d'enseignements car les agences travaillent souvent avec des clients très différents qui forment autant de problématiques. Ceci permet de se constituer une expérience solide très rapidement.

Cette méthodologie peut être difficile à appliquer pour un digital analyste ayant peu d'expérience, mais certaines structures de données peuvent aider à appliquer cette méthode. C'est le cas du tunnel de conversion. Le tunnel de conversion représente de façon synthétique la navigation type sur un site en fonction des usages. Toujours dans le cadre d'un site e-commerce, l'usage de base est le suivant : le visiteur arrive sur une page du site, puis recherche un produit parmi une liste, consulte la page produit, l'ajoute à son panier, puis procède enfin à l'achat. Le tunnel de conversion est capable de représenter rapidement ce comportement avec les données extraites des sites. Ensuite, de par l'analyse, il sera possible de visualiser les étapes où les visiteurs quittent le site afin de déterminer les pages qui posent problème.

Le schéma ci-dessous présente le tunnel de conversion réalisé pour le Jardin de Catherine, un site de vente de matériel destiné au jardinage et dont l'activité de vente se fait uniquement à partir du site. Le site a subi une refonte et le client a demandé de réaliser une analyse comparative de la performance du site entre les deux sites différents. Ce tunnel de conversion doit alors présenter la performance actuelle du site comparée à la performance du site pré-refonte.

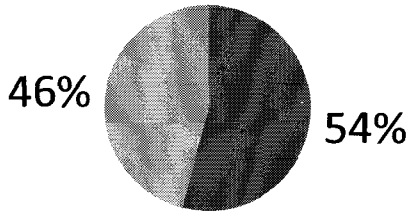


Les rapports en noirs sur la gauche expriment le nombre total de session arrivant à une étape de navigation par rapport à toutes les sessions. Les rapports en noirs sur la droite définissent le pourcentage de fuite entre deux étapes. Les rapports en couleur expriment le delta de croissance entre les données extraites de la période retenue sur 2016 et les données extraites de la période retenue sur 2015. La méthode de calcul du taux de croissance est la suivante :

$$\frac{(\textit{Données de la période courante} - \textit{Données de la période précédente})}{\textit{Données de la période précédente}}$$

Le résultat sera en rouge si l'indice est négatif et vert si l'indice est positif. Sans même se concentrer sur les deltas, nous pouvons observer un taux de rebond important entre la page d'atterrissage et la phase de shopping. L'objectif est alors de comprendre pourquoi ce taux est aussi élevé. A partir de cette donnée, plusieurs hypothèses se forment et après investigation des données, la véracité de ces hypothèses pourra être mesurée. Parmi les hypothèses établies, nous avons désiré savoir si les visiteurs arrivent sur un type de page dont le comportement logique augmente de manière significative ce taux de rebond. L'objectif est alors d'analyser quelles sont les différentes pages d'atterrissage.

Ce Graphique représentant le pourcentage de sessions accédant au site par Google Shopping, un comparateur de produit présent sur le moteur de recherche Google.



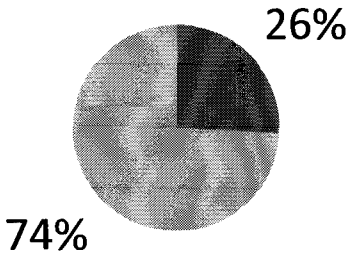
En 2016, 54% des sessions accèdent au site par cette fonctionnalité et 46% des sessions accèdent au site d'une autre manière.

Après analyses, nous avons constaté qu'une part importante des sessions arrivant sur le site est obtenue suite à l'utilisation de Google Shopping. Cette fonctionnalité présente sur le moteur de recherche de Google permet de comparer rapidement des

produits. Les visiteurs qui comparent accèdent au site Le Jardin de Catherine pour visualiser plus facilement l'article et quittent le site pour comparer d'autres articles présents sur d'autres sites en utilisant de nouveau le comparateur. Il est alors normal de rencontrer un grand taux de rebond sur ces pages d'atterrissage.

Nous sommes tout de même en droit de nous poser une autre question: pourquoi le delta de croissance est négatif entre 2016 et 2015 ? Le taux de rebond a augmenté de 21%, mais Google Shopping existait en 2015. Si les chiffres sont les mêmes, cela ne pourra pas expliquer l'augmentation du taux de rebond.

En 2015, seulement 26% des sessions accèdent au site par cette fonctionnalité et 74% des sessions accèdent au site d'une autre manière.



En 2015, l'utilisation de Google Shopping sur le site était beaucoup moins marquée. Seulement 26% des sessions accèdent au site par cette fonctionnalité face à 54% en 2016. Cette information apporte une réponse à la forte augmentation du taux de rebond entre les deux périodes.

L'analyse ascendante est très utile dans certains cas, car elle apporte des réponses qu'il n'est pas possible d'obtenir lorsque nous réalisons l'analyse uniquement sur le site. Une compréhension totale des données et la maîtrise de l'outil est importante pour réaliser ce type d'analyse. Malgré tout, même avec une investigation ascendante, il est important d'effectuer une analyse rigoureuse du site. Même si le raisonnement présenté dans cette partie apporte des éléments de réponse, une analyse in situ va permettre d'apporter d'autres pistes permettant d'augmenter la performance du site.

1.b Méthodologie de type top-down

A contre-courant de la méthodologie d'analyse précédente, la méthodologie top-down ou descendante permet d'infirmer ou d'affirmer de manière objective des analyses qui sont réalisées directement à partir de la navigation sur le site. La première étape de

cette méthodologie consiste donc à naviguer sur le site afin de réaliser des hypothèses sur les problèmes que pourraient rencontrer les visiteurs des sites et qui pourraient avoir un impact négatif sur la performance des entreprises. Cette méthode est aussi liée à des difficultés. L'analyste doit savoir adopter un regard neutre, en ayant la capacité d'adopter le regard d'un internaute type du site. L'analyste a l'habitude de naviguer et d'analyser les sites et à cause de cette expérience, son accommodation se fait bien plus rapidement. Il aura la capacité de comprendre plus rapidement le fonctionnement des interfaces pour obtenir l'information souhaitée. A cause de cette expérience, ce professionnel peut passer à côté de problèmes majeurs rencontrés par des internautes ayant moins d'expérience de navigation.

Lors des investigations, cette méthode est souvent introduite suite à la méthode ascendante, car elle est théoriquement moins rigoureuse. Cette méthode est tout de même très importante car l'analyste doit être capable d'adopter une réflexion objective sans travailler uniquement avec données récoltées.

Lors des analyses menées sur certains clients, des problèmes de tracking ont étés détectés. Notre exemple le plus représentatif est pris sur le client Speedy, entreprise Française d'entretiens et de réparation automobile. Un système de double tracking a été détecté suite à la réalisation d'un audit de tracking préalable (à voir dans la partie **4.b Démarche de type audit)**. Nous avons déjà vu dans la partie sur les biais que la présence d'un double tracking rend la majorité des données remontées dans les outils d'analyse d'audience fausses et inexploitables. Cette méthodologie top-down est donc envisagée pour fournir un travail convainquant et exploitable par l'entreprise.

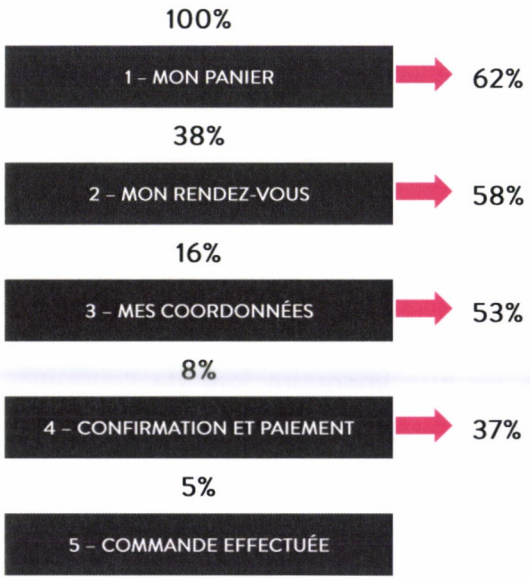
Toujours dans le cadre de l'analyse du site Speedy, une analyse du site a été effectuée et tous les bugs rencontrés sur le site ou les soucis pouvant poser un problème ergonomique ont étés reportés sur un document. A la suite, les différents problèmes sont catégorisés en fonction de leurs natures (bug, problème ergonomique) puis ces problèmes ont été priorisés en fonction de l'impact théorique sur les performances de vente du site. Pour exemple, deux soucis majeurs rencontrés vont être développés:

- nous avons détecté un bug important lors de l'utilisation de la fonctionnalité de recherche de pneu qui ne propose pas de produit lorsque l'utilisateur ne réalise pas une action spécifique. Cette fonctionnalité est pourtant un élément central du processus d'achat sur ce site. Dans le cas d'un bug ou d'un problème technique, il est plus facile de présenter ce problème sans obligatoirement faire appel aux données

car la démonstration du problème se fait simplement. Il est toutefois d'usage d'amener des preuves autres que les analyses réalisées car il est important pour les entreprises d'étudier l'impact de ces bugs sur la performance de leurs sites. Dans notre cas, les seules preuves apportées sont les enregistrements vidéo des sessions utilisateur ayant rencontré ce type de problème car dans le cas d'un double tracking, le taux de rebond et le taux de sortie n'est pas correct.

- Le second problème est présent dans le tunnel d'achat. Ce tunnel comprend, la page panier, la page de sélection du centre de montage pour l'achat de pneu et la page permettant de s'identifier. Ces pages ont posé beaucoup de problèmes lors de la navigation. Les informations ne sont pas affichées clairement, l'interface n'est pas claire et il est difficile de visualiser les informations importantes. Dans ce cas précis, il est complexe de mettre en avant les problèmes sans fournir des données illustrant ces propos.

En partant de notre analyse, il est logique que les visiteurs du site rencontrent des problèmes dans ce check-out et les données devraient concorder avec les analyses pour valider cette hypothèse. Dans la méthodologie descendante, l'analyse des données se fait après l'analyse du site et il est important de pouvoir lier les problèmes rencontrés aux données du site. Heureusement, lorsque le site possède un problème de double tracking, nous pouvons nous reposer sur certaines données qui sont théoriquement fiables. Le nombre de sessions n'est pas impacté par le double tracking est il est possible de voir si notre hypothèse est correcte.



Le schéma suivant représente le nombre de sessions passant par les différentes étapes du processus d'achat. Visiblement, sur la totalité les sessions accédant à la page panier, seules 5% des sessions procèdent à un achat. De plus, le nombre de fuite par étapes est en général supérieur à 50%, plus de la moitié des sessions interrompent leurs processus d'achat. Plus globalement, la métrique de conversion est de 0.15%. L'hypothèse est donc en partie validée. Les différentes étapes vont subir une analyse plus poussée permettant de mettre en avant les points bloquants.

Comme nous l'avons expliqué, cette analyse descendante nécessite l'utilisation d'outils d'analyse d'audience pour valider les hypothèses mais ce n'est pas tout. Réaliser une analyse descendante mobilise des connaissances approfondies en ergonomie permettant de détecter les problèmes et des outils permettant d'illustrer facilement les soucis de navigation rencontrés par les utilisateurs. La partie suivante va permettre de présenter le savoir mobilisé.

Attention, les deux méthodologies présentées ont été présentées séparément pour pouvoir facilement les distinguer. Dans le cadre d'une enquête sur un site, ces méthodologies sont toujours utilisées de concert afin de réaliser des analyses complètes. En fonction du site et en fonction des problèmes rencontrés dans la récolte de données, une des deux méthodes sera préférée mais, certaines recommandations ne peuvent être réalisées que par l'approche ascendante ou descendante.

II.2 Analyse Ergonomique

2.a Analyses par un dérivé de test utilisateur: Hotjar

Réaliser une analyse d'interface en impliquant les utilisateurs permet souvent de récolter plus d'informations qu'en effectuant les analyses soi-même. Cette implication de l'utilisateur permet d'évaluer le dispositif de manière moins subjective car les personnes ayant participé à la création du site ont une bonne connaissance de l'environnement et des interfaces. Dans le cadre de son activité, le digital analyste utilise aussi un outil permettant de réaliser un dérivé de tests utilisateurs.

Dans le cadre des analyses, il n'est toutefois pas possible de faire intervenir un panel d'utilisateur pour des raisons économiques car le coût d'un test utilisateur est assez important. Le logiciel Hotjar est alors utilisé pour appuyer les recommandations en fournissant un gage de fiabilité tout en permettant aux entreprises de réaliser des économies importantes.

Hotjar est un outil légal permettant d'extraire des vidéos de navigation à l'insu de l'usager afin d'analyser sa navigation sur le site. Ce logiciel permet de mettre en avant des problèmes rencontrés sur le site en visualisant ces vidéos. Cette application ne permet pas de remplacer complètement le test utilisateur et il est important de bien définir les avantages et les inconvénients de cette technique avant de pouvoir utiliser cet outil :

Avantages :

- Enregistrer le parcours d'un visiteur sur un site à son insu permet de pallier à une certaine subjectivité induite par le changement de cadre obligatoire pour enregistrer les vidéos avec un utilisateur choisi. Ceci passe par la modification du comportement et une altération de son attention. Par exemple, son attention peut être plus importante car l'usager n'est pas dans des conditions habituelles ce qui peut stimuler ses sens, et lui éviter de faire des erreurs. Dans notre cas, le visiteur est dans des conditions de tests avantageuses puisqu'il navigue sans savoir que sa navigation est enregistrée.

- Utiliser ce système permet de s'affranchir des coûts de tests. Le seul investissement nécessaire est le temps pris pour analyser les vidéos. Nous pouvons réaliser une étude plus qualitative mais sur un panel important de personnes.

- L'étude peut s'affiner au cours des lectures. En effet, les vidéos peuvent être filtrées à l'aide d'expressions régulières permettant de choisir les vidéos en rapport avec des pages précises. Lorsqu'un problème est décelé, il est alors possible de rechercher d'autres vidéos du même type pour voir si ce cas est isolé ou non.

Inconvénients :

- Les seules informations que nous pouvons extraire sont les déplacements de la souris et le temps pris pour réaliser certaines tâches. Sans voir les expressions du visage, il est difficile de savoir si la personne doute ou si son attention est focalisée/partagée.

- Il est impossible d'obtenir un feedback de la navigation sur le site ce qui est problématique car nous ne savons pas réellement ce que le visiteur pense de sa navigation sur le site ou encore quel est son profil.

- Le seul investissement nécessaire est en fait important car il est difficile de cadrer l'analyse selon nos besoins ou nos objectifs. Pour extraire des informations pertinentes, il est obligatoire d'observer beaucoup de vidéos.

Hotjar est un outil très pratique qui peut être utilisé de manières différentes. A l'instar de l'interprétation des données quantitatives, il peut être utilisé selon une logique ascendante ou descendante. L'analyste peut l'utiliser soit pour analyser le site au travers de la navigation des utilisateurs puis voir si les données concordent avec les problèmes rencontrés, soit, repérer des données singulières et voir si les utilisateurs rencontrent des problèmes sur les points relevés.

Utiliser cet outil permettra d'appuyer les recommandations ergonomiques en fournissant les différentes vidéos prouvant que les visiteurs ont rencontré certaines difficultés. Par contre, nous ne pourrons jamais obtenir des informations aussi précises que les tests utilisateur. L'utilisation de cet outil est tout de même assez aléatoire car il est possible de visionner plusieurs vidéos sans pour autant détecter de problème. L'enregistrement des sessions de navigation est aussi aléatoire puisque certaines sessions ne durent que quelques secondes et d'autres n'effectuent pas d'activité sur le site. Les recherches peuvent prendre beaucoup de temps.

Dans l'interface listant toutes les vidéos enregistrées, il est possible de filtrer ces vidéos en fonction du temps de navigation ou encore en fonction des pages visitées. Dans le cadre du stage, nous avons décidé d'utiliser cet outil selon une méthode en deux temps. Cette méthode a été appliquée au site Le Jardin de Catherine :

- La découverte du site s'effectue avec cet outil, pour se faire, nous sélectionnons les vidéos les plus longues et celles montrant des parcours naviguant sur beaucoup de pages afin de voir le site au travers des utilisateurs. Ceci permet de découvrir le site,

et de détecter les premiers problèmes. Par exemple, pour Le Jardin de Catherine, nous avons sélectionné des vidéos proches de cet enregistrement :

http://insights.hotjar.com/p?site=181371&recording=205064385&token=9203ab8f82ba3ae14f5ae69c46faaffe

Tout au long de la lecture de la vidéo, les analyses réalisées sont consignées dans un document. Ces remarques vont constituer les premières analyses du site :

*Vidéo 6 le **13/06/2016** sur **desktop** durée **13:20**:*

Le visiteur arrive sur la page d'accueil et effectue des recherches en naviguant par le menu et en utilisant la fonctionnalité de recherche. À 3 minutes 20, un bug est présent sur la fonctionnalité de zoom, l'image s'affiche dans sa résolution native et le visiteur est obligé de revenir en arrière car il n'arrive pas à fermer l'image. Nous pouvons observer des difficultés dans l'utilisation du menu à 5 minutes 20 (le visiteur clique tout simplement sur le logo du site car il n'arrive pas à cliquer sur une entrée du menu). à 6 minutes 44, il n'y a pas d'options de filtre alors que la page comporte beaucoup de produits. À 12 minutes, la fonctionnalité de zoom n'est pas pratique, car les boutons permettant d'afficher l'image suivant ne sont pas positionnés de manière fixe ce qui créé de la friction dans la navigation.

Suite à cette première étape, le travail habituel du digital analyste est réalisé. Les premières analyses top-down sont effectuées à partir du document rédigé pendant les analyses.

La seconde étape est effectuée pour valider les observations retenues durant les analyses ascendantes afin d'appuyer les hypothèses. Dans cette étape, nous filtrons les vidéos qui respectent une URL précise permettant ainsi de voir si les visiteurs rencontrent bien les problèmes détectés. Dans notre exemple, nous avions remarqué que le taux de rebond des pages produits était anormalement élevé et que la fonctionnalité de zoom était difficilement utilisable. L'hypothèse à évaluer et que cette fonctionnalité de zoom a un impact négatif sur le taux de rebond et fait fuir le visiteur.

http://insights.hotjar.com/p?site=181371&recording=205438640&token=41d247f8ec157302d5724fbbc2e6760d

Plusieurs vidéos comme celle présentée ci-dessus sont sélectionnées et sont ensuite transmises à l'entreprise pour prouver que ce problème a un réel impact sur la performance du site.

Hotjar est un outil permettant de faciliter le travail d'analyse ergonomique à condition de l'utiliser au travers d'une méthode précise afin de ne pas rendre l'analyse trop chronophage. L'analyse ergonomique passe aussi par la navigation sur le site et par la détection de problèmes qu'il n'est pas possible d'obtenir sur les vidéos, le digital analyste doit alors faire intervenir d'autres types de connaissances, plus théoriques.

2.b Apports théoriques

« L'ergonomie est vue comme l'étude scientifique des relations entre l'homme et ses moyens, méthodes et milieux de travail et l'application de ces connaissances à la conception de systèmes »[13]. L'objectif principal de l'ergonomie d'un site web est de permettre aux utilisateurs d'utiliser le dispositif avec un maximum de confort tout en permettant d'accéder rapidement à l'information recherchée.

Le professionnel en charge d'analyser les sites ne doit alors pas seulement détecter les problèmes rencontrés lors de la navigation mais il doit être aussi capable d'être force de proposition pour améliorer l'expérience utilisateur (appelé User eXpérience, UX en anglais). Les analyses vidéo des tests utilisateur permettent de mettre en avant les points bloquant mais ne permettent pas de savoir quel point pourrait améliorer la conversion sur le site et le rendre encore plus performant sans savoir relatif à l'ergonomie des sites.

L'expérience en matière d'ergonomie d'un digital analyste est souvent apportée par l'expérience et la lecture d'articles de recherches théoriques ou la lecture d'articles sur des sites spécialisés dans l'analyse ergonomique de sites existants. Ce travail est important car il va permettre de constituer plus souvent des hypothèses vérifiables proches des problèmes rencontrés sur les sites. L'idéal est de se constituer une bibliographie et ou une webographie complète sur les problématiques ergonomiques rencontrées largement sur le web. Cette base de données de connaissance peut être constituée :

[13] Citation empruntée à Fabien Duchateau, Maître de conférences et chercheur dans le domaine de l'ergonomie et de la conception des interfaces Homme-machine.

- D'articles concernant le comportement des visiteurs sur un site. Ces articles permettent par exemple de comprendre pourquoi un visiteur a confiance ou non dans un dispositif. Dans les investigations, il est ainsi possible de comprendre pourquoi le site n'est pas attractif. Ceci permet de vérifier si les visiteurs quittent le site dès la page d'accueil à cause de facteurs tels que les éléments de réassurance. Ces articles permettent aussi de donner des conseils pour mettre plus en confiance le visiteur.
- D'articles orientés design web et science cognitives permettant de conseiller l'utilisation **une police sans-serif** qui permet d'être plus facilement lue sur un dispositif numérique qu'**une police serif** qui rends la lecture plus difficile et plus fatigante. Ces articles permettront de formuler des recommandations plus pertinentes sur le design d'interface.
- D'articles orientés communication interculturelle permettant d'expliquer la différence de performance entre deux sites similaires mais situés dans des pays différents.
- D'études de cas. Des PDF rédigés par des professionnels sur des sujets très divers. Une grande partie des études de cas recherchés ici concernent la comparaison de la performance des interfaces dans le cadre **de tests A/B ou de tests multivariés**[14].

Il existe beaucoup de ressources présentes sur Internet mais pour obtenir des ressources de bonne qualité et s'assurer de la véracité des propos, il est conseillé de s'orienter vers les articles de recherche.

L'apprentissage de l'ergonomie reste assez difficile et il n'existe pas de méthode rapide pour apprendre l'ensemble des règles régissant cette science. L'ouvrage de référence pouvant être conseillé dans le cadre d'une première approche d'analyse d'un site ou d'une application reste Bastien, J.M.C., Scapin, D. (1993) Ergonomic Criteria for the Evaluation of Human-Computer interfaces. Ce livre rassemble l'ensemble des différents critères permettant d'analyser l'ergonomie d'un site sous de nombreuses facettes.

[14] Types de tests effectués sur un site permettant de proposer deux ou plusieurs versions d'une même page d'un site aux visiteurs. En fonction de la performance des versions, la version gagnante est sélectionnée pour être l'unique version qui sera accessible par les visiteurs.

II.3 Intelligence économique

3.a Benchmarking concurrentiel

L'intelligence économique peut être vu comme « la recherche et l'interprétation systématique de l'information accessible à tous, afin de décrypter les intentions des acteurs et de connaître leurs capacités.» [15]Il est alors tout naturel de prendre en compte ces actions dans le cadre de l'activité d'analyse pour avoir une vision plus large de la performance des sites analysés.

Dans cette optique, il n'est pas rare de réaliser des études comparatives ou *benchmarking* en anglais consistant à évaluer la concurrence d'un site analysé. Ceci dans le but de collecter des informations liées aux données et aux sites de la concurrence. L'activité de benchmarking en général est très importante pour une entreprise car concevoir et intégrer de nouvelles fonctionnalités mobilise de très nombreuses ressources. La comparaison et l'analyse de la concurrence sont des actions qui permettent d'apporter une réponse efficace à cet obstacle. Tout d'abord car ceci permets d'être à jour en minimisant les coûts de recherche et développement. Ensuite, si le benchmark est réalisé de manière efficace, ceci permet théoriquement de définir les meilleures fonctionnalités mises en place par la concurrence et ainsi, se mettre à niveau sans avoir à imaginer de nouveaux concepts.

Dans le cadre d'analyses numériques, ceci permet de proposer des solutions qui fonctionnent et pouvant répondre rapidement aux problèmes rencontrés. Dans un contexte où, le web mute très rapidement, il est important d'avoir la capacité d'interpréter efficacement les informations accessibles à tous. Cette stratégique permet aux entreprises de réaliser des économies et d'optimiser leurs stratégies.

La démarche d'analyse comparative consiste à naviguer sur les sites de la concurrence, de voir et d'évaluer les solutions apportées aux problématiques similaires aux sites sur lesquels nous travaillons. Cette activité débute après la rédaction d'un document permettant de structurer l'analyse selon les besoins puis de sélectionner les sites comparés. Dans notre cas, deux types de benchmarks sont mobilisés.

---

[15] Propos empruntés à Christian Harbulot, directeur de l'école de guerre économique à Paris à propos de l'intelligence économique.

Tout d'abord, le benchmark compétitif va nous permettre de comprendre le fonctionnement des sites concurrents opérant dans le même secteur d'activité. Si ce benchmark n'est pas concluant ou si les concurrents présents sur le marché sont inexistants, l'analyse comparative s'étendra aux sites n'opérant pas dans le même secteur d'activité pour trouver des situations similaires comme l'ajout au panier ou le formulaire de contact. Ce second benchmark mobilisé est appelé benchmark horizontal.

Cette activité de benchmarking permet de répondre à des questions auxquelles l'analyse de données centrée sur le site ne répond pas. La piètre performance d'un site, n'est pas forcément liée à son propre fonctionnement. Sa performance peut aussi être liée à la concurrence qui répond plus efficacement au besoin que ça soit par des prix plus attractifs, une meilleure ergonomie ou encore des fonctionnalités innovantes.

Durant ce stage, un benchmarking concurrentiel a été réalisé pour l'enseigne Speedy. L'objectif principal de cette enseigne était de comprendre pourquoi l'activité de vente de pneus en ligne n'était pas assez performante. L'organisation de ce benchmarking a été relativement simple : Nous avons tout d'abord décidé de sélectionner des acteurs majeurs de la vente de pneus en ligne et d'analyser la fonctionnalité de recherche de pneu et le tunnel de commande. La principale difficulté de cet exercice est constituée par l'absence de données des sites concurrents. Dans le cadre d'un benchmarking, il n'est pas possible de communiquer avec les différentes marques afin d'obtenir les données permettant d'évaluer l'efficacité des différentes fonctionnalités. La qualité d'un benchmarking est alors en partie liée aux connaissances de l'analyste.

Allopneus.com, Feu vert, et Norauto, constituent les acteurs majeurs de la vente de pneumatique en France. reifen.com, un site allemand, tirerack.com un site américain et tyrepac.com un site chinois ont aussi été sélectionnés. Nous avons décidé d'intégrer des acteurs internationaux au benchmark car même si la législation concernant les pneus est différente de celle mise en place au sein de l'union européenne, il peut être intéressant de voir comment se présentent les sites dans le monde. Choisir des sites internationaux permet d'obtenir une perception de la diversité culturelle et permettront de voir si des fonctionnalités sont différentes et théoriquement plus performantes que celles qui sont mises en place sur les sites nationaux.

Il est important de sélectionner des sites présents sur le territoire mais aussi présents dans le monde afin de pouvoir analyser les différences culturelles et obtenir un point de vue global de la présence numérique des entreprises opérant dans le même

secteur d'activité. Ces sites devraient nous permettre de relever de nouvelles façons de présenter les produits et de les rechercher ce qui est très important dans le cadre de notre benchmark. Il n'est pas possible de détailler l'ensemble du benchmark réalisé dans ce mémoire, mais il nous a permis d'obtenir beaucoup d'informations stratégiques.

Par exemple, Speedy possède des problèmes techniques présents dans la fonctionnalité de recherche qui ne permettent pas aux visiteurs de rechercher de produits. De par l'analyse, nous avons pu observer que les différentes fonctionnalités de recherche présentes sur les sites concurrents sont bien plus ergonomiques tout en fonctionnant parfaitement. Nous avons pu aussi observer que pour commander un produit sur Speedy, le client doit réaliser 55 actions différentes alors que la moyenne d'actions nécessaires à la commande sur les autres sites est de 40. Un benchmark peut alors mettre en avant les problèmes rencontrés sur un site et appuyer les recommandations en présentant les observations réalisées sur les sites concurrents afin d'apporter une réponse claire et facilitant la mise en place d'une stratégie adaptée.

3.b Veille multiforme

Alors que l'activité de benchmarking permet d'obtenir des informations stratégiques essentiellement orientées sur l'analyse des sites web, une action différente va être menée permettant l'acquisition d'informations stratégiques pertinentes. Le principal problème du benchmark est le fait que le travail est réalisé de manière passive. En effet, les différentes informations que nous pouvons obtenir de ce travail sont des informations qui sont déjà mises en place sur les sites. Les informations stratégiques tirées d'une veille doivent permettre d'obtenir bien plus d'informations que celles que nous pouvons extraire à partir de la comparaison des sites web.

L'objectif de la veille est d'être proactif afin d'obtenir des informations stratégiques directement liées aux différentes problématiques de l'analyse numérique. L'idée est de mener des recherches sur le web ou non et sur des supports très différents afin de tirer tout enseignement utile à la réalisation de l'activité d'exploration. Lors de ce stage, le travail de veille a vite été important car les compétences nécessaires à la réalisation des différents travaux devaient être acquises. Ce travail de veille a donc permis de répondre à deux problématiques différentes. Tout d'abord celle de la mise à niveau sur les pratiques

actuelles de l'activité d'analyse et ensuite sur les pratiques qui pourraient apporter une valeur ajoutée au travail réalisé au sein de l'agence.

Pour pouvoir réaliser une veille couvrant la majeure partie des besoins nécessaires à l'analyse numérique, plusieurs sources d'information sont mobilisées. Ces différentes sources permettent de réaliser une veille en rapport avec l'ensemble des domaines constituants l'analyse numérique. La liste suivante permet de répertorier les différents types de veille qui ont été effectués lors du travail au sein de d'altima° :

- Blog de digital analyst reconnus : centres d'intérêt prioritaires, les blogues d'analystes digitaux permettent de connaître de nouvelles stratégies d'analyses ou encore des méthodes permettant d'extraire des informations stratégiques à l'intention des entreprises. Ces blogues permettent aussi d'éviter de passer du temps dans la documentation des outils d'analyse d'audience pour connaître les nouvelles fonctionnalités car elles font souvent l'objet de billets de blogues assez bien détaillés pouvant faire office de tutoriels. Le travail de veille sur ces blogues a été nécessaire au début du stage car les compétences dans l'utilisation des outils et l'analyse des données quantitatives n'étaient pas satisfaisantes. Le blogue http://measuringthedigitalworld.com/ dont le contenu est rédigé en partie par Gary Angel, professionnel reconnu dans le domaine de l'analyse numérique a permis d'obtenir une bonne base de connaissances.

- Sites sur les nouvelles tendances en design/ergonomie : Ces sources d'information sont aussi importantes car elles permettent de connaître les interfaces à la mode et faciles d'utilisation. Réaliser une veille sur ces nouvelles tendances est moins pertinente que de lire des articles de recherche en ergonomie et design par exemple, mais permet d'obtenir des informations de bonne qualité plus rapidement. Ces sources sont donc utiles lorsque le digital analyste veut réaliser rapidement des **wireframe**[16] dans le cadre de préconisation sur l'ergonomie d'une interface ou pour simplifier un design.

  Cette activité peut se faire soit sur les sites communautaires de développement front-end comme codepen.io, soit sur les sites attribuant des notes aux meilleurs designs présents sur le web. Enfin, cette veille peut être effectuée sur les sites spécialisés en

[16] Wireframe : En Français, maquette fonctionnelle est un schéma très simplifié d'une interface. Cette maquette permet de positionner et de définir simplement les différentes zones de contenu que doit contenir une interface.

expérience utilisateur comme UXpin proposant le téléchargement de PDF très complet sur différents aspects comme les nouvelles tendances de design ou d'interface.

- Les livres concernant l'analyse digitale ou l'utilisation des outils manipulant les données. Altima° possède beaucoup d'ouvrages pertinents sur l'analyse digitale et l'ergonomie. La recherche de documents sur le web permet d'obtenir rapidement des informations variées et de qualité. Malheureusement, à partir d'un certain niveau de maîtrise des logiciels ou de connaissance sur un sujet précis, il n'est plus possible de trouver des ressources en ligne permettant d'obtenir de nouveaux savoirs gratuitement. La première possibilité est d'acheter des ressources en ligne mais il n'est pas possible d'être sûr de la qualité des savoirs apportés. La stratégie adoptée a simplement été de sélectionner les livres qui étaient présents à l'agence afin de réaliser ce travail de veille.

La veille réalisée au début du stage était essentiellement opérationnelle. Etant donné la nécessité de ce mettre à niveau dans beaucoup de domaines de compétence différents, la veille a été essentiellement utilisée pour constituer le savoir nécessaire à la maîtrise de l'activité. Au fur et à mesure du temps passé dans l'entreprise et de la constitution des connaissances, la veille s'est ensuite transformée en veille stratégique permettant d'anticiper les besoins et d'y répondre au travers d'actions concrètes. Cette veille a permis par exemple d'anticiper les problèmes liés à l'extraction de données.

II.4 Analyse Technique

4.a Justifications

L'analyse de l'implémentation et des caractéristiques techniques des sites n'est pas censée entrer dans le périmètre d'analyse assuré par le digital analyste. Toutefois il apparaît (d'un point de vu subjectif) important de disposer d'un minimum de connaissances techniques. Beaucoup d'études ont démontré que la qualité technique d'un site, facteur de performance en terme de chargement des pages ou de fluidité de navigation avait un réel impact sur la conversion des visiteurs mais aussi sur leurs fidélisations. De plus, les études en psychologies cognitives tendent à démontrer que les temps de latence au chargement

des pages peuvent être un facteur de perte d'attention des internautes durant le processus d'achat.

Pourquoi alors ne pas intégrer une dimension d'analyse technique permettant d'apporter des réponses claires et précises sur les problèmes techniques rencontrés sur les sites. Car l'objectif majeur des analyses est de permettre l'amélioration de la performance du site peu importe les moyens.

Les études réalisées sur la conversion des visiteurs ont permis d'apporter un savoir intéressant pour motiver l'entrée de l'analyse technique dans le périmètre d'analyse numérique permettant aux sites d'être toujours plus performants. L'exemple pris sur le terrain le plus simple reste le cas d'Amazon. Cette entreprise de commerce électronique américaine a réalisé des études démontrant la corrélation entre le temps de chargement d'une page et le chiffre d'affaire. Ces études ont démontré qu'un temps de chargement des pages moins important permettait de convertir plus facilement. Pourquoi alors, ne pas faire entrer le savoir faire technique dans l'analyse des sites afin de les rendre plus fluides, plus rapides à charger pour concourir à la réalisation d'un objectif omniprésent de l'analyse numérique : la conversion des visiteurs et l'augmentation du chiffre d'affaires.

Un background technique permet non seulement de remonter les différents bugs provoquant une baisse de performance du site mais aussi de fournir rapidement une solution simple et adaptée pour pallier aux problèmes rencontrés. Sans cette connaissance des environnements techniques, il peut être difficile de comprendre le problème et de le remonter efficacement. Il faut toutefois se mettre à l'évidence et bien prendre en compte que les consultants ne sont pas censés détenir de compétences en programmation ou dans la compréhension du fonctionnement des serveurs web. Il est alors complexe pour ces derniers de réaliser ce type d'analyse. Deux démarches différentes ont alors été développées permettant à tout type de profil de réaliser des investigations techniques de qualité.

4.b Démarche de type audit

L'audit est un processus d'analyse visant à évaluer la qualité d'un service. Ce travail d'audit est facilité par l'utilisation de référentiels faisant autorité et permettant d'appliquer une méthode rigoureuse. Le référentiel appliqué durant ce stage est la check-list. Une

check-list est une grille d'analyse permettant d'évaluer les bonnes pratiques d'un site web. Ces bonnes pratiques doivent évaluer l'ergonomie mais aussi d'autres aspects comme la pertinence des contenus ou la qualité du code.

Cette démarche d'analyse est dans notre cas, uniquement appliquée à l'évaluation des aspects techniques des sites web. Les aspects techniques sont séparés en deux parties distinctes. Tout d'abord, il y a l'analyse structurelle des sites (taille des fichiers chargés, organisation du chargement des fichiers du site) permettant d'optimiser la performance des sites pour augmenter la vitesse de chargement et optimiser le fonctionnement des serveurs sur lequel le site est consulté. Ensuite, un audit est réalisé sur le tracking mis en place sur le site. Durant le stage, plusieurs bugs de tracking étaient présents et il était difficile de comprendre pourquoi le taux de rebond était anormalement bas par exemple. Un audit de tracking permet de vérifier les erreurs les plus courantes rapidement.

La démarche de type audit fournit une expertise professionnelle effectuée à l'aide d'un référentiel fort et intuitif pour l'analyste aboutissant à un jugement pragmatique. L'idéal étant que cette démarche puisse être appliquée par des professionnels n'ayant nécessairement des compétences dans ces domaines. Le référentiel utilisé doit alors présenter une liste exposant les différents points de vérification tout en apportant la solution permettant de corriger ce bug. Ci-dessous, nous pouvons voir un exemple d'élément à vérifier lors de la réalisation d'un audit donnant le critère à analyser, la méthode permettant de détecter ce problème et la mise en place d'une solution appropriée

N°1 - Les feuilles de style sont minifiées.

Objectif

- Diminuer le volume de données à télécharger

Solution technique

Supprimer les espaces non nécessaires et les commentaires dans les fichiers CSS, par exemple en recourant à des outils tels que developer.yahoo.com/yui/compressor/

Moyen de controle

Vérifier l'absence d'espaces non nécessaires et de commentaires dans les fichiers CSS, par exemple à l'aide de l'extension Firefox ySlow

Thématique

CSS

Niveau

1

Crée le

22/03/2012 06h32

Automatisation

Automatisé

Pour réaliser ces audits, nous avons utilisé deux check-lists différentes :

- Le référentiel d'Opquast, une check-list utilisée par les professionnels du web pour évaluer les bonnes pratiques mises en œuvre au sein d'un site. Ce document constitue un référentiel fort puisqu'il est réalisé de manière collaborative dans une optique d'amélioration continue.
  La check-list se situe à l'adresse suivante : http://checklists.opquast.com/fr/

  Dans l'utilisation de ce référentiel pour réaliser les audits, il n'a pas été nécessaire de prendre toutes les catégories de la check-list car l'analyse se focalise uniquement sur la performance technique des sites. Nous analysons donc uniquement les points relatifs à la performance du site situés à cette adresse :
  http://checklists.opquast.com/webperf/

- Un autre référentiel dont la légitimité est discutable car réalisé à partir de lecture d'articles sur le web et réalisé uniquement pour l'agence. Ce référentiel liste une grande partie des erreurs de tracking que nous pouvons rencontrer sur les différents sites, la manière de les détecter et comment rétablir la situation. Au final, une quinzaine de points différents sont listés permettant de mettre en exergue des erreurs de tracking tel que le double tracking ayant été présenté dans la partie **1.c sur les biais des données** ou sur d'autres erreurs permettant d'optimiser le tracking sur des sites comprenant des sous domaines par exemple.

L'application de check-list donne des outils supplémentaires aux analystes permettant d'apporter des réponses claires aux problèmes présents sur les sites sans pour autant devoir comprendre parfaitement les environnements techniques. Ces check-lists permettent d'optimiser la qualité d'analyse et la vitesse de travail. Elles permettent aussi de minimiser la charge cognitive, car il n'est pas nécessaire de se rappeler des différents points à analyser. Réaliser ce travail d'analyse technique sans utiliser de check-list pourrait prendre plusieurs heures et il serait sans doute très difficile pour une personne n'ayant pas l'habitude de porter son regard sur le côté technique des analyses.

Réaliser l'audit permet aussi de ne pas se focaliser sur l'expérience de navigation réalisée à un instant T. Prenons un exemple simple : même si un visiteur n'a aucun problème pour naviguer sur un site de vente en ligne en temps habituel, les périodes de forte d'affluence peuvent devenir problématiques, car les serveurs sont bien plus sollicités.

Réaliser une analyse technique de type audit permettra donc d'anticiper au mieux les soucis pouvant être rencontrés lors de ces périodes.

4.c Démarche exploratoire

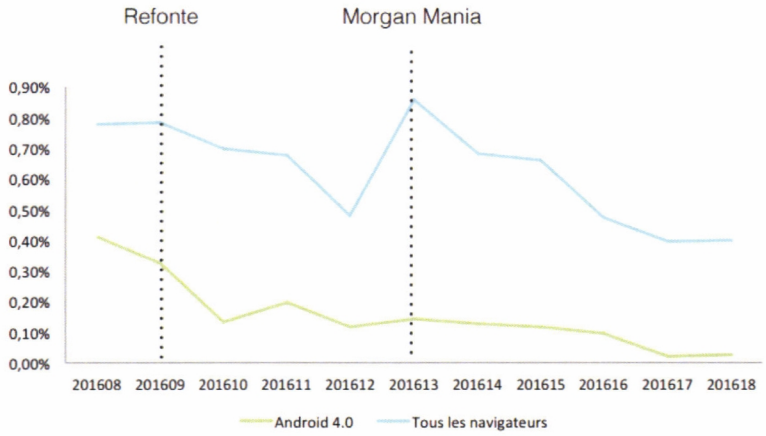
La seconde méthode dite, exploratoire, est très différente d'une méthode rigoureuse appliquée lors d'un audit. Cette méthode plus subjective est essentiellement liée à l'expérience des environnements techniques et des langages de programmation de la personne chargée de l'analyse. Cette démarche demande donc une connaissance technique plus importante que la démarche de type audit car les solutions aux problèmes présents sont souvent explicitées au sein des grilles d'évaluation. Dans la démarche exploratoire, Il est important de bénéficier d'une expérience en programmation tout en réalisant une veille constante sur les l'actualité des techniques de programmation.

Il est bien sûr impossible de connaître parfaitement l'ensemble des langages utilisés pour la réalisation des sites, et cette méthode ne peut pas être appliquée par tous les analystes. Mais cette démarche a son intérêt puisqu'elle permet d'apporter des réponses concrètes aux problèmes rencontrés sur le site. En 2016, plusieurs technologies sont largement répandues pour réaliser les sites : Le langage CSS permettant de mettre en forme les sites et le JavaScript permettant d'ajouter des fonctions interactives sont des langages de base qui peuvent être appris pour réaliser des analyses pertinentes sur la majorité des sites. Une bonne connaissance de ces langages permet d'apporter une solution face à un bug graphique ou à un bug de navigation par exemple.

La démarche exploratoire répond à des besoins d'analyses des problèmes techniques du site qui ne sont pas détectables par des outils. Ces bugs informatiques sont uniquement visibles par les visiteurs du site et il est difficile de définir ces problèmes sans mener d'analyses sur le site. Pour détecter les différents problèmes que les visiteurs peuvent rencontrer et avant même de débuter l'exploration du site, il est possible de faciliter le travail de plusieurs manières.

Tout d'abord, en prenant le design comme point d'analyse, il existe un certain nombre de restrictions liées aux différents moteurs de recherche. Par exemple, Internet explorer 8, n'est pas compatible avec certaines règles de mise en forme de contenu. Dans d'autres cas, la mise en forme des informations ne se fait pas correctement sur les

navigateurs mobiles. La première étape est donc de voir si certaines données liées aux différents navigateurs sont anormales.



Sur le graphique ci-dessus, nous pouvons observer que les sessions naviguant sur le site avec Android 4.0 convertissent peu et que la courbe de conversion décroît en fonction du temps. Nous pouvons supposer d'après ces données que le site possède des bugs empêchant le visiteur de passer commande. Il est toujours important de valider ces suppositions avec des éléments concrets. C'est à ce moment que la méthode exploratoire est mise en place car il est impossible de détecter les problèmes autrement qu'en analysant le site en utilisant le navigateur posant problème.

Après exploration du site sur le navigateur posant problème, nous avons détecté un bug qui ne permet pas au visiteur d'ajouter un produit au panier. Pour ajouter un produit au panier, il est obligatoire de choisir une taille mais la photo de l'article masque le choix de la taille car les règles CSS n'ont pas été correctement implémentées.

Cette démarche reste intéressante dans le cadre de l'analyse d'un site possédant beaucoup de problèmes techniques, mais elle reste tout de même aléatoire car l'ensemble des problèmes remontés sont souvent trouvés grâce à la navigation sur les sites. Pour

détecter les problèmes, il est nécessaire de d'effectuer plusieurs navigations sur le site en testant toutes les fonctionnalités proposées de manière différente ce qui prends beaucoup de temps lorsque l'analyse se porte sur un site e-commerce car beaucoup d'outils sont proposés aux visiteurs. Par moments, certains problèmes sont trouvés par pur hasard. Il n'est donc pas conseillé de mobiliser trop de temps pour cette démarche.

Conclusion de partie II

Nous avons vu que l'analyse digitale mobilise des savoirs très différents. Le digital analyste doit savoir adopter une multitude de points de vue afin de réaliser des analyses complètes permettant de renforcer efficacement la stratégie des entreprises. Posséder l'ensemble de ces connaissances n'est pas du tout obligatoire et l'analyse doit théoriquement se focaliser sur l'analyse des données récoltées dans les outils d'analyse. Malgré tout, ces connaissances peuvent aider les entreprises à faire confiance au professionnel car il apportera des réponses complètes rendant son profil plus attractif car ses analyses auront plus d'impact sur le pilotage des activités.

Cette partie permet de montrer la complexité liée à cette activité. Cette partie permet aussi d'apporter une réponse à la difficulté de définir précisément à la fois le poste et les compétences attendues sur ce poste. Les parcours possibles permettant d'accéder au poste de consultant digital analyst sont très différents et les compétences qui seront apportées suivront cette tendance. Les champs d'expertise varient fortement d'un professionnel à l'autre, en fonction de l'entreprise mais aussi en fonction de l'affinité avec les différents domaines d'expertise applicables à l'analyse.

Les attentes des commanditaires varient tout autant. Certains besoins demanderont un profil plus technique, ou au contraire un profil plus orienté marketing. Il n'a pas été possible de présenter toutes les compétences mobilisables dans le cadre des analyses. Nous aurions pu ajouter une partie orientée communication interculturelle car les analyses des sites peuvent être renforcées par la connaissance des différences culturelles inhérentes aux secteurs d'activité et aux pays pour lesquels l'analyse s'effectue.

III. CAPITALISATION ET TRANSMISSION DU SAVOIR

L'analyse des sites réalisée selon des points de vue très divers, la dernière étape peut débuter.

Cette étape concerne tout d'abord la transmission des connaissances à l'intention des entreprises bénéficiaires de l'expertise. Cette action a un impact fort sur l'objectif principal de son activité : faciliter la prise de décision des entreprises dans le cadre de leurs activités en ligne. Ceci permettra au bénéficiaire d'évaluer sa performance et les actions à entreprendre pour améliorer l'efficacité de son activité. Lors de cette étape, le consultant va synthétiser ses analyses, les vulgariser et enfin les organiser afin de transmettre efficacement le résultat du travail réalisé en amont.

Cette étape concerne ensuite, l'activité de capitalisation de ce savoir. C'est-à-dire, l'ensemble des actions mises en place qui permettent de préserver les connaissances concourant à rendre l'activité d'analyse toujours plus performante. Cette capitalisation permet aussi de partager cette connaissance avec son équipe. Cette activité va permettre de mettre en place une gestion documentaire des connaissances permettant de réaliser des analyses de qualité en conservant une productivité et une qualité optimale.

La qualité des analyses faites par le consultant est enfin liée à un facteur qui a toujours été important dans l'histoire du travail : la gestion du temps. Les compétences et l'expérience ne sont pas les seuls facteurs à prendre en compte pour évaluer la qualité d'une production car le temps est une variable essentielle permettant de régir l'ensemble de ces activités.

III.1 Transmission des connaissances : une vulgarisation rigoureuse

1.a Supports de communication utilisés

La transmission du savoir du consultant vers les différents commanditaires est une étape tout aussi importante que l'ensemble du travail réalisé en amont afin de faire en sorte que la situation de communication soit la plus confortable possible. Afin de transmettre les analyses qui résultent de la réflexion de manière détaillée avec une perte d'informations minimales, ce professionnel ne doit pas simplement communiquer. Il doit être capable d'effectuer un travail de médiation. Pour tenter de définir précisément cette activité dans le cadre de la rédaction du mémoire, la médiation pourrait, à notre sens se définir comme :

« *La pratique ou discipline qui vise à faciliter la circulation d'information, éclaircir ou rétablir des relations au travers de dispositif ou de processus dans le cadre de la transmission des connaissances issues de l'analyse d'un site web ou d'une application.* »

Cette partie est donc dédiée aux dispositifs utilisés dans la médiation. Dans ce cadre, il est nécessaire de procéder une sélection d'outils qui peuvent devenir des dispositifs de médiation efficaces et une sélection des formats de données qui pourront faciliter la compréhension des données jugées pertinentes. Durant le stage, l'utilisation de certains des outils ont été conseillés par l'agence et d'autres outils ont été testés pour tenter de transmettre plus efficacement les analyses.

Auparavant, Excel était utilisé par l'agence pour transmettre les résultats des différents diagnostics effectués sur les sites. Ce tableur édité par Microsoft est très pratique pour travailler et mettre en forme les données mais ce n'est pas l'outil idéal pour synthétiser les analyses car ce logiciel n'est pas assez pratique pour structurer les informations. Cet outil a rapidement été délaissé pour utiliser un dispositif de présentation très connu : PowerPoint.

Ce logiciel de présentation aussi édité par Microsoft permet de présenter les résultats plus simplement. Cette simplicité est utile pour le consultant qui pourra présenter une idée par diapositive, et structurer les informations de manière linéaire. Elle sera aussi

utile pour les récepteurs de la communication qui ne feront pas face à une surcharge informationnelle qui rend tout processus d'apprentissage plus difficile. Toutefois l'utilisation de cet outil nous amène face aux principales contraintes : Les informations doivent être sélectionnées et la quantité, limitée. Les propos tenus doivent aussi être simplifiés. L'objectif est de montrer le maximum d'informations avec un minimum de données en sélectionnant précisément les différents éléments permettant d'illustrer le problème.

Excel n'a pas été délaissé pour autant car il n'est pas possible de réaliser des graphiques statistiques sur PowerPoint. Cette solution est toujours utilisée pour réaliser les différents traitements sur les données et pour générer les graphes qui seront intégrés dans PowerPoint.

D'autres dispositifs ont été utilisés de manière plus expérimentale. Des logiciels de Wireframing ont été choisis pour transmettre les connaissances. Enfin, nous pouvons rappeler l'utilisation d'Hotjar en tant que dispositif de médiation. En effet la présentation de vidéos permet d'illustrer simplement et efficacement les problèmes rencontrés par les utilisateurs.

1.b Vulgarisation des analyses

La vulgarisation désigne le : *« Fait d'adapter des notions, des connaissances scientifiques ou techniques afin de les rendre compréhensibles au non-spécialiste; reformulation d'un discours spécialisé qui consiste généralement à le débarrasser de ses difficultés spécifiques, de ses caractères techniques afin de le rendre accessible au grand public »* [17]

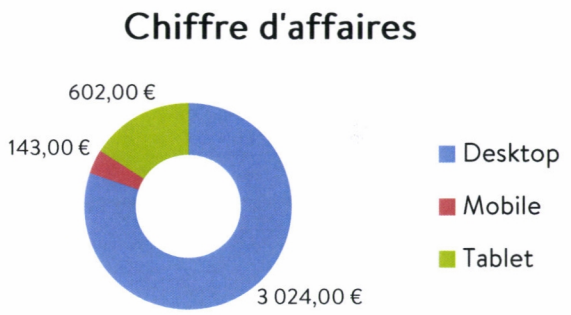
Dans le cadre des analyses, cette vulgarisation vise à simplifier les analyses réalisées sur les sites faisant appel à des connaissances spécifiques ou non. L'objectif est de fournir aux entreprises les informations présentées de façon très simple tout en veillant à ce que le savoir soit transmis avec la même qualité. Cet aspect non négligeable de l'activité d'analyse numérique est la dernière étape qui permettra de rendre tout le travail en amont effectif et efficace pour la stratégie des entreprises.

[17] Cnrtl.fr. (2016). *VULGARISATION : Définition de VULGARISATION*. [online] Consultable à l'adresse : http://www.cnrtl.fr/definition/vulgarisation [Version du 1 Sep. 2016].

Cette vulgarisation des analyses passe en premier lieu par une étape de traduction des différents discours spécifiques aux différentes expertises. Cette traduction vise à synthétiser les analyses en remplaçant tous les mots complexes par leurs équivalences. Il n'est pas possible de transmettre le contenu d'un article de recherche par exemple. Il est nécessaire de faire en sorte que le discours puisse être compris par tous les destinataires du message car dans la situation de restitution des diagnostics par exemple, plusieurs personnes sont présentes et possèdent des expertises très différentes. Dans le cadre des présentations, il est nécessaire d'intégrer des éléments plus graphiques comme des captures d'écran permettant d'illustrer les problèmes.

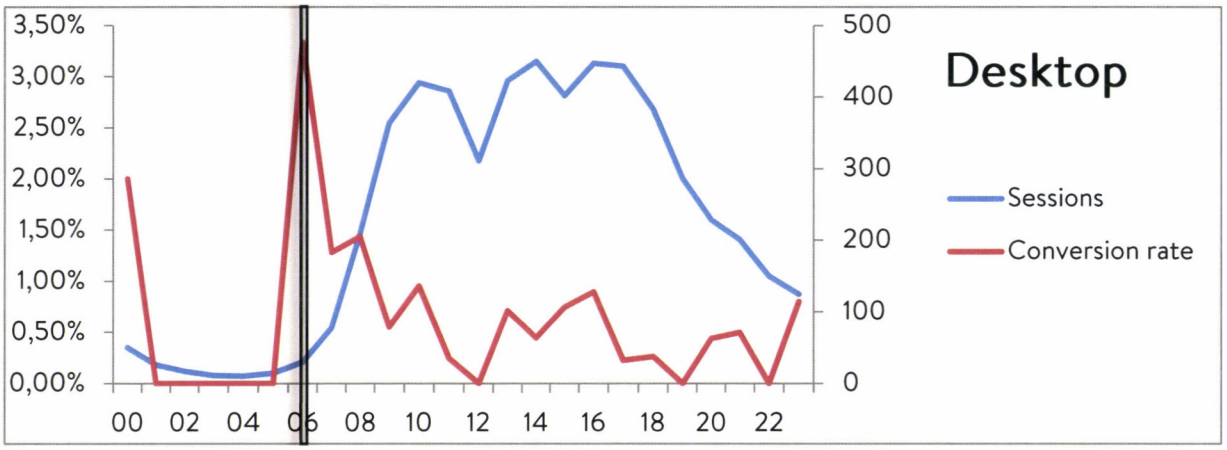
Pour comprendre au mieux la performance des sites web, il est aussi important de pouvoir fournir des données chiffrées permettant de faciliter la prise de décision à partir de données numériques. Dans cette optique, les présentations intègres des graphiques statistiques. Le choix des graphiques doit permettre de transmettre facilement les informations le plus rapidement possible.

La comparaison un jeu de données conservant une relation entre les différentes variables analysées pourra par exemple être représentée par un graphique en anneau.



Le lecteur pourra rapidement appréhender le rapport entre les données pour s'en faire une idée globale.

Dans des cas de transmission d'informations plus complexes, il sera important de faciliter au mieux la lecture en réfléchissant au format idéal. Dans l'exemple choisi ci-dessous, le format de données doit représenter à la fois le taux de conversion (ici désigné par l'expression anglaise *Convertion rate)* et le trafic sur le site en fonction du temps. Ce graphique doit permettre de transmettre l'information aux différentes personnes assistant à la présentation. Mais ce format de données permettra aussi d'être facilement interprété par d'autres personnes qui n'ont pas assisté à la présentation. Ces données peuvent par exemple être utilisées ensuite par les professionnels chargés du *Search Engine Advertising* (termes désignant le référencement payant) pour adapter la stratégie de présence en ligne de l'entreprise.



D'autres formats peuvent aider à synthétiser des recommandations. Plusieurs *wireframe* ont étés réalisés durant les diagnostics afin de fournir des solutions claires permettant d'améliorer les interfaces posant problème car expliquer toutes les modifications peut devenir vite trop verbeux. Dans le cadre de l'analyse du site Mr Bricolage, un wireframe a été produit pour formuler une recommandation de la page de livraison permettant de schématiser les propositions d'amélioration.

Utiliser un graphique permet de ne pas simplement énoncer les problèmes du site et fournir une idée d'amélioration pour que la situation de communication soit plus agréable. Sans la réalisation de wireframe, plusieurs diapositives auraient été nécessaires pour formuler des recommandations portées à la fois sur le design, l'ergonomie et la structure de l'information du site.

1.c Structurer les connaissances transmises

L'association des dispositifs et les informations vulgarisées ne permettent pas de transmettre les informations de manière optimale. Il est important de pouvoir ordonner ce savoir, en le catégorisant et en lui faisant suivre un parcours spécifique permettant de former un discours homogène pour que *« Le tout puisse avoir plus de valeur que la somme des parties le constituant»* [18]

[18] Aristote, *La Métaphysique*, trad. Annick Jaulin, PUF, 1999

Tout aussi important que sur les sites web, la structuration des informations permettra de s'assurer que les analyses seront transmises efficacement. La transmission du savoir vulgarisé passe essentiellement par la structuration de la présentation et par des rappels permettant de synthétiser un ensemble d'informations conservant des caractéristiques communes.

Ce travail de structuration a été grandement simplifié par la lecture de documents produits par l'équipe des analystes de l'agence. La structure des informations qui seront délivrées suivent généralement la structure suivante :

- Introduction du site avec présentation du contexte de l'analyse. Cette introduction doit permettre de comprendre pourquoi le diagnostic est réalisé.

- Présentation du périmètre de l'analyse et des dates sur lesquelles les données vont être analysées. Cette étape permet de savoir de quelle manière le site va être analysé.

- Présentations des données clés du site permettant avant tout d'avoir un aperçu synthétique des performances globales. Puis au fil des diapositives, les données se focalisent sur des éléments de plus en plus précis permettant de mettre en avant des informations beaucoup plus précises.

- Présentation des analyses techniques et ergonomiques du site. Les données précédentes permettent de faire facilement le lien avec les différentes recommandations. Cette présentation rassemble en premier les analyses techniques sur le plan de tracking et les bugs les plus importants. Puis les analyses ergonomiques et les recommandations sont classées selon l'ordre de visite. Pour synthétiser, la présentation se focalisera tour à tour sur : la page d'accueil, la page listant les produits, la page présentant un produit, le panier, et la page paiement.

- Enfin, un ensemble de diapositives permettra de présenter les 3 actions qui devraient apporter le meilleur retour sur investissement à l'entreprise puis de synthétiser l'ensemble des informations présentées.

Structurer de cette manière les différentes présentations n'est pas uniquement utile pour présenter le travail. Il permet d'harmoniser les pratiques professionnelles et permet de

faciliter la lecture des travaux réalisés par les différents consultants de l'agence pour optimiser la circulation du savoir.

III.2 Capitalisation sur les connaissances

2.a Les types de savoirs constituant les connaissances

Les différents travaux menés doivent permettre de participer à l'amélioration des compétences du consultant. Il est nécessaire dans le cadre d'une étude de la circulation de l'information d'expliciter les différents enseignements tirés des expériences afin de lier ces expériences aux différents types de savoirs générés. Ces types de savoirs permettent d'optimiser et de capitaliser au mieux les informations et les connaissances qui en découlent. L'activité permettant d'avoir la meilleure vision des types d'enseignement acquis reste celle du diagnostic car elle mobilise plus de compétences que la réalisation de *reporting* et que la transmission des données. L'analyse d'un site sous tous ses aspects et la transmission de ces analyses vont permettre d'y distinguer les trois types d'élaboration du savoir au sens de François-Marie GERARD, Expert en éducation, en formation et en gestion de projets, au Bureau d'Ingénierie en Éducation et en Formation (BIEF) :

L'ensemble des nouvelles connaissances liées au savoir transmis par les différents collaborateurs de l'agence et certains documents faisant office de tutoriels permet de constituer **Le savoir reproduire**. Dans ce type de savoir, nous pouvons citer par exemple l'utilisation des techniques d'audit permettant de transposer les informations contenues sur le document en un savoir utile à l'entreprise. Nous pouvons aussi y intégrer le savoir lié aux différentes formations destinées à découvrir de nouvelles procédures d'investigation au sein d'un site. En effet, certaines connaissances acquises permettent simplement la réalisation d'analyse de site au travers de la reproduction d'une même méthodologie sur des données ou des dispositifs différents.

Le second type de savoir est le **savoir-faire.** Ce type de savoir qui est essentiel pour intégrer la notion d'analogie. Une analogie est un processus mental permet d'associer un savoir acquis, plus précisément, un savoir reproduire avec une situation différente et non pas uniquement des données. Ces situations différentes présentant des caractéristiques similaires mais le champ d'application ne sera pas du tout le même. Ce savoir-faire pourra par exemple intégrer la maîtrise des outils d'analyse d'audience. Cette

maîtrise sera utile à l'investigation dans les données et à la création de nouvelles méthodologies d'extraction d'information permettant d'obtenir des solutions à des problèmes qui étaient auparavant insolubles. Ce savoir-faire pourra ensuite alimenter le capital du savoir reproduire car la procédure d'analyse n'aura (théoriquement) plus à être développé de nouveau.

Enfin, l'adaptation constante aux nouvelles problématiques et aux nouveaux types de clients permet de se constituer **le savoir-être.** Ce dernier savoir est essentiel dans l'activité globale d'analyse numérique, de l'analyse des dispositifs jusqu'à la transmission du savoir en général. Ce type de savoir régit l'ensemble des différents savoirs développés précédemment. Le savoir-être désigne la façon de réagir et la façon de s'adapter face à de nouvelles situations. Ce savoir très large permettra de connaître plus précisément ses propres connaissances tout en permettant de déterminer le type de savoir doit être mobilisé pour réaliser une investigation précise. Ce savoir très important permettra aussi la remise en question afin de constituer de nouvelles connaissances. Ainsi, l'analyste saura qu'il n'a pas les capacités pour réaliser ce type d'analyse puis, décidera de constituer un nouveau savoir reproduire ou savoir faire pour être à même de répondre efficacement aux différentes problématiques rencontrées. Grâce à ce savoir-être, il pourra être aussi capable de communiquer avec des profils très différents pour optimiser la transmission d'informations.

Pouvoir ainsi catégoriser les différents types de savoirs permet d'optimiser le processus de production des connaissances et facilite leurs assimilations. Les savoirs nécessaires à l'activité d'analyse ne sont pas uniquement liés à la maîtrise des différents outils. Dans l'activité d'analyse numérique, il est important de pouvoir faire face à l'évolution du web, et aux nouvelles attentes des différentes entreprises. Ainsi, les trois types de savoir présentés sont très importants car ils forment de manière égale l'expérience du consultant.

Malgré tout et pour que ces savoirs soient facilement réutilisables, il est important de pouvoir capitaliser sur l'ensemble des enseignements.

2.b Processus de capitalisation

Développer les différents types de savoir constitue une activité essentielle permettant l'évolution des différents professionnels de l'entreprise qui détient l'expertise. Cette activité est alors essentielle pour la performance de l'entreprise et son attractivité. Ce développement permettra de proposer des prestations de valeur tout en minimisant l'impact de la concurrence sur les activités de manière globale. Malheureusement, il existe beaucoup de freins au développement des connaissances au sein d'une même entreprise :

- Certains savoirs ne sont maîtrisés que par une infime partie des professionnels opérant dans leurs secteurs. Deux consultants disposeront de connaissances et de savoir-faire individuels. Dans le cas d'une retraite par exemple où, un des consultants serait amené à quitter l'entreprise, le savoir acquis par ce consultant serait théoriquement perdu pour l'entreprise.

- Lorsqu'un nouvel employé arrive à un poste, il est utile pour l'entreprise de pouvoir le former rapidement afin qu'il puisse être opérationnel le plus tôt possible. Plus les connaissances et les outils nécessaires à l'activité sont nombreux, plus la formation risque d'être difficile. Difficile à la fois, dans le cadre de l'organisation de la transmission des connaissances mais aussi, dans le cadre l'assimilation des différents savoirs.

- Plus les pratiques mobilisent des connaissances complexes, plus les risques de faire des erreurs sont importants. De plus, les différents types de savoir ne sont pas toujours parfaitement acquis et si les connaissances ne sont pas réutilisées durant un certain temps, la probabilité pour que ces connaissances soient oubliées croît.

Ces différents arguments plaident en faveur d'une pratique de plus en plus stratégique et importante pour les entreprises : la capitalisation sur les connaissances. Cette activité permet de répondre en partie aux différentes problématiques exposées ci-dessus. La capitalisation sur les connaissances est une activité intellectuelle permettant la définition, la traduction et l'organisation des connaissances en vue d'une réutilisation ultérieure par toute personne pour qui ces connaissances pourraient être utiles.

Au sein d'altima°, les différents professionnels n'hésitent pas à partager leurs connaissances. De plus, sur les serveurs de l'entreprise, toutes les productions réalisées

pour les clients sont stockées et organisées au mieux pour former une base de connaissances solide. Par contre, il existe trop peu de documents permettant d'améliorer les savoirs faires et les pratiques des différents acteurs au regard justement, des nombreux savoirs faire présents. Par exemple, il n'existe pas de documents permettant d'apprendre à utiliser les divers outils d'analyse d'audience ou de documents permettant de réaliser un type d'analyse spécifique.

Il serait intéressant pour une entreprise de cette envergure de disposer d'une base de connaissances permettant aux différents collaborateurs d'accéder à de nouvelles pratiques d'analyses, d'apprendre à utiliser un nouvel outil ou tout simplement de développer de nouvelles compétences. Cette pratique est tout de même difficile à mettre en place car elle demande un investissement important à la fois sur le plan des ressources humaines spécialisées dans la documentation et à la fois au niveau économique.

Même si la capitalisation collective est encore en développement au sein de l'agence, un travail de capitalisation individuel a été mis en place pour rassembler le maximum d'informations permettant d'être plus productif dans cette activité. Ce travail de capitalisation individuelle facilite simplement l'accès à toute information pouvant être utile pour l'activité. Elle permet aussi d'organiser les fichiers, les documents de manière stratégique afin d'économiser un temps précieux lors de l'élaboration et de la recherche des documents à destination des entreprises bénéficiaires des diagnostics.

Après le travail sur l'extraction automatique des données, le regard va surement se réorienter sur l'organisation d'une capitalisation sur les connaissances collectives de l'activité d'analyse. Tout d'abord pour que la charge cognitive nécessaire à l'activité de l'équipe soit minimisée et ensuite pour que l'organisation du savoir se fasse plus rigoureusement. En prenant un exemple simple, certains consultants rencontrent des certaines difficultés dans l'utilisation des expressions régulières. Les expressions régulières sont générées à partir d'un langage informatique qui permet de réaliser des recherches dans des fichiers texte mais l'élaboration d'une expression peut-être compliquée pour les personnes n'ayant pas l'habitude du langage informations. Ces expressions sont par moment très difficiles à lire et demandent beaucoup d'énergie pour être générées par un consultant plus orienté marketing. Pourtant l'utilisation de ce langage est obligatoire lorsque le consultant veut faire des recherches sur une URL[19] précise dans les outils d'analyse d'audience.

[19] Uniform Resource Locator : Désigne l'adresse d'un site ou d'une page d'un site web.

Le principe a été de profiter d'un background en informatique fort afin maîtriser ce langage rapidement dans son intégralité. Ensuite, cette maîtrise a permis de produite un document synthétisant l'ensemble des cas rencontrés lors de la recherche d'une URL précise sur un site. Cet exemple représente une activité de capitalisation. Ce document permet de stocker le savoir acquis sur les expressions régulières tout en le partageant. Le savoir contenu pourra ensuite être utile à toute l'équipe même si les compétences ne sont plus dans l'entreprise.

III.3 Gestion du temps liée au travail d'analyse

3.a Définition des différentes activités

Sans prendre en compte la qualité des réponses apportées à un problème donné en fonction des différents types de savoir impliqués, la principale différence entre un analyste confirmé et un analyste débutant se situe dans la rapidité avec laquelle le professionnel est capable d'analyser puis de proposer une solution adaptée.

La problématique du temps dans cette activité s'est posée après avoir réalisé les premières tâches durant le stage. Même avec une bonne expérience, beaucoup d'analyses sont réalisées à partir d'un travail d'investigation. Peu importe la méthode utilisée, le processus d'obtention de certains résultats est lié à un facteur aléatoire qui peut être très gênant. Il n'est pas possible de connaître à l'avance le site, les différents problèmes et comment nous allons pouvoir transposer ces problèmes en préconisations.

La question du temps a suscité une appétence particulière tout au long du stage et ce, dès la réalisation du premier diagnostic critique. Alors qu'un analyste confirmé est censé proposer l'analyse du site tout en rédigeant un document structuré permettant de transmettre le savoir en l'espace de quarante heures, le premier diagnostic réalisé en autonomie pour le site Le Jardin de Catherine a été achevé en plus de 120 heures. Ce constat plutôt négatif a tout de même un intérêt car nous allons pouvoir analyser objectivement les différentes étapes de l'analyse et celles qui prennent le plus de temps :

- Les outils d'analyse d'audience et les fonctionnalités avancées des tableurs sont difficiles à maîtriser. De plus, l'apprentissage de ce type d'outil se fait parallèlement au travail d'analyse. Il est alors naturel que les analyses prennent plus de temps.

- L'investigation dans le site et dans les données, activité reposant sur un facteur aléatoire fort.
- L'élaboration d'un fichier d'extraction de données réalisé parallèlement à l'analyse du site a considérablement augmenté les temps d'analyse. Une réflexion a été menée sur la présence des différentes données nécessaires à la rédaction du cahier des charges.
- Visionnage des tests utilisateurs avec Hotjar, outil permettant d'effectuer un dérivé de test utilisateurs.
- L'apprentissage de la communication. Le vocabulaire utilisé en sciences de l'information et de la communication est radicalement différent du vocabulaire utilisé dans le domaine du marketing. Un temps supplémentaire est nécessaire pour adapter le discours vers un public totalement nouveau et développer ses compétences en communication.
- Les présentations à l'intention des entreprises doivent idéalement suivre une structure précise et la charte graphique de l'entreprise afin de respecter son identité.
- Des facteurs exceptionnels sont à prendre en compte lors du travail avec des clients. Certaines informations ne sont pas transmises à temps et certaines de ces informations modifient la manière dont les analyses doivent s'effectuer. Par exemple, dans le cadre de comparaison de la performance d'un site entre deux périodes, il n'est pas possible de comparer deux périodes ayant profité d'offres différentes. Si une analyse est effectuée entre une période de solde et une période standard, il est normal de trouver des différences importantes. Si le client fournit ces informations après que l'analyse des données ai été faite, ces analyses devront être rectifiées ou même supprimées.
- Le travail de veille qui sera effectué pour mettre à jour ses connaissances et proposer des solutions innovantes.
- La configuration des outils.

Ces différentes étapes sont ensuite catégorisées en fonction de la nature des besoins en matière de temps. Plus simplement, nous allons séparer les différentes étapes entre celles qui nécessitent un temps défini comme l'apprentissage des outils et qui nécessitent un temps indéfini comme le travail d'investigation :

Défini	Non-défini
-Apprentissage des outils -Mise en forme de la présentation -Configuration des outils -Elaboration et utilisation du fichier d'extraction de données	-Investigation -Visionnage des tests utilisateurs -Communication -Facteurs exceptionnels -Travail de veille

L'idée est de minimiser le temps pris par les étapes dont le temps est défini. Certaines d'entres elles prendront de moins en moins de temps puisqu'elles concernent l'apprentissage de l'outil. D'autres peuvent être automatisées car elles ne nécessitent pas de travail de réflexion. Il est important d'optimiser le temps pris sur ces étapes qui n'impactent pas directement la stratégie des entreprises à l'aide des outils adaptés. Il sera ensuite possible d'orienter le temps restant vers des actions qui prennent un temps théoriquement indéfini et qui permettent d'apporter le plus de valeur aux présentations réalisées.

3.b Optimisation du temps des différentes activités

Un chapitre a déjà été réalisé sur le programme d'extraction automatique des données mais ce cas illustre parfaitement les différentes actions mises en place permettant d'optimiser le temps lors du travail d'analyse. Ces outils permettent d'automatiser les tâches redondantes afin de réduire les différents temps d'exécution en économisant de l'énergie. Ce choix d'outils permet de rendre les analyses plus fiables d'un point de vu théorique.

Pour chaque étape, un travail d'analyse a ensuite été mené permettant de choisir les stratégies et les outils adaptés permettant d'obtenir le meilleur rendement sur les différentes activités menées dans le cadre des analyses. Il n'y a pas de méthodologie-type d'optimisation du temps d'une activité car chaque activité possède une manière d'optimiser son temps qui lui est propre.

Concernant l'apprentissage des outils, l'optimisation du temps peut être facilitée par le choix pertinent des sources facilitant l'apprentissage. Un travail de veille plus efficace permet donc d'optimiser le temps mobilisé par cette activité.

Pour rendre l'activité de veille plus efficace, Il est nécessaire de sélectionner les sites puis de les visiter régulièrement être au courant des nouveautés. Un agrégateur de flux permet de réaliser l'extraction d'informations plus facilement. Plutôt que de se souvenir des sites sur lesquelles la veille s'effectue, un agrégateur va optimiser l'activité de veille par l'ajout de fonctionnalités très pratiques. Feedly a été utilisé en tant qu'agrégateur de contenu. Cette application en ligne permet une double action. Tout d'abord, celle de rassembler les sources utiles à la veille selon des catégories définies. Ensuite, cet outil permet d'extraire et d'accéder rapidement aux différentes informations extraites sur les sites. Cet outil est assez puissant car il permet d'agréger des informations ne relevant pas uniquement des **flux RSS**[20] mais aussi des liens directs vers des billets de blogue, ou des articles. Enfin, cet outil propose automatiquement des sites populaires correspondant aux sites sélectionnés pour élargir les recherches.

Pour la mise en forme des informations dans les présentations, un fichier de base a été réalisé permettant de concentrer toutes les structures types de diapositives que nous trouvons dans une présentation. Ce fichier permet de définir différentes architectures de diapositives normalisées permettant de faciliter la mise en forme des informations à l'intérieur des présentations. Le fichier d'extraction de données permet aussi de générer automatiquement les graphiques standards des analyses.

Le temps mobilisé pour la configuration des différents outils d'analyse d'audience est optimisé par le travail de capitalisation et par le document permettant de réaliser l'audit de tracking. Ces documents permettent de s'assurer qu'il n'y aura pas d'erreur lors de la configuration et fournissent toutes les informations nécessaires à une configuration rapide.

Il n'est pas possible d'optimiser le temps pris par toutes les étapes du travail d'analyse au travers d'outils. Pour optimiser le temps pris par une analyse en général, nous allons encore devoir faire appel à une méthode pour organiser les différentes étapes afin de les organiser et de faire en sorte que chaque étape puisse servir la suivante.

[20] Rich Site Summary (R.S.S.) : Fichiers permettant d'obtenir rapidement un sommaire du contenu des sites afin de les consulter rapidement.

3.c Procédure d'optimisation du temps

Dans cette course contre la montre, ne pas mettre en place de stratégie permettant d'organiser et de structurer l'analyse risque d'avoir un impact majeur sur la qualité des connaissances générées. Sensibilisé par les problématiques en informatique, l'objectif est de dérouler un algorithme permettant d'effectuer les tâches de manière cohérente pour en retirer le maximum d'informations tout en permettant de prendre en compte des variables inconnues.

Cette recherche d'une procédure idéale d'organisation scientifique du travail doit permettre d'organiser les différentes étapes afin d'obtenir le maximum de temps d'analyse et d'amélioration des compétences. Cette organisation doit aussi permettre à chaque étape de générer des informations qui pourront servir l'étape suivante et la rendre plus rapide. Cette méthodologie est en cours d'élaboration, puisque certaines connaissances relatives à ce métier doivent encore être solidifiées.

Nous allons donc présenter une partie de l'organisation des étapes appliquées à un site qui suit cette méthodologie. Les choix sont justifiés, mais restent bien sûr discutables étant donné le caractère inachevé de la réflexion.

1- Préparation des dossiers et fichiers permettant de réaliser l'analyse: structure des dossiers pour organiser les captures d'écran, les tableurs, et tout document permettant de réaliser les analyses. Cette étape est effectuée à partir d'un copier-coller d'un dossier-type et prends 1 minute.

2- Tout d'abord, il est important de s'assurer que les données sur le site sont fiables. L'audit de tracking est alors réalisé. Cette étape doit prendre au maximum 30 minutes.

3- Si l'application a été installée sur le site, Hotjar est utilisé. Pour découvrir le site au travers des yeux de l'utilisateur, nous sélectionnons 5 vidéos d'environ 10 minutes pouvant illustrer une navigation type. Cette étape devrait prendre une heure et les principaux problèmes de navigation sont consignés directement sur le document de présentation.

4- Phase d'exploration du site afin d'obtenir les URL des différents types de page. Ces URL permettront de configurer le fichier d'extraction automatique des données afin d'obtenir les données quantitatives des différentes pages. Cette phase d'exploration permet aussi d'obtenir les premiers éléments d'une méthodologie d'analyse descendante. La rédaction de ces observations se fait aussi sur le document de présentation. En fonction de la structure des sites et des objectifs d'analyse, cette étape possède une difficulté variable et peut prendre entre 2 heures et 4 heures.

5- Configuration du fichier d'extraction de données d'une durée de 40 minutes

6- Phase d'extraction de données qui prend entre 1 et 2 heures. Cette phase étant entièrement automatisée, elle est effectuée parallèlement à la réalisation de l'audit technique et la poursuite de la méthodologie d'analyse descendante.

7- Intégration des graphiques types réalisés grâce au fichier d'extraction de données dans le document de présentation nécessitant 10 minutes.

Cette première partie de la méthodologie permet de faciliter grandement le travail d'analyse. Le temps mobilisé pour réaliser toutes les étapes se situe entre 6h20 et 8h20. La réalisation de ces étapes va permettre d'obtenir une grande partie des éléments nécessaires aux différentes analyses dans un temps plus court que si les différentes étapes n'avaient pas été organisées de la sorte. Cela va permettre au consultant de se concentrer par la suite exclusivement sur les analyses qualitatives.

Par la suite, la recherche de la bonne méthodologie s'orientera vers les différentes analyses (ascendantes, descendantes, Hotjar, benchmark) pour réfléchir à une organisation optimale. L'application de cette méthodologie à ces étapes risque tout de même d'être plus difficile car les informations extraites par ces différentes analyses sont liées à une variable temporelle aléatoire.

Conclusion de partie III

Cette dernière partie nous permet d'obtenir une vision plus globale et décentrée du travail d'analyse numérique réalisé le consultant. Cette vision est nécessaire car même si elle ne représente pas directement le travail d'analyse, elle permet d'appréhender l'activité

dans la globalité et représenter la diversité des pratiques que nous pouvons retrouver au sein de ce métier. Enfin, il est important de détailler toutes les activités périphériques destinées à faciliter le pilotage des activités numériques dans le cadre d'une analyse de la circulation des informations.

Cette vision plus générale permettant de visualiser les différentes activités a un impact direct sur la stratégie des entreprises seulement lorsqu'il est question de la transmission des informations. Malgré tout, la capitalisation et l'organisation du temps vont être deux facteurs indirects qui vont avoir une incidence sur la qualité des analyses. Au final, les entreprises profiteront en grande partie des différentes innovations apportées par ces activités ce qui leurs permettra d'orienter toujours précisément leurs stratégies.

CONCLUSION & OUVERTURE

Depuis que les entreprises ont la possibilité d'être présente sur le web, ces entreprises se trouvent confrontées à une multitude de problématiques. En effet, une stratégie de présence en ligne implique de nombreux besoins en matière de compétences mais aussi en matière de compréhension de l'univers numérique. Une présence sur le web efficace exige d'une part la maîtrise des différentes technologies nécessaires au fonctionnement optimal des dispositifs et d'autre part, la compréhension des comportements, besoins et attentes des usagers. De plus, le web évolue et se transforme plus rapidement que n'importe quel média et oblige les entreprises à s'adapter sans cesse pour rester attractives face à la concurrence. Dans ce cadre, nous avons voulu constituer un document permettant d'obtenir une vision générale de l'activité d'analyse.

Tout au long de ce mémoire, nous avons essayé de transmettre une vision objective de l'analyse des activités en ligne et de la circulation de l'information. L'activité du digital analyst est théoriquement liée à l'analyse et l'interprétation des données quantitatives. Pourtant, nous avons désiré pousser la réflexion en mettant en avant d'autres types d'analyses puisque ce consultant ne se contente pas uniquement de réaliser les analyses sur ce type de données. Il a paru important que ce mémoire reflète l'activité étudiée dans sa globalité.

Chaque consultant *digital analyst* présent dans l'agence possède les compétences permettant de réaliser les analyses à partir des données quantitatives. Pourtant, chaque analyste possède des compétences qui surpassent ce besoin de base formant ainsi une équipe pluri compétente et complémentaire. Cette équipe possède des compétences très diverses lui permettant d'être à l'aise en ergonomie, en communication, en programmation ou encore en marketing. Ces compétences supplémentaires permettent de compléter les analyses et permettent ainsi détecter plus précisément les différents problèmes afin d'orienter le plus efficacement la stratégie des entreprises.

Ainsi, nous avons pu obtenir une vision plus précise du parcours de l'information dans le cadre son activité partant des différentes analyses jusqu'à la constitution du savoir global. Le spectre des compétences mobilisé par le digital analyste est au final assez large et c'est pour cette raison que cet acteur est un interlocuteur particulier des entreprises car ses connaissances et ses conseils permettent d'orienter de manière pragmatique leurs différentes stratégies.

Malgré tout, vouloir traiter de l'analyse des activités en ligne au sens large était un risque étant donné la multitude des connaissances nécessaires pour traiter la problématique dans son intégralité. Ce risque était connu et a été pris délibérément puisqu'il est important d'apporter des éléments de réponse permettant de constituer une idée de ce qu'est l'analyse digitale. Pour faciliter l'élaboration de ce document et ne pas se perdre dans la diversité des éléments à prendre en compte, nous avons limité les recherches en fonction de l'activité du *digital analyst*. Nous avons ensuite structuré les recherches en fonction du type d'information générée et de sa circulation. Ceci nous a permis d'organiser plus facilement la réflexion sur le sujet afin de proposer une étude la cohérente permettant de faciliter la lecture.

Ce mémoire n'a donc pas eu pour objet de réaliser une réflexion cloisonnée sur un sujet en particulier de l'analyse numérique. Il est à notre essentiel de réaliser une étude permettant de mettre en avant la diversité des informations traitées qui ont toutes un impact sur l'orientation des stratégies numériques. Dans le sens où le l'activité d'analyse de dispositifs numériques a trop souvent été étudiée sous des aspects spécifique, il était important de réaliser une autre étude plus générale en associant des points de vus tout aussi variés que les analyses pouvant être faites sur les sites. Ainsi, nous avons impliqué :

- Les sciences de l'information et de la communication qui ont permis de définir ce qu'est l'information au sens large et la diversité de ses parcours.
- L'ingénierie documentaire qui a permis de définir plus précisément la forme que prennent ces informations et les différents dispositifs utilisés permettant rendre la transmission du savoir plus efficace tout en favorisant son archivage.
- L'informatique qui a permis d'apporter un regard différent en essayant de lever l'ambiguïté des différentes étapes de l'analyse tout en essayant de les organiser pour en simplifier la compréhension.

Il serait intéressant de compléter ce document grâce aux méthodes d'analyse empruntées à d'autres métiers pour obtenir une vision plus complète des différentes analyses réalisées sur les sites. Des connaissances en acquisition de trafic, en référencement naturel ou encore en communication interculturelle, permettraient à ce document d'être plus complet et intéressant. Intéressant, pour les personnes désirant connaître les différentes façons d'analyser un site et intéressant, pour les entreprises qui pourraient connaître les différents besoins d'un site et orienter leurs stratégies plus efficacement.

BIBLIOGRAPHIE

Livres :

Arson, B. (2012). *Web analytics*. Paris: Pearson.

Kaushik, A. (2010). *Web analytics 2.0*. Indianapolis, IN: Wiley.

Foreman, J. (2013). *Data Smart*. Somerset: Wiley.

Bastien, J. (1993). *Ergonomic criteria for the evaluation of human-computer interfaces*. Rocquencourt: Institut national de la recherche en informatique et en automatique.

Daumal, S. (2012). *Design d'expérience utilisateur*. Paris: Eyrolles.

Revues :

Smart Insights. (2016). *Ecommerce analytics - Smart Insights Digital Marketing Advice*.

Lemoine, J. (2012). À la recherche d'une meilleure compréhension du comportement des internautes. *Management & Avenir*, 58(8), p.116.

Gerard, F.-M. (2000). Savoir, oui... mais encore !, *Forum – pédagogies*, mai 2000, 29,35

Kalika, H. (2016). Surcharge informationnelle, urgence et TIC. l'effet temporel des technologies de l'information. *Management & Avenir*, [online] n° 13(3), pp.149-168.

WEBOGRAPHIE

Smart Insights. (2016). *Ecommerce analytics - Smart Insights Digital Marketing Advice*. [online]
Consultable à l'adresse : http://www.smartinsights.com/ecommerce/ecommerce-analytics/ [Version du 2 Juillet 2016].

Bricebottegal.com. (2016). *Définition et histoire du digital analytics | Blog digital analytics*. [online] Consultable à l'adresse : http://www.bricebottegal.com/definition-histoire-web-analytics/#7 [Version du 20 Juillet 2016].

Occam's Razor by Avinash Kaushik. (2008). *Six Web Metrics / KPI's To Die For*. [online] Consultable à l'adresse : http://www.kaushik.net/avinash/rules-choosing-web-analytics-key-performance-indicators/ [Version du 2 Juin 2016].

Analyticsacademy.withgoogle.com. (2016). *Analytics Academy*. [online] Consultable à l'adresse : https://analyticsacademy.withgoogle.com/course/1/unit/3/lesson/2 [Version du 15 Août 2016].

Opquast.com. (2016). *La marque des professionnels du Web - Opquast*. [online] Consultable à l'adresse : http://opquast.com/fr/ [Version du 30 Juin 2016].

Studio by UXPin. (2016). *e-Books*. [online] Consultable à l'adresse : https://www.uxpin.com/studio/ebooks/ [Version du 18 Juillet 2016].

Enterprise Ecommerce Blog - Enterprise Business Marketing, News, Tips & More. (2016). *How to Perform a Google Analytics Audit for Your Ecommerce Site – Shopify*. [online] Consultable à l'adresse : https://www.shopify.com/enterprise/101158918-how-to-perform-a-google-analytics-audit-for-your-ecommerce-site [Version du 20 août 2016].